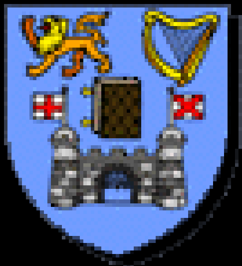
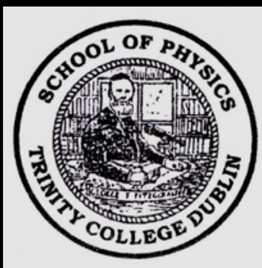
Semiconductor Devices - 2014

Lecture Course
Part of
SS Module PY4P03

Dr. P. Stamenov

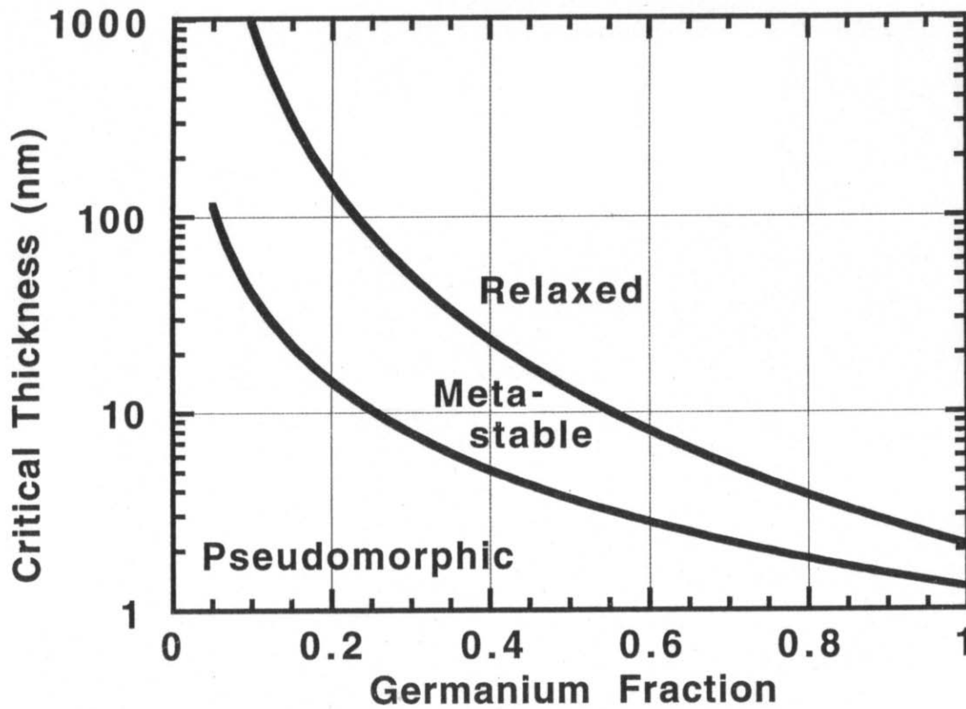
School of Physics and CRANN, Trinity College,
Dublin 2, Ireland

