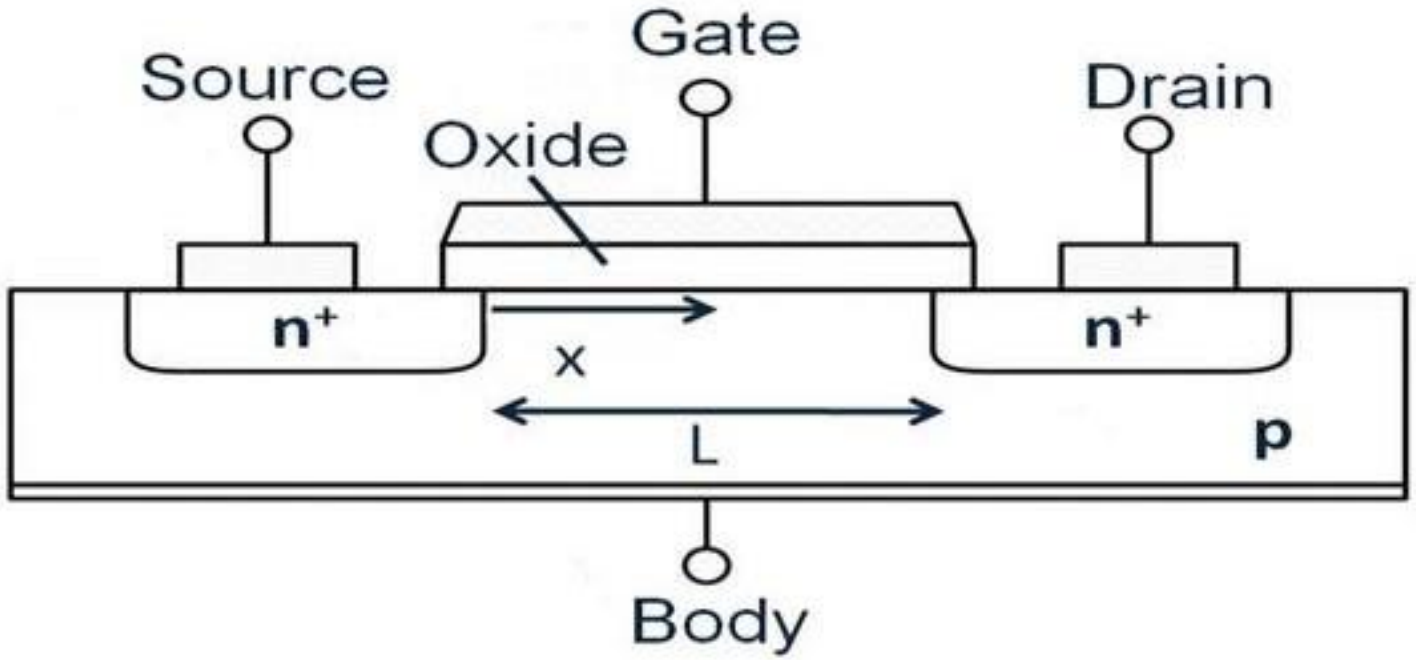
EC 402
NANOELECTRONICS
Module 4

MOSFET

- The MOSFET (Metal Oxide Semiconductor Field Effect Transistor) transistor is a semiconductor device which is widely used for switching and amplifying electronic signals in the electronic devices.
- The MOSFET is a core of integrated circuit and it can be designed and fabricated in a single chip because of very small sizes.
- The MOSFET is a four terminal device with source(S), gate (G), drain (D) and body (B) terminals.
- The body of the MOSFET is frequently connected to the source terminal so making it a three terminal device like field effect transistor.


Gate
Source
Drain
Oxide
n+
n+
x
L
p
Body

- MOS structure is obtained by growing a layer of silicon dioxide (SiO_2) on top of a silicon substrate and depositing a layer of metal or polycrystalline silicon (the latter is commonly used).

- its structure is equivalent to a planar capacitor, with one of the electrodes replaced by a semiconductor.

- The n-type (MOSFET) consists of a **source** and a **drain**, two highly conducting n-type semiconductor regions which are isolated from the p-type **substrate**.

- A metal (or poly-silicone) **gate** covers the region between source and drain, but is separated from the semiconductor by the **gate oxide**.

- The source and drain regions are identical.

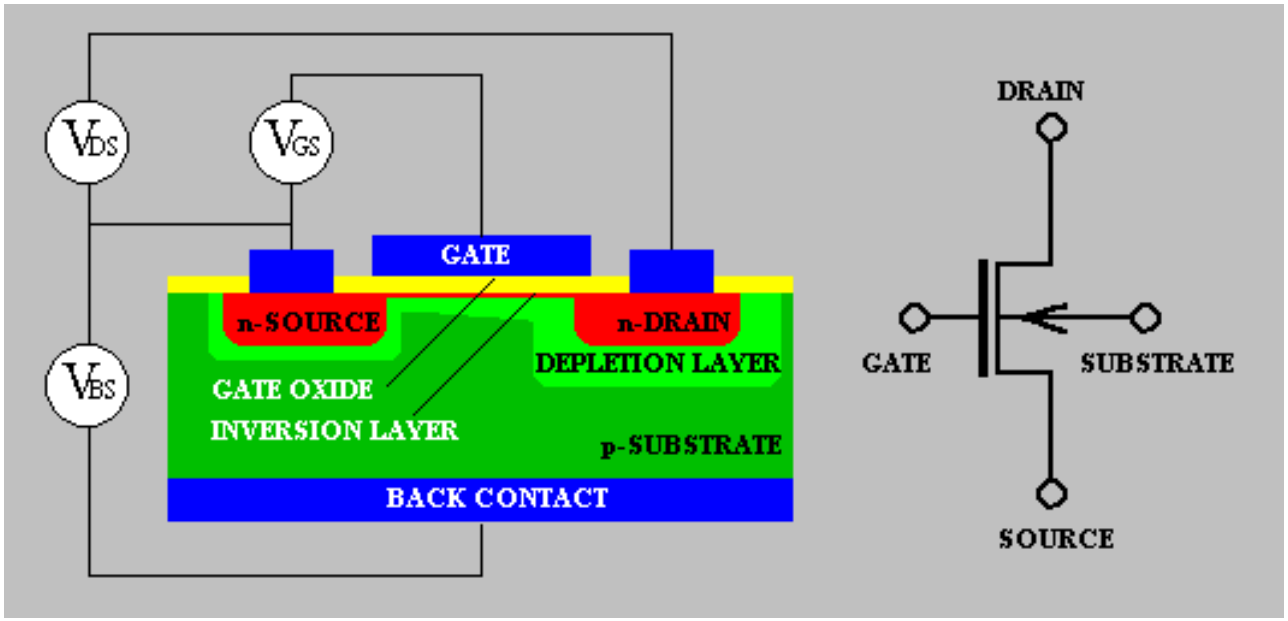
- It is the applied voltages which determine which n-type region provides the electrons and becomes the source, while the other n-type region collects the electrons and becomes the drain.

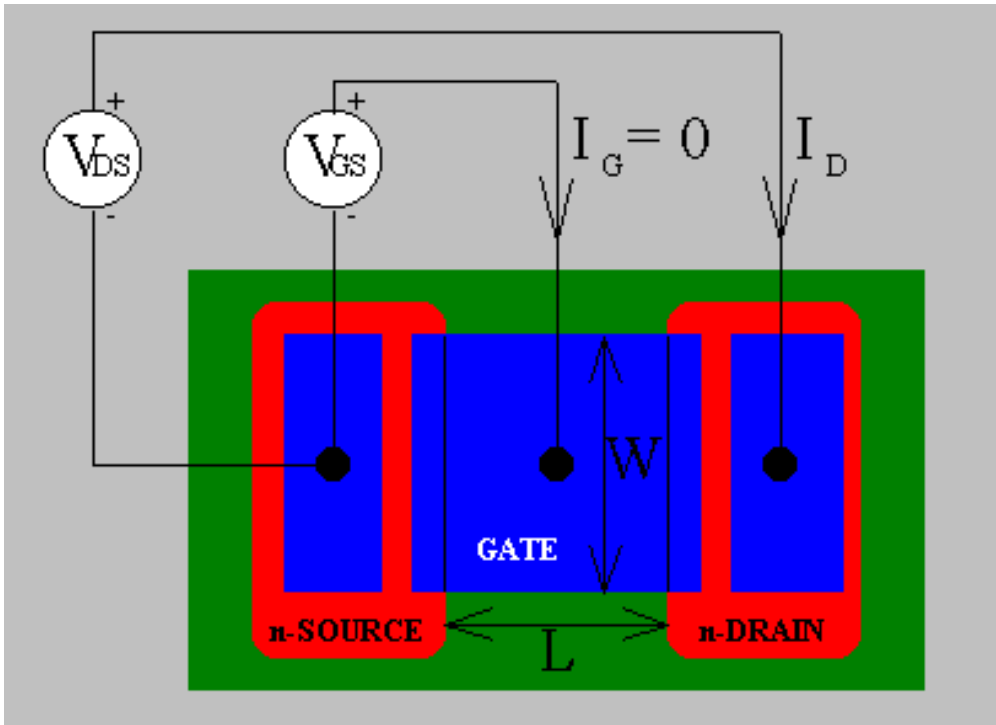
- The MOSFET works by electronically varying the width of a channel along which charge carriers flow (electrons or holes).

- The charge carriers enter the channel at source and exit via the drain.

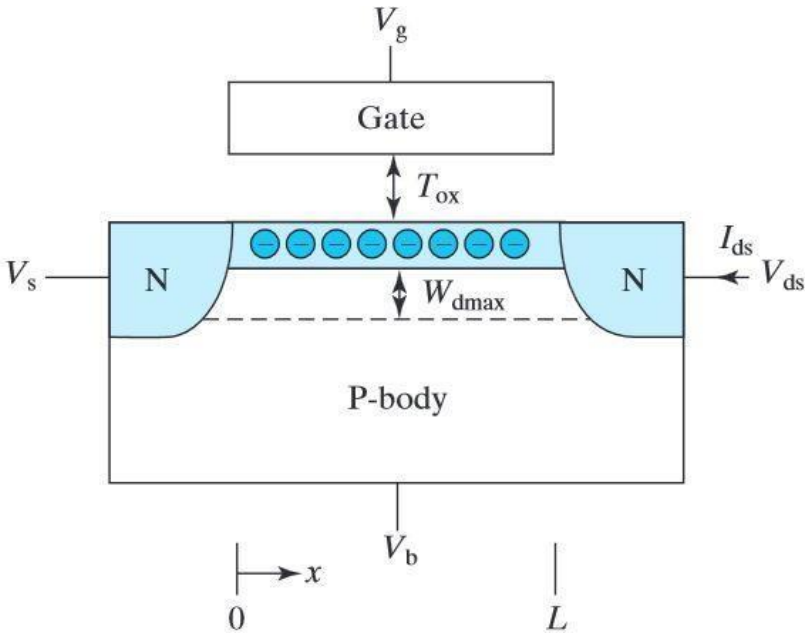
- The width of the channel is controlled by the voltage on an electrode is called gate which is located between source and drain.

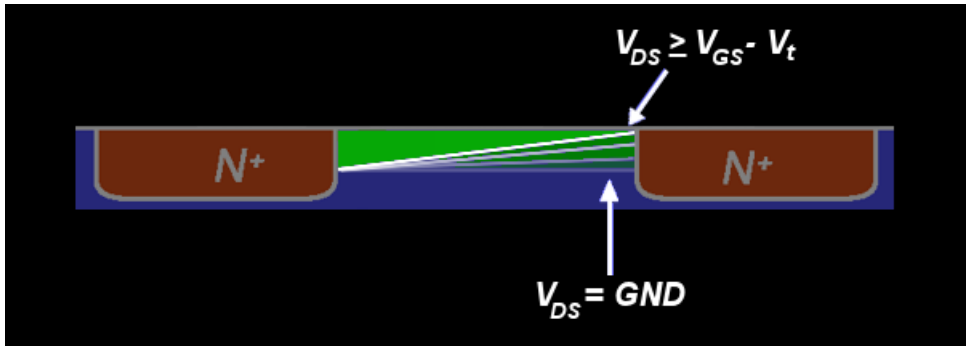
- It is insulated from the channel near an extremely thin layer of metal oxide.

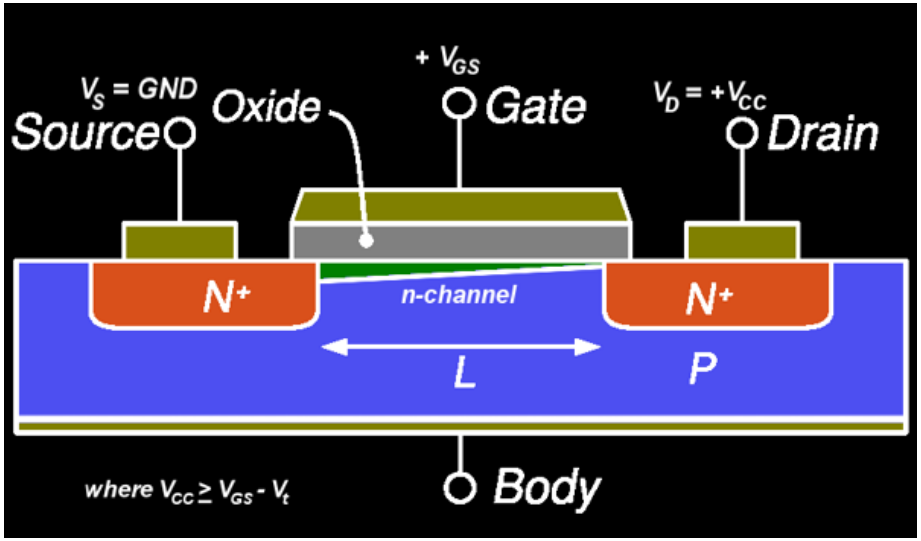

VDS
VGS
VBS
GATE
n-SOURCE
n-DRAIN
DEPLETION LAYER
GATE OXIDE
INVERSION LAYER
p-SUBSTRATE
BACK CONTACT
DRAIN
GATE
SUBSTRATE
SOURCE