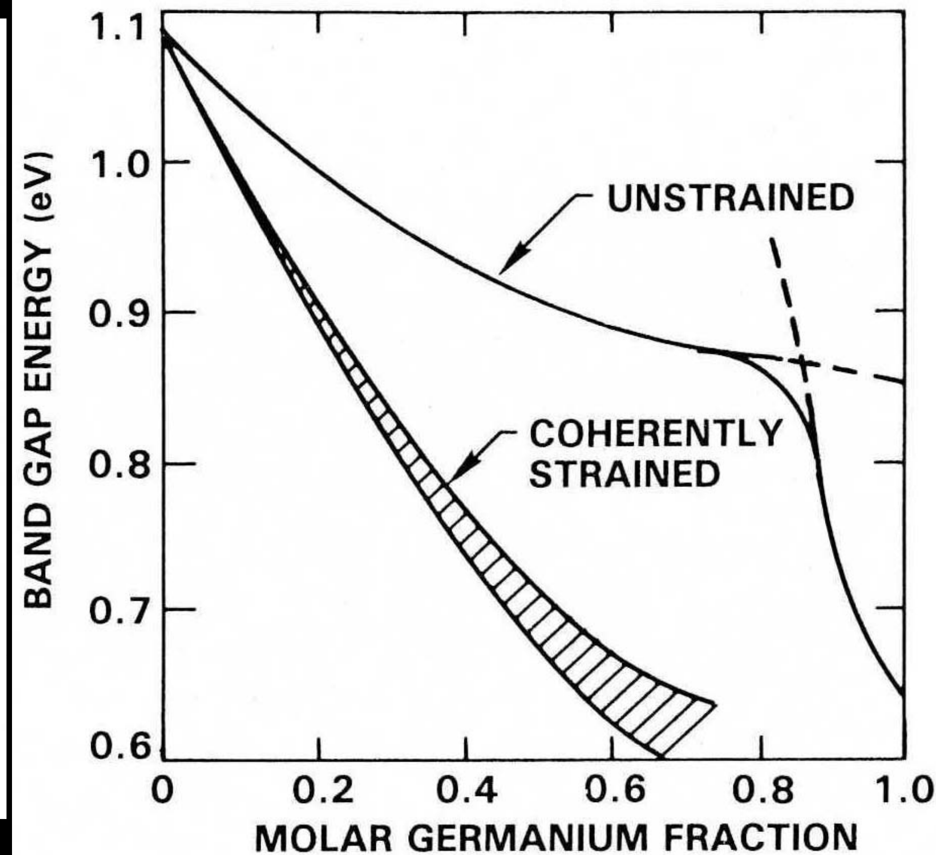




Hilary Term, TCD
17th of February '14

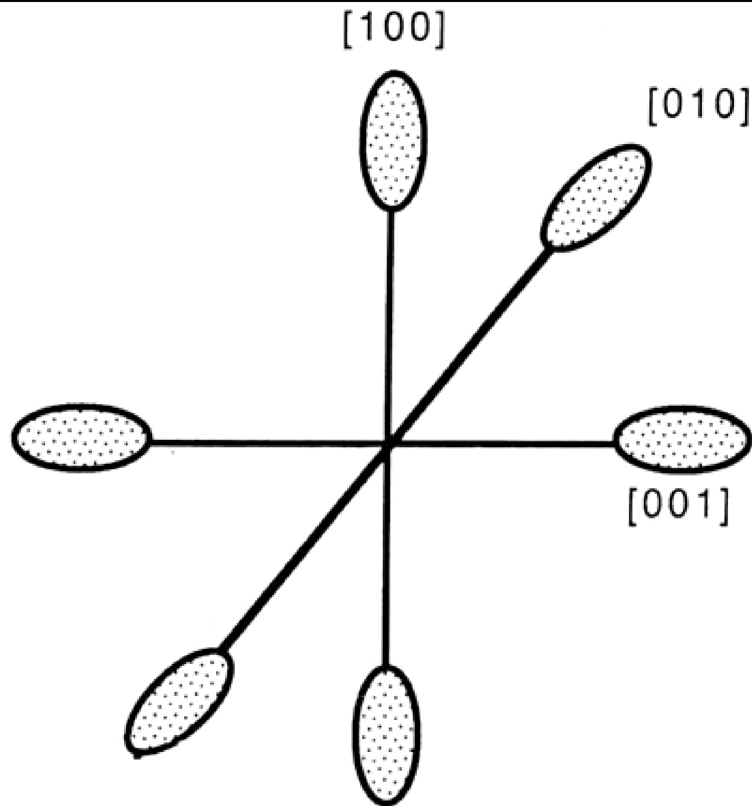
Strain Control - $Si_{1-x}Ge_x$

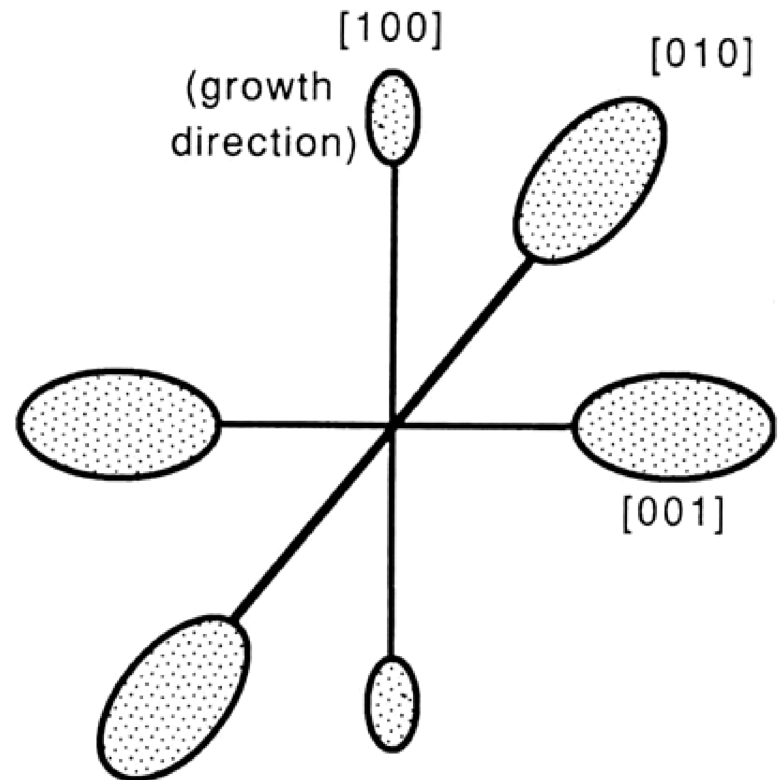




- $Si_{1-x}Ge_x$ is attractive, as it can be processed without large alternations by silicon device technology fabs.
- The ‘chemical pressure’ relies on the fact that the Ge lattice has ~4 % larger constant. It alters the band-gap and other important ‘constants’.