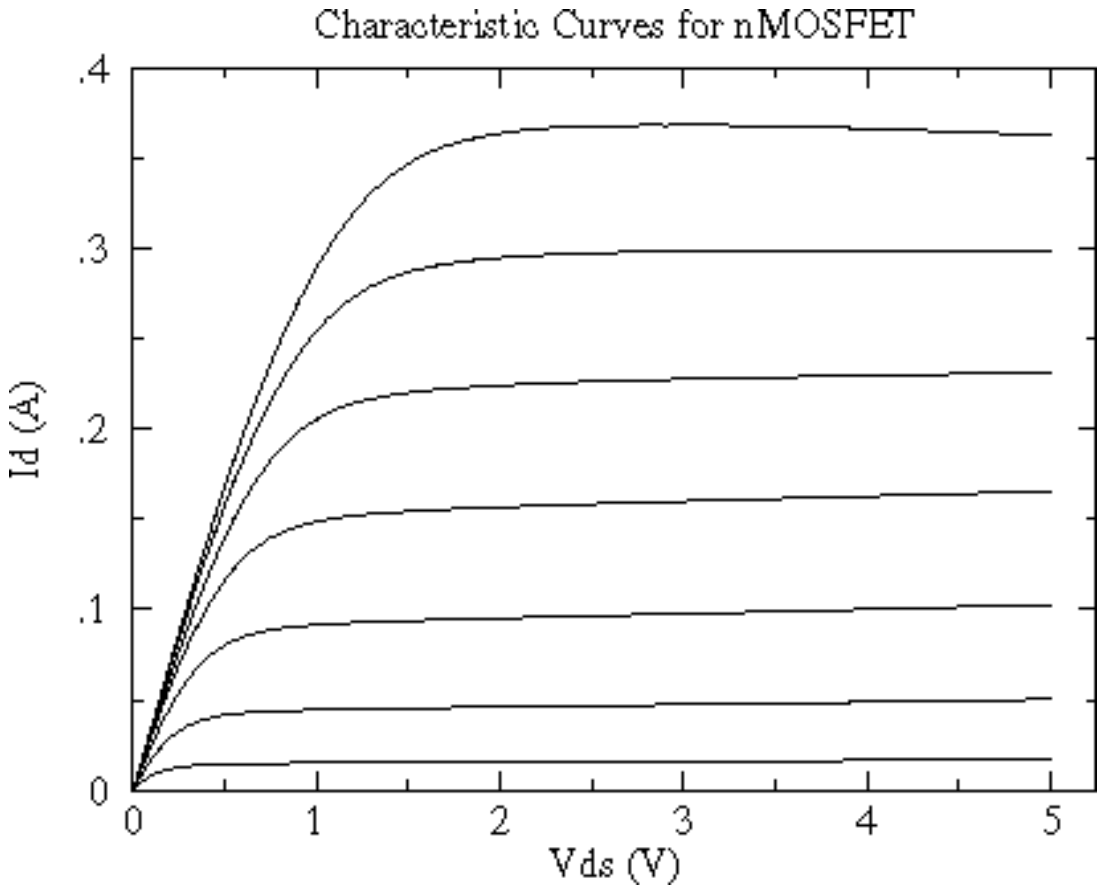

+
VDS
-
+
VGS
-
I G = 0
I D
W
GATE
n-SOURCE
L
n-DRAIN

- Consider NMOS.
- First, while applying positive V_{GS}, it causes a field that starts to repulse the majority carriers at the substrate.
- This causes "the depletion region" i.e depleted from majority carriers.
- On increasing V_{GS} more, the field starts to attract minority carriers from the substrate and form the channel. This is called "inversion".
- Now, if a voltage is applied between the drain and source, the current flows freely between the source and drain and the gate voltage controls the electrons in the channel


V_g
Gate
T_{ox}
V_s
N
W_{dmax}
N
I_{ds}
V_{ds}
P-body
V_b
x
0
L


$V_{DS} \geq V_{GS} - V_t$
N+
N+
V_{DS} = GND


V_S = GND
Source
Oxide
+ V_{GS}
Gate
V_D = +V_{CC}
Drain
N+
n-channel
N+
L
P
where $V_{CC} \geq V_{GS} - V_t$
Body


Characteristic Curves for nMOSFET
Id (A)
.4
.3
.2
.1
0
0
1
2
3
4
5
Vds (V)

- Inversion layer can be considered a 2D system of electrons immersed in a triangular-shaped quantum well.
- Quantized values for the energy of confinement is given as

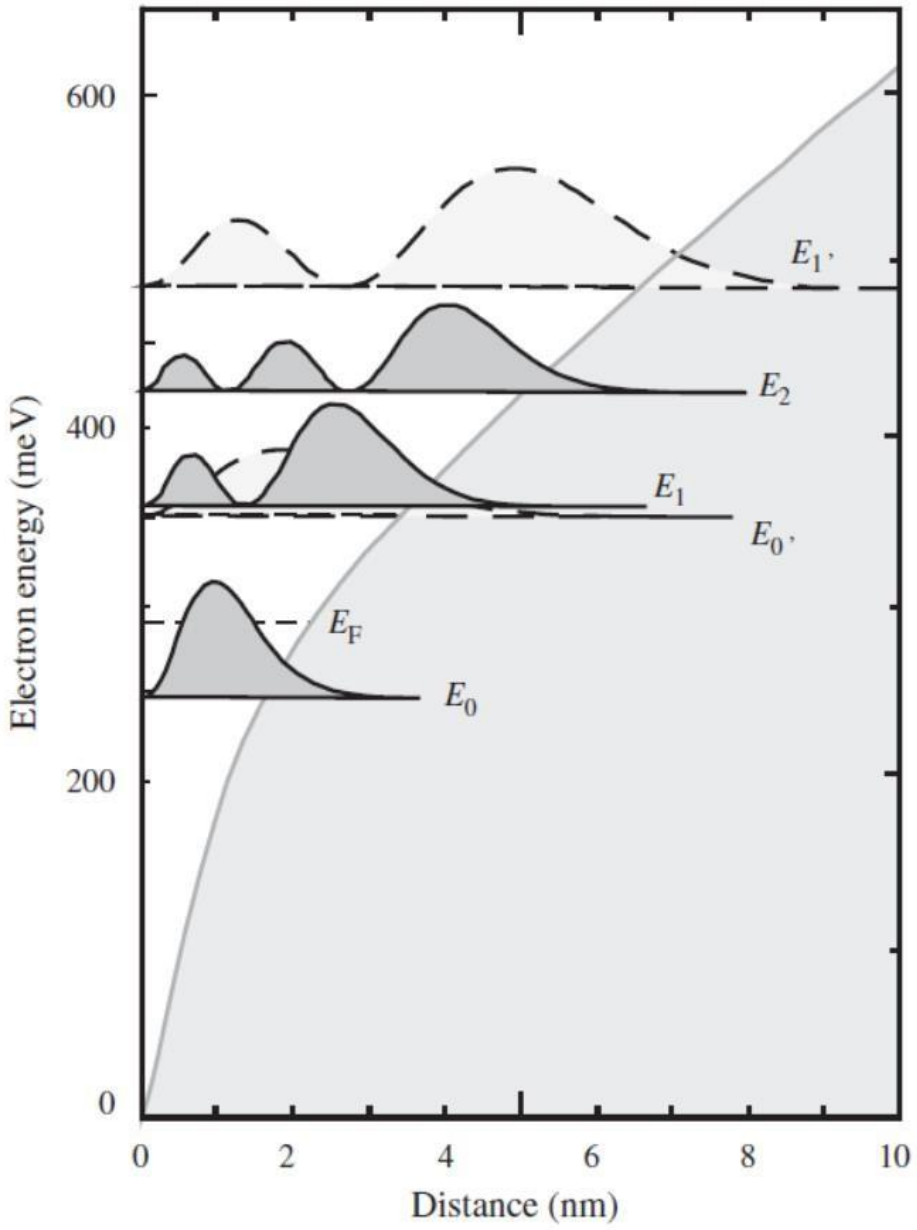
$$E = E_n + \frac{\hbar^2}{2m_x^*}k_x^2 + \frac{\hbar^2}{2m_y^*}k_y^2$$