Constant Energy Ellipsoids and Strain



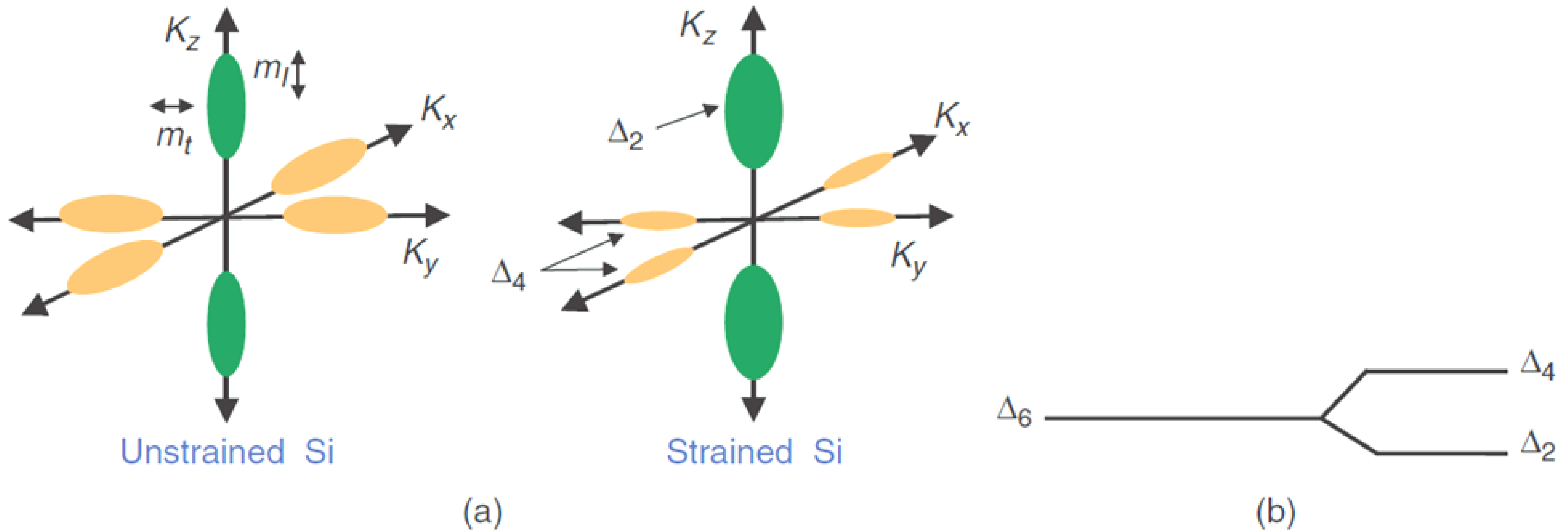


- Strain lifts the degeneracy of the in-plane and out-of-plane longitudinal and transverse effective masses.
- For SiGe on Si – the lattice expands and filling along the growth direction drops, leading to smaller *E* = const ellipsoids along <100>.

Why is Strain Beneficial?

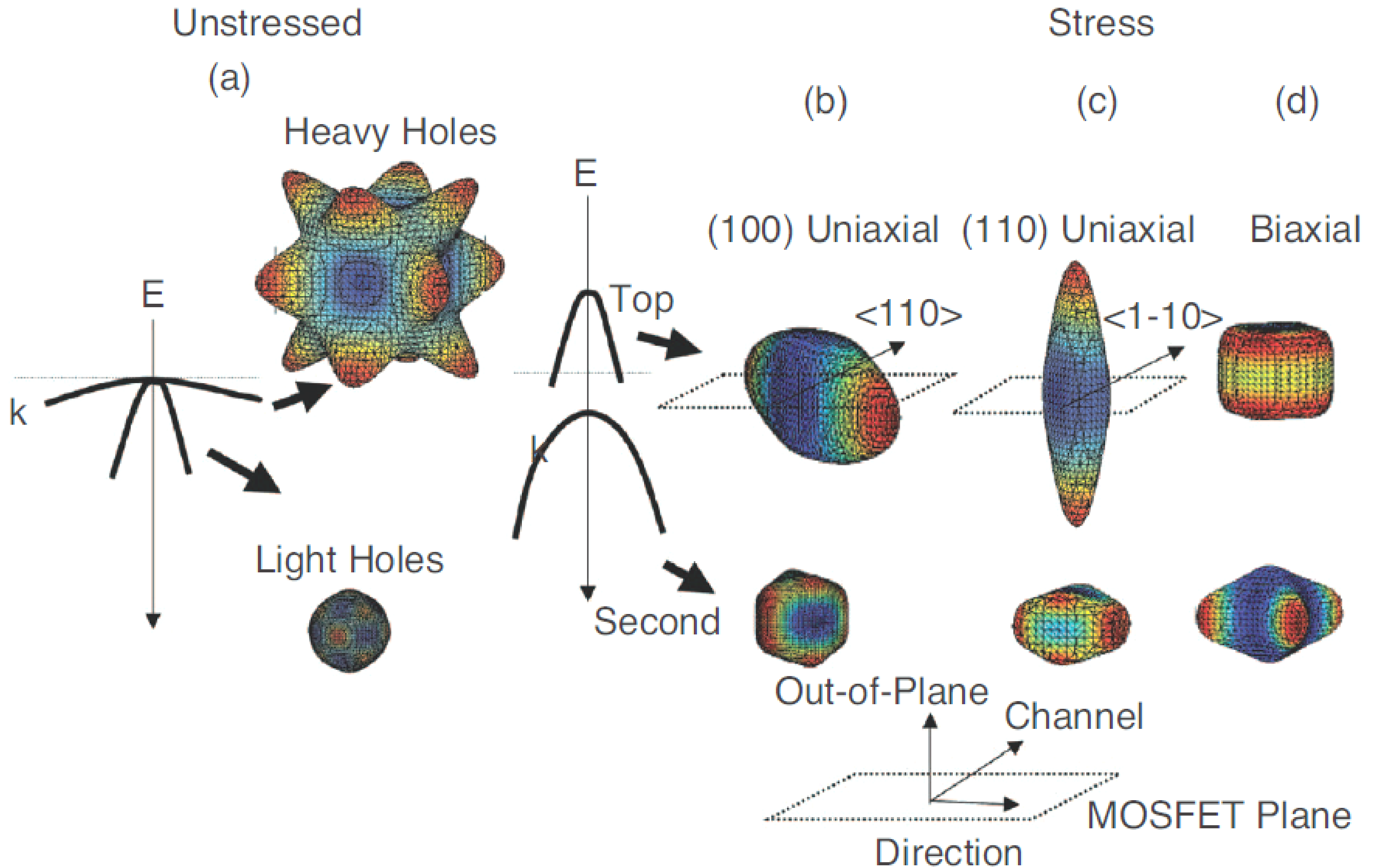
- Not primarily because it alters the size of the ellipsoids, but because it alters the effective mass along the growth direction by up to a factor of about two.
- Correspondingly, epilayer mobility along the growth direction – out-of-plane (perpendicular to the wafer surface) is increased substantially.
- The narrower *k*-vector distributions mean also lower overall electron scattering, i.e. longer electronic lifetimes.
- The carrier (electron) diffusion coefficient is also increased, in agreement with the general Einstein relations.
- For bipolar transistors (structured vertically) the increase in mobility implies lower base transit times (for the *n*-*p*-*n*, or minority electron case), leading to much faster operation.
- Lateral control is also possible by simply bending the wafer or (better) by a careful design of the doping laterally.