- where *En* corresponds to the quantized energy for the triangular well

$$E_n \approx \left[\frac{3}{2}\pi\left(n - \frac{1}{4}\right)\right]^{2/3}\left(\frac{e^2F^2\hbar^2}{2m_z^*}\right)^{1/3}, \quad n = 1, 2, \ldots$$

- The density of states (DOS) function corresponds to the 2D case and is given by

$$g(E) = g_{\mathrm{v}} \frac{m_{\mathrm{T}}^*}{\pi \hbar^2}$$


600
400
200
0
Electron energy (meV)
0
2
4
6
8
10
Distance (nm)
E_1'
E_2
E_1
E_0'
E_F
E_0

HETEROJUNCTIONS

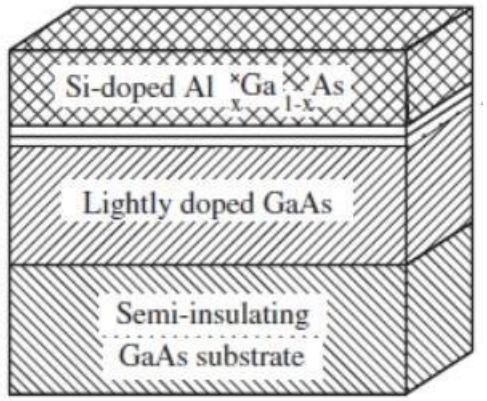
- Interfaces between two semiconductors of different gaps are called hetrojunction.

- If the dopants of the two sides are of the same type then it is called isotype heterojunction else called anisotype heterojunction.
- The most studied heterostructure is the one formed by n-type $Al_xGa_{1-x}As$ and almost intrinsic or lightly doped p-type GaAs

AlGaAs–GaAs interface

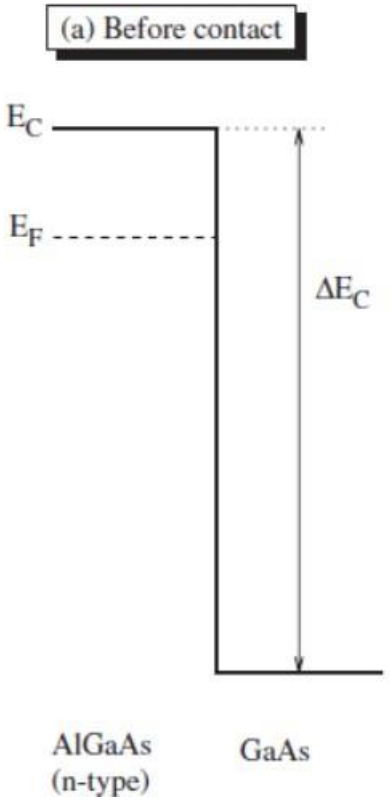
- Here the left material is gallium arsenide doped with aluminium and the right one is near-intrinsic GaAs.

- This structure is called a modulation-doped heterojunction and the method to produce it is known as modulation doping.

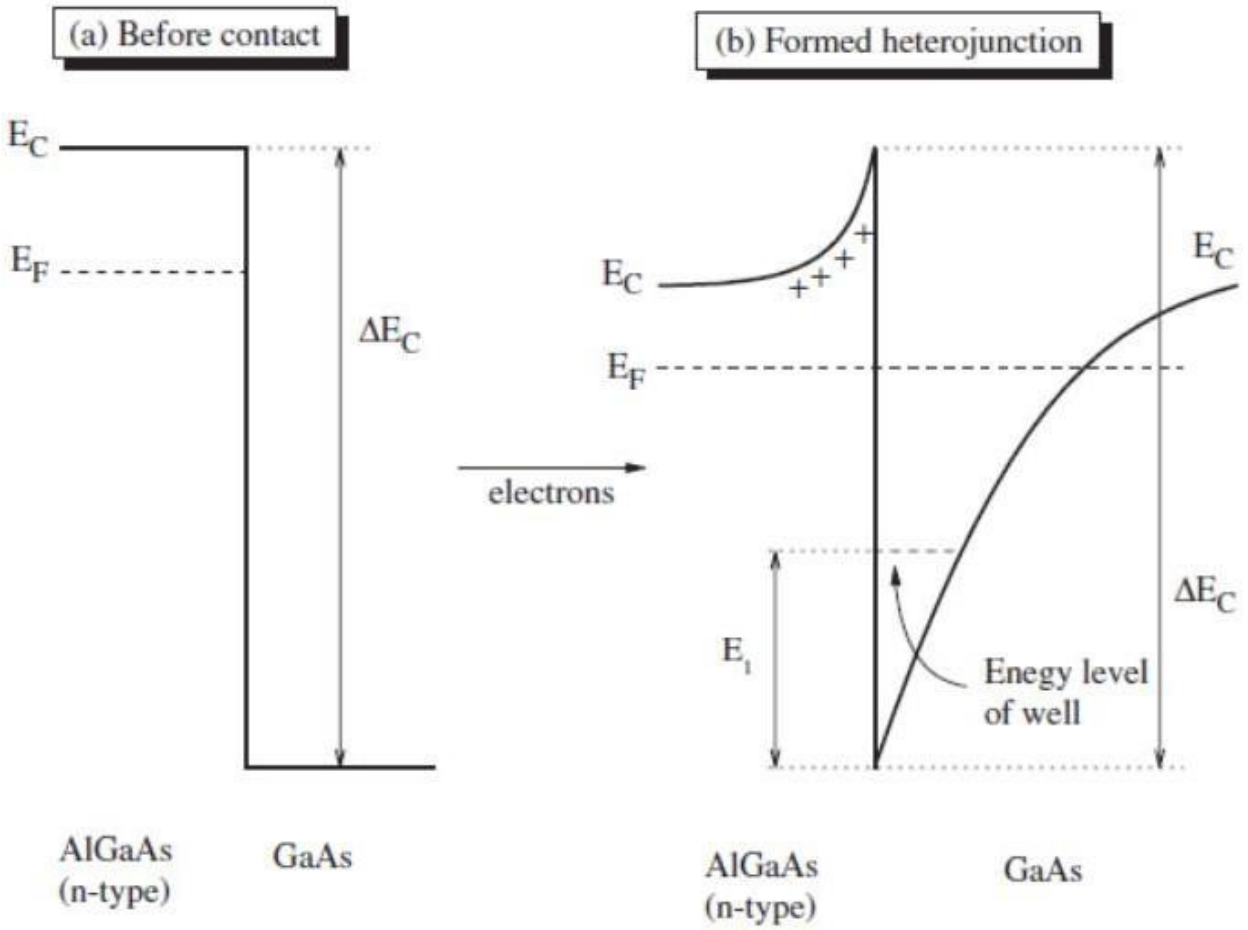


Before contact

- Fermi level of n-type AlGaAs is close to the conduction band, and for lightly p-doped GaAs is located close to the middle of the gap.
- Here the bands are flat because the materials have uniform doping.
- The barrier between them in the conduction band, vE_c, can be found using Anderson's rule and is approximately 0.35 eV



After Contact



- When both materials, AlGaAs and GaAs, enter in contact, some of the electrons from the donors of the n-material will cross the interface reaching the undoped GaAs.