Strain on the Quasi-Electrons

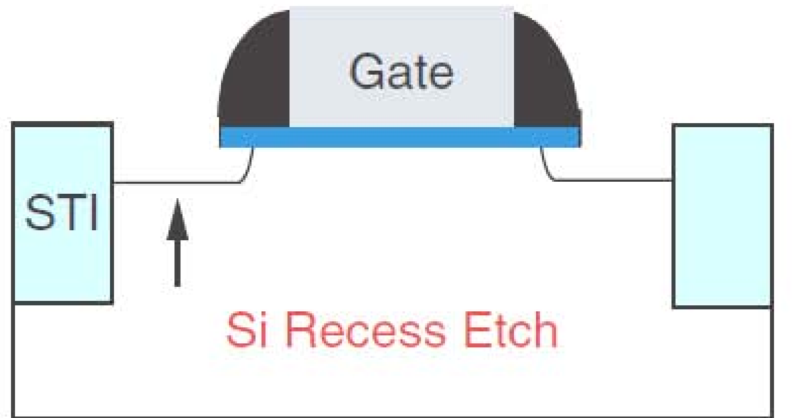


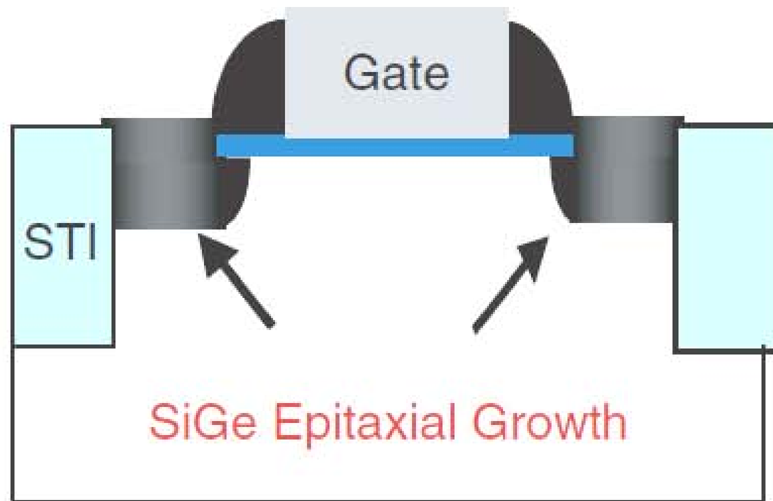
- Control has to be exercised in the plain for FETs. This requires more ingenuity, compared with the vertical devices.
- Just orienting the channels and relying on macroscopic stress in not an option – CMOS capability is necessary.
- These are the pictures for small fillings – may be altered at very high doping concentrations.

Strain on the Holes

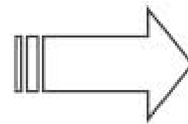


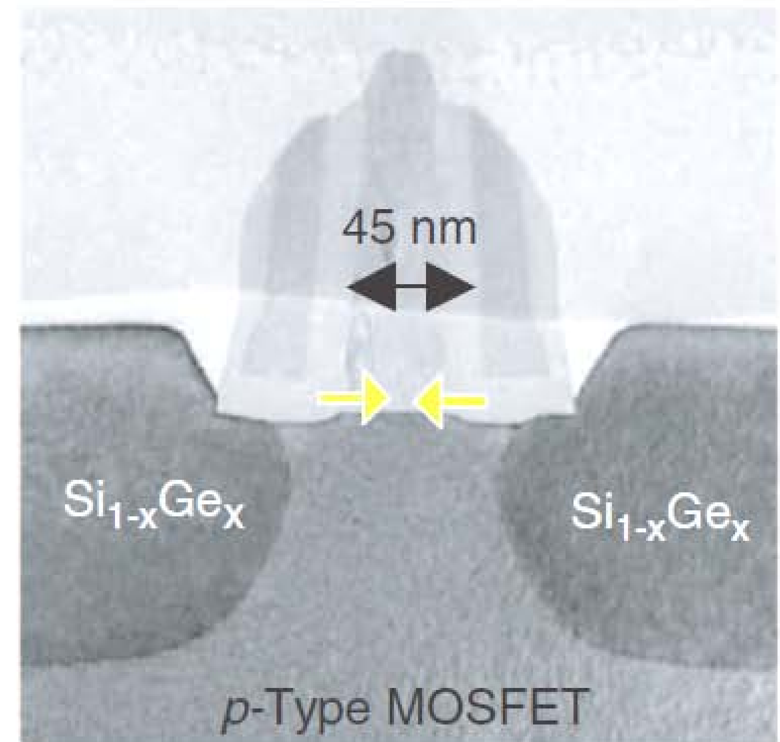
Compressive (SiGe) and Tensile (SiC)





(a)

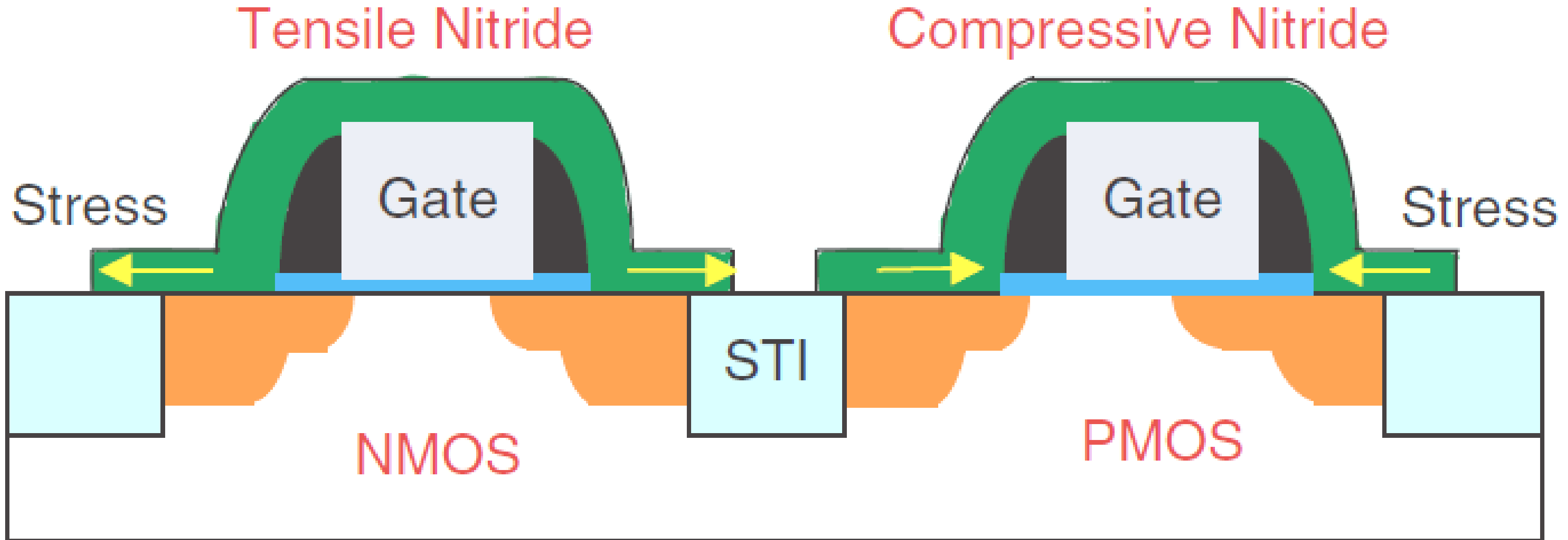




(b)

STI = Shallow Trench Insulation

Use of Tensile or Compressive Nitride



- Nitride (typically used for passivation layers) can also be utilised for strain control in NMOS and PMOS – helping make the scaling of CMOS easier.
- Can also be used as an additional measure in devices relying on SiGe.

Notes on Strain

- Strain-induced in-plane mobility enhancement can be used to improve FET speed.
- The pseudomorphic SiGe MODFET (strained SiGe on Si or *vice-versa*) reaches 50 GHz, 10 times faster than an *n*-MOSFET, yet it can still be made in a silicon fabrication plant.
- For the highest speeds, the field along the channel can become so high that the electron velocity saturates.
- Alternatively the channel lengths are made so low (~100 nm) that the electron transport from source to drain can become 'ballistic', i.e. with no electron scattering.
- In either case the output drain current I_{DSAT} can be now shown to vary (neglecting V_T) not as V_{GS}^2 but as V_{GS}.
- The mutual conductance g_m now becomes independent of V_{GS}, not, as before, $\propto V_{GS}$. However, it is still important to reduce the electron effective mass by straining the channel.

Future Scalability of Strain

$$I_{d(\text{lin})} = \frac{\mu W C_{\text{ox}}}{L}\left(V_g - V_t - \frac{V_d}{2}\right) V_d$$

$$I_{d(\text{sat})} = \frac{\mu W C_{\text{ox}}}{2L}\left(V_g - V_t\right)^2 .$$

The conventional expressions for the drain current and its saturation.

$$I_{d(\text{sat})} = W C_{\text{ox}} < \upsilon(0) > (V_g - V_t)$$

Ballistic Saturation

$$\upsilon_T = \sqrt{\frac{2 k_B T_{\text{L}}}{\pi m_t^*}}$$

Injection Velocity

$$I_d = W C_{\text{ox}} \sqrt{\frac{2 k_B T_{\text{L}}}{\pi m_t^*}} (V_g - V_t)$$

Ballistic Limit

… And hi-k's

- It turns out that with a strained Si channel MOSFET the change in band structure increases the barrier height between the Si channel and gate SiO_2.
- For the thinnest oxides (a few nm), the gate leakage current is due to electron tunnelling. The really fundamental limit is 0.7 nm, below which the SiO_2 bulk band gap is not fully formed.
- Oxide thicknesses < 1.5 nm are precluded because of electron tunnelling and breakdown due to defect accumulation in the oxide.
- The increase in barrier height reduces this current by x 2 or x 3 – a much welcome 'bonus'.
- Need for hi-*k*'s - 'constant field scaling rules', which preserve MOSFET long-channel behaviour, dictate that as the channel length L falls to ~ 100 nm the gate SiO_2 thickness d must be reduced into the nm region.
- Voltages also need to scale down, and doping densities scale up.