- At equilibrium, the two Fermi levels line up, the bands are bent like in the case of the p–n junction.

- a quantum well for the electrons has been formed which is limited by a potential well of height v$E_{c.}$

- The quantum well for the electrons produced at the AlGaAs–GaAs interface has a shape close to a triangle as in the case of the MOS structure.

- The wells cannot be assumed to be of infinite height, since in our case $E_c \approx 0.3$ eV.

- So as in first module energy for the motion along *z* is quantized in potential well.

- The most important aspect of this heterojunction is that the charge carriers are located in a region (mainly in the GaAs), spatially separated from the AlGaAs semiconductor which originates the free electrons.
- The electrons in the well should have very high mobility for their motion along the (*x*, *y*) plane, since they move within the GaAs which is free of dopant impurities and it is well known that impurity scattering is one of the main factors which limit carrier mobility.

- Devices based on AlGaAs structures can be used to much higher frequencies than silicon devices due to the high mobility of electrons in GaAs.

- Transistors made with above hetrojunctions are called modulation-doped field effect transitor (MODFET) or high electron mobility transistor (HEMT).

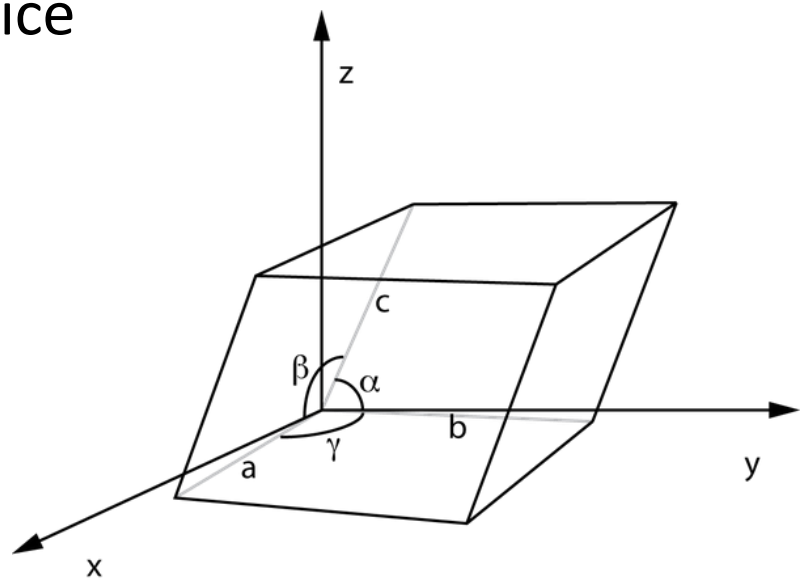
- MODFET can use in many high-frequency applications due to the very high electron mobility of the electrons in the channel.

Lattice constant

- The lattice constant is the physical dimension of the smallest repeating unit that possesses all the symmetry of the crystal structure.

- A 3D crystal is considered to have 3 lattice constants in terms of length of the unit cell edges *a, b* & *c* and 3 lattice constants in terms of the angles between them α, β & γ.

- This set of 6 constants are termed lattice parameters and give the entire structure of the lattice



STRAINED LAYERS

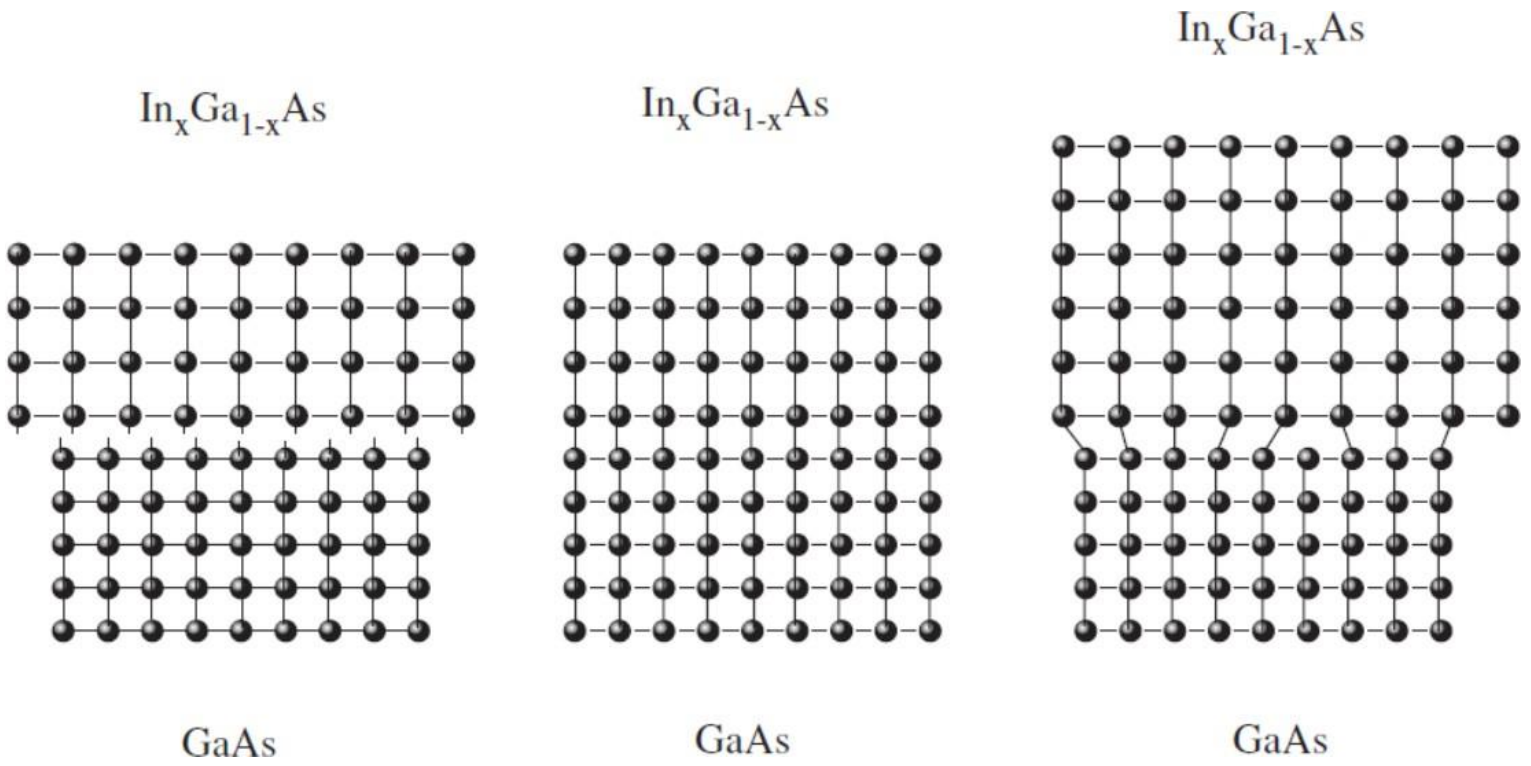
- The quality of an interface between two materials depends greatly on the relative size of the lattice constants.