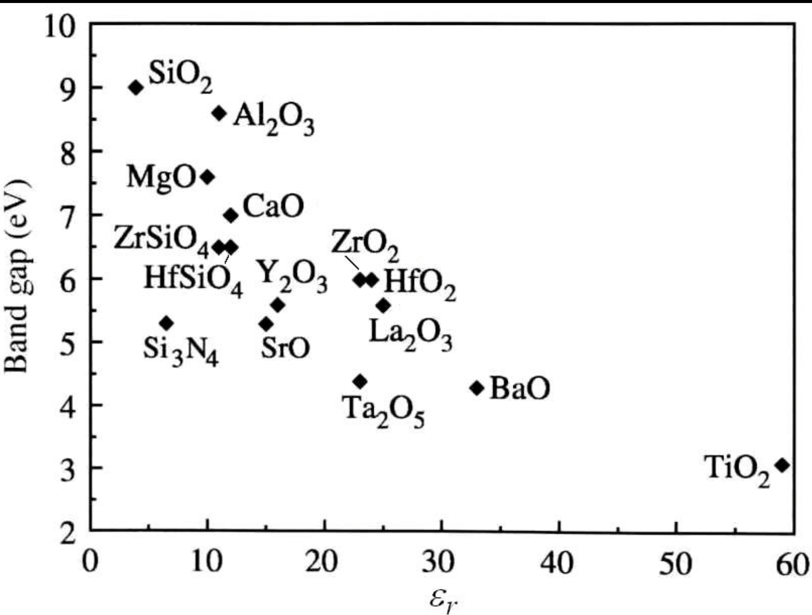
Hi-k and Low-k...

- The expression given earlier for the MOSFET mutual conductance g_m suggests the alternative approach:

$$g_{\mathrm{m}} = \left(\frac{\partial I_{\mathrm{dsat}}}{\partial V_{\mathrm{gs}}} \right)_{V_{\mathrm{ds}}} = \frac{\varepsilon_{\mathrm{r}} \varepsilon_{\mathrm{o}} w \mu_{\mathrm{n}}}{Ld} V_{\mathrm{gs}}$$

- The use of a 'high-*k* dielectric' instead of SiO2 for the gate insulator can therefore keep g_{m} high by raising ε_{r} above the value of 3.9 for SiO_2, without having to reduce *d* to too low values.
- This also keeps the output current I_{d} and the gate capacitance high!
- Since 2007 MOSFETs have been made on a production basis using HfO_2, $HfSiO_4$, ZrO_2 and $ZrSiO_4$ for the gate dielectric. These are in 45 nm-node CMOS microprocessors which have up to almost 10^9 MOSFETs. Polysilicon gates are not used because of the way they would interact with high-*k* dielectrics, and instead they are once again being made of metal. Low-*k* oxides are also used to lower parasitics.

Hi-k's



The figure shows the tendency for a trade-off between band gap and dielectric constant, as plotted for SiO_2 and for 14 candidate dielectrics.

Several requirements must be fulfilled:

- The dielectric must not form thick silicide layers or intrinsic defects at the Si/dielectric interface.
- The channel carrier mobility must not be impaired.
- The energy band gap of the dielectric (9 eV in SiO_2) must be large enough for leakage currents to be low.
- The material must be compatible with CMOS processing requirements.
- The excellent material and electrical properties otherwise possessed by SiO_2 should be maintained as far as possible.

Why low-k_{eff} ?

Characteristic inductance and capacitance per unit length...

$$l_0 := 0.25 \cdot \frac{\mu H}{m} \qquad \mu_0 = 1.257 \frac{\mu H}{m}$$

$$c_0 := 100 \cdot \frac{pF}{m} \qquad \varepsilon_0 = 8.854 \frac{pF}{m}$$

Group velocities Are limited by *c* – Note that the phase velocity it not!