- If the lattice constants are very similar, as in the case of the Al*x*Ga1−*x*As–GaAs heterojunctions (0.2% diffrence), the thermal expansion coefficients are similar and no stresses are introduced at the interface.

- Heterojunctions with differences in lattice constants up to 6% are fabricated (for instance, In*x*Ga1−*x*As–GaAs). In this case strong stresses appear at the interface.

- In the above case, only very thin films of a few monolayers can be grown on a given substrate.

STRAINED LAYERS FORMATION



- (a) The layer is not yet deposited over the substrate;
- (b) the layer is subjected to a compressive stress;
- (c) Dislocations are formed at the interface.

- Suppose that a layer of lattice constant aL is grown on a substrate of lattice constant aS. The strain ε of the layer is defined as

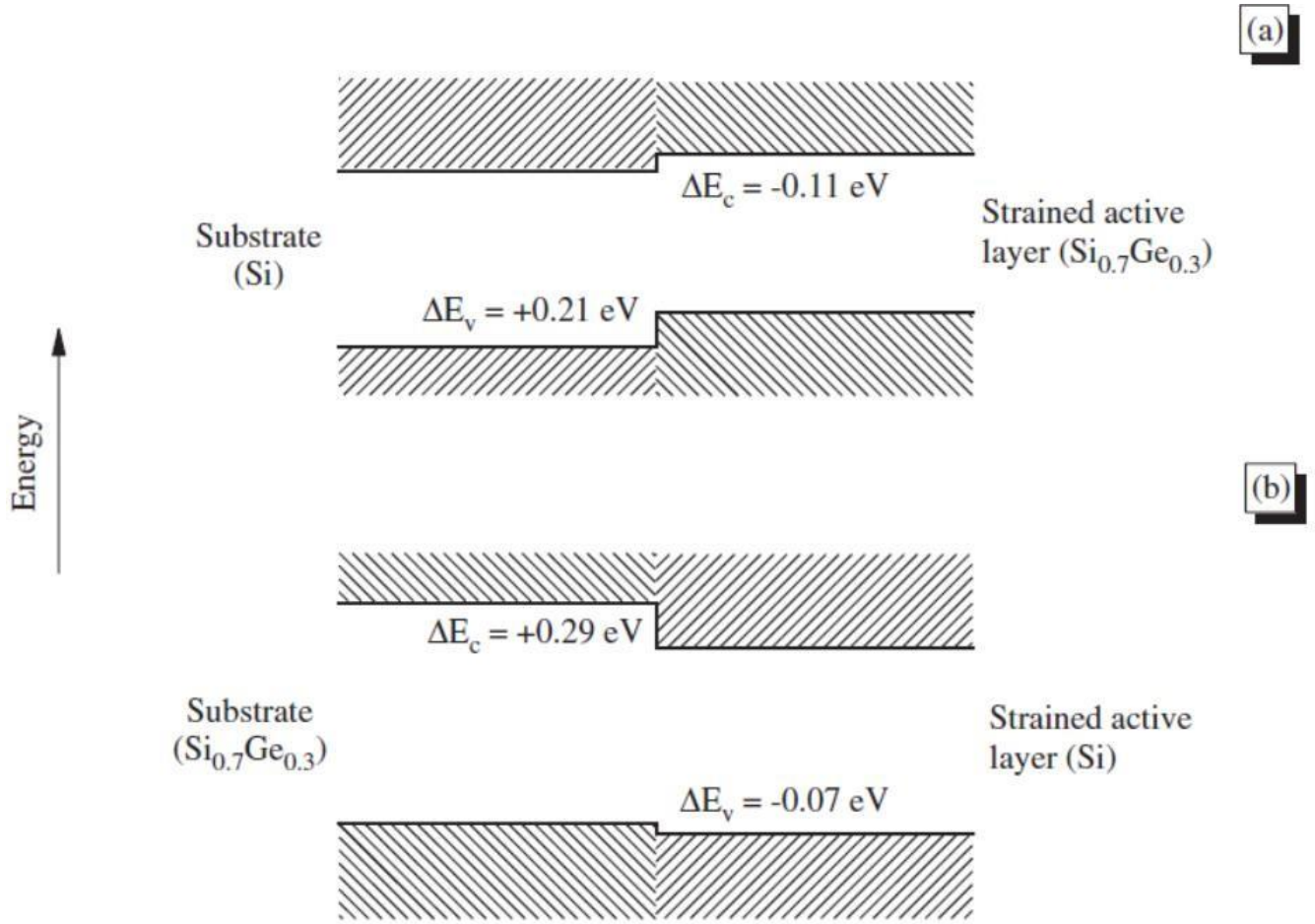
$$\varepsilon = \frac{a_{\mathrm{L}} - a_{\mathrm{S}}}{a_{\mathrm{S}}}$$

SiGe strained heterostructure

- SiGe heterojunctions did not attract too much attention at first because of the large lattice constant difference between Si and Ge (around 4%).

- This means that the layers grow strained over the substrate.

- SiGe heterostructures have found interesting applications in several fields such as high frequency transistors and IR photodetectors.

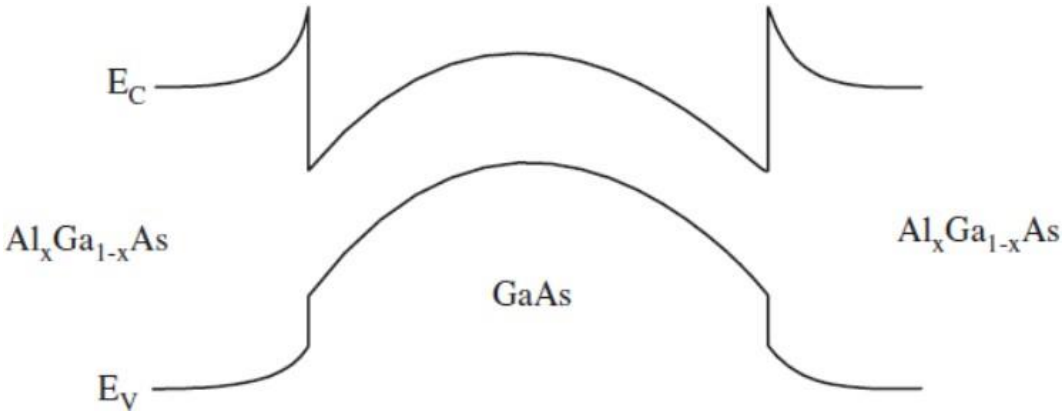


Simplified band diagram of two strained SiGe heterostructures. In (a) the active layer is SiGe, while in (b) the SiGe acts as a substrate.