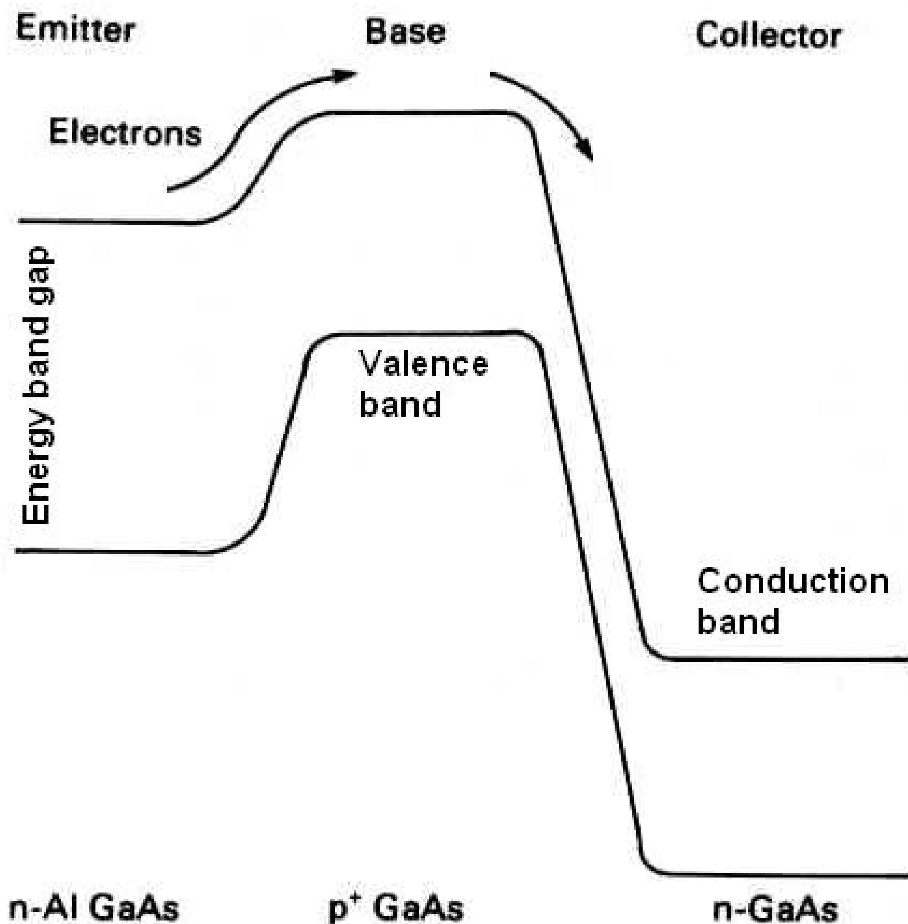
$$v_p := \frac{1}{\sqrt{l_0 \cdot c_0}} = 0.667\,c \qquad \frac{1}{\sqrt{\varepsilon_0 \cdot \mu_0}} = 1\,c$$

$$k_{eff} := 2 \qquad v_{eff} := \frac{c}{\sqrt{k_{eff}}} = 0.707\,c$$

The characteristic impedance – Usually wise to keep it about 50 Ω!

$$Z_0 := \sqrt{\frac{\mu_0}{\varepsilon_0}} = 376.73\,\Omega \qquad Z_{eff} := \sqrt{\frac{l_0}{c_0}} = 50\,\Omega$$

In Bipolar Transistors



- Heterojunctions and strained-layer epitaxy can both greatly improve the performance of bipolar transistors (always *n-p-n* for the fastest performance).
- The heterojunction BPT (HBT) uses a wide bandgap emitter and a base with a bandgap narrower by ~ 0.25 eV.
- This *reduces* the barrier to electrons diffusing from emitter to base, and *increases* the barrier to holes from base to emitter.

Notes on Bipolar Transistors

- The emitter efficiency for minority electron injection into the base is almost ideal (negligible hole current).
- The emitter can be *lightly* doped (reduces base-emitter capacitance) and the base *heavily* doped (reduces base resistance). This is the reverse of the doping requirements in a homojunction BPT.
- The heavily doped base can be very thin, as the Early effect can now be neglected – i.e. the base-collector depletion layer extends into just the collector, so punch-through in the base cannot occur.
- The common emitter current gain can be very high, $\gg 10^3$, as is the frequency response, even though carrier lifetimes in GaAs are short (direct gap semiconductor).
- In a Si 'homojunction' BPT, with emitter doping above 10^{23} m^{-3} the band gap shrinks by ~ 2 or 3 x kT/e. This gives an emitter-base *heterojunction*, but this time with an *increased* barrier to minority carrier injection into the base from the emitter.

Thank You Very Much for Your Attention!