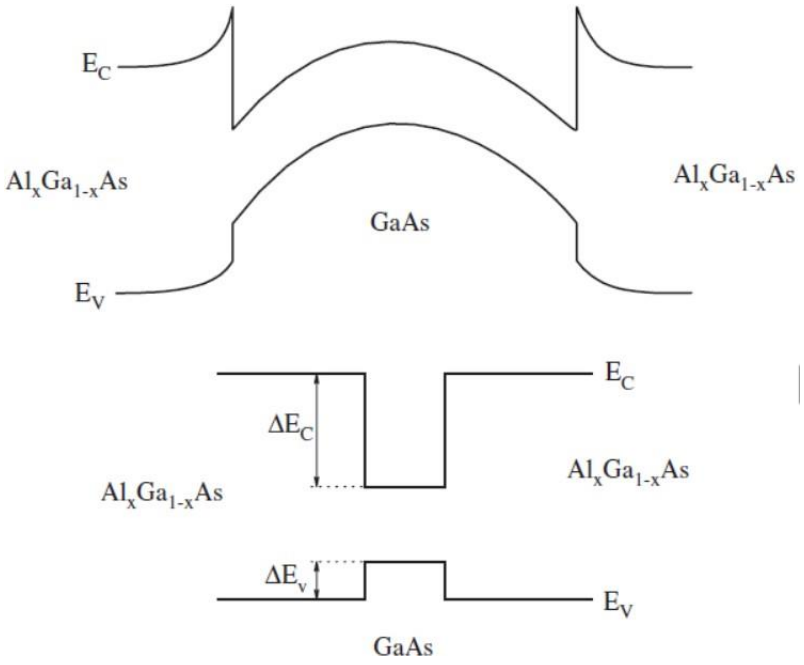
- In first case the conduction band offset is rather small, in contrast to the valence band offset.
- In second case the discontinuity in the conduction band is fairly large.

Modulation-doped quantum well

- Let us consider that we build a symmetric well by facing two AlGaAs–GaAs heterojunctions opposing each other.
- The wide gap semiconductor material Al*x*Ga1−*x*As is located at the ends and the GaAs in the middle.



- The distance between the two interfaces is made sufficiently small.
- Then the resulting well for electrons and holes would be almost square with a barrier on each side of the same height.



- The electrons inside the well, which originated at the neighbouring AlGaAs donor-material, can move into the GaAs region or channel with very high mobility.

- Quantum well structures with either high or low mobility for electrons can be fabricated by introducing a controlled amount of impurities.

- A double quantum well structure with high and low mobilities constitutes the base of the velocity-modulation transistors.

- Velocity-modulation transistors can be operated at very high frequencies.

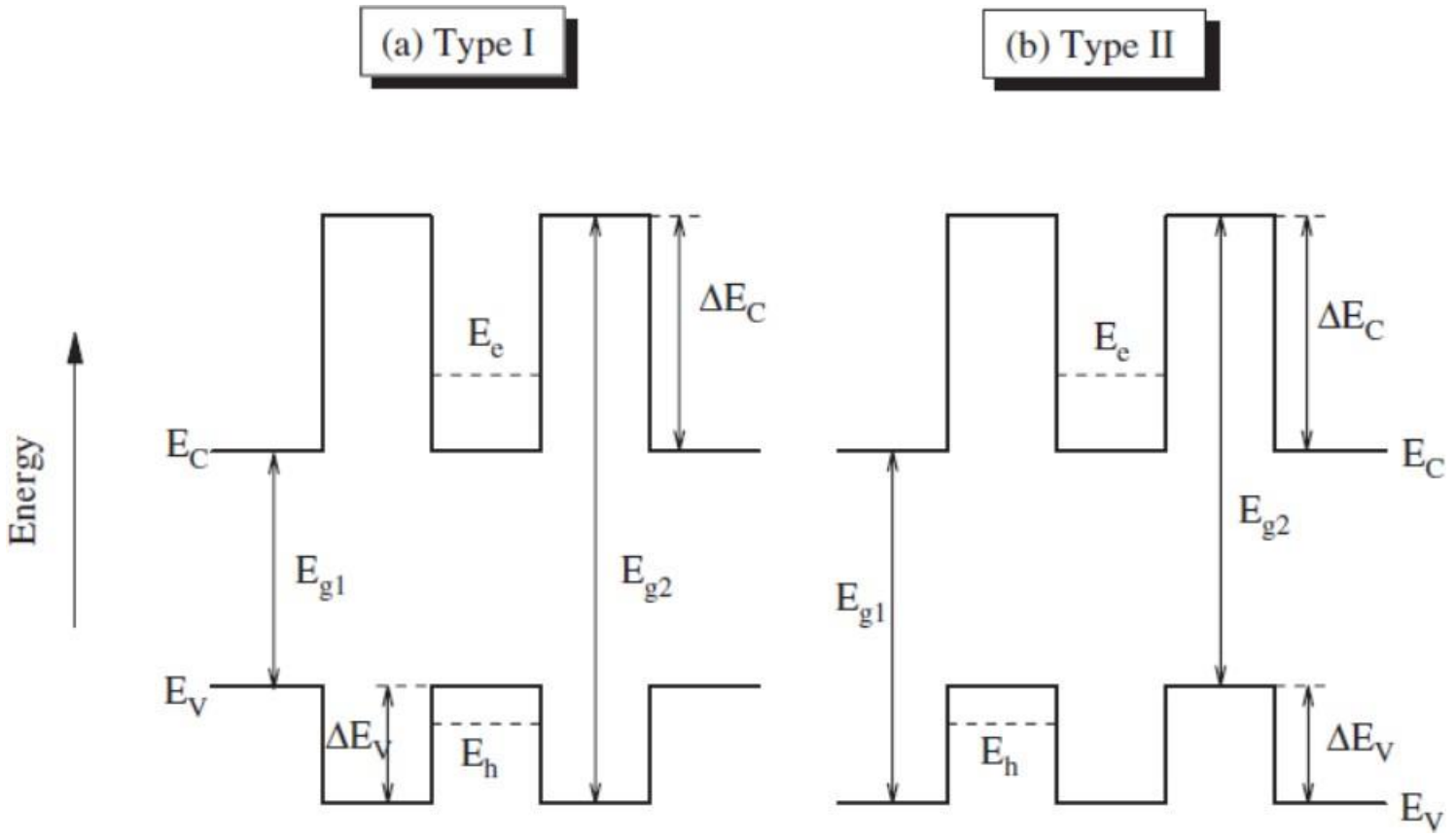
- The occupation of levels depends on the electron concentration in the well, and for low concentrations usually the first level is the only one occupied.

Multiple quantum wells (MQW)

- It is an array of quantum wells formed by several single quantum wells.

- If the wells for electrons and holes are located in the same space location, the MQW is called *Type I*

- while the name *Type II* is used when the corresponding wells are located alternatively as in Figure.


(a) Type I
(b) Type II
Energy
E_C
E_V
E_e
E_h
E_{g1}
E_{g2}
ΔE_C
ΔE_V

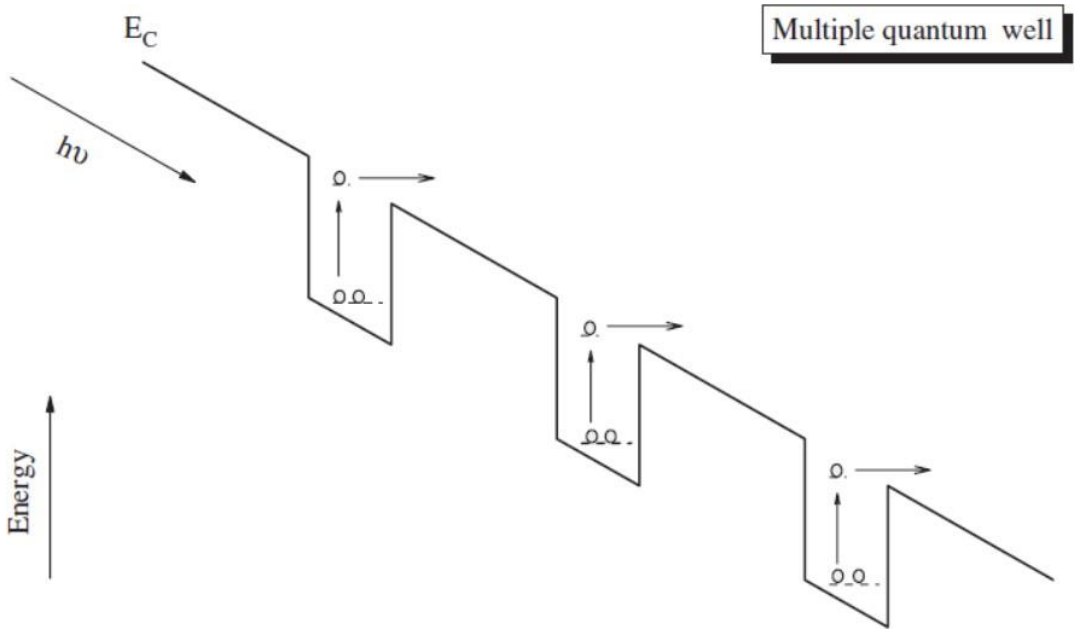
- In a MQW system it is assumed that there is no interaction between neighbouring quantum wells, because the barriers separating the wells are thick enough, usually more than about 10 nm.

- However, if the energy barriers between consecutive wells are thin enough, the wells will be coupled to each other by tunnelling effects.

- the discrete energy levels of the quantum wells are then transformed into energy bands.

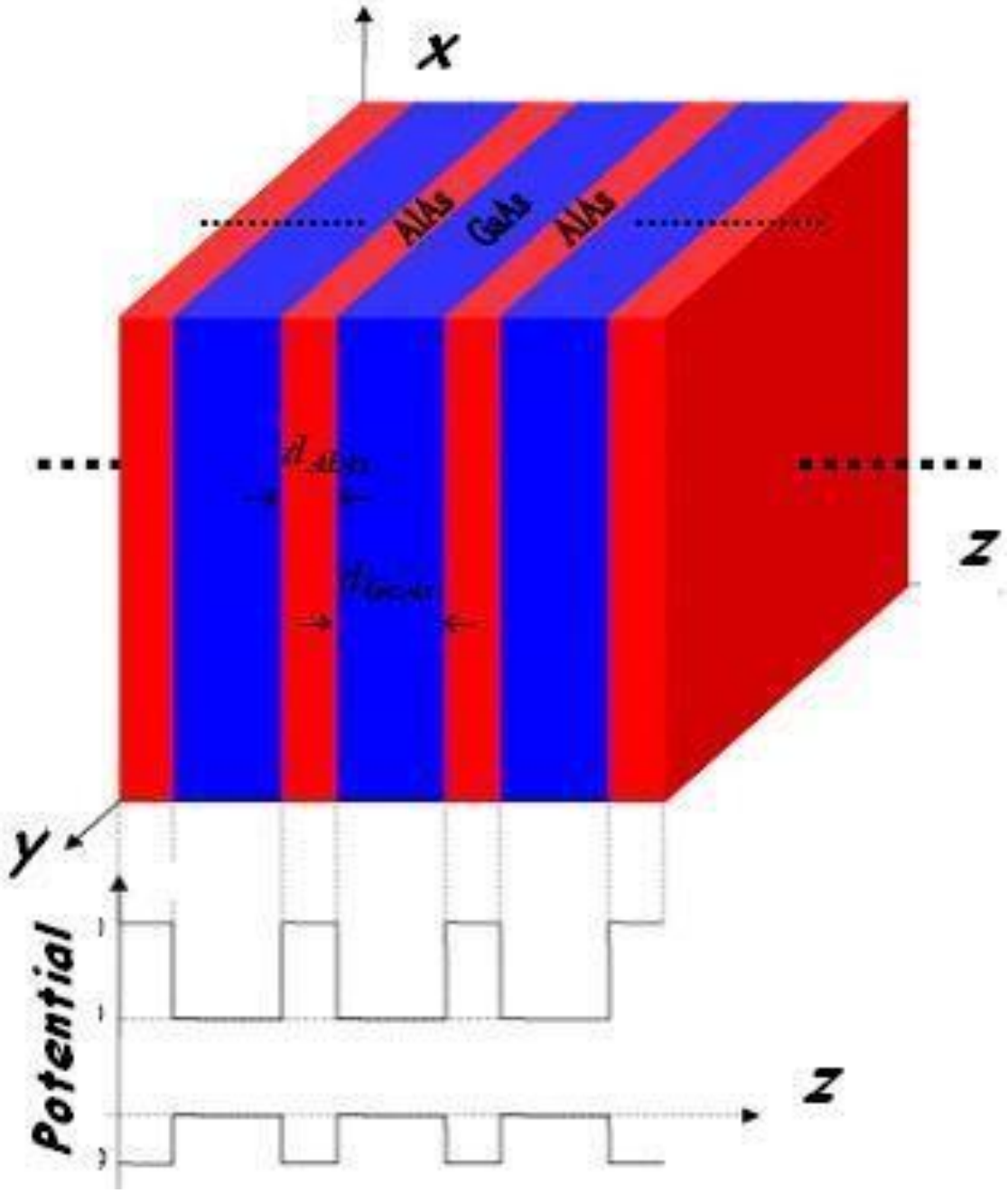
- In this case, the system of MQWs is called a superlattice and the energy spectrum shows very interesting new features.

- In optoelectronics, MQWs are frequently used and are made of about 50 single wells.
- thickness of each well is not exactly the same for all.
- Figure shows the band structure of a typical MQW for applications in IR photo detectors.



SUPERLATTICES

- A **superlattice** is a periodic structure of layers of two (or more) materials.
- Typically, the thickness of one layer is several nanometers.
- It can also define as an array of quantum wells.
- Layers are fabricated by molecular beam epitaxy (MBE) techniques.
- Conditions of each quantum well will affect charge flow through the structure.


x
AlAs
GaAs
AlAs
z
y
Potential
z

- A superlattice consists of a periodic set of MQW in which the thickness of the energy barriers separating the individual wells is made sufficiently small.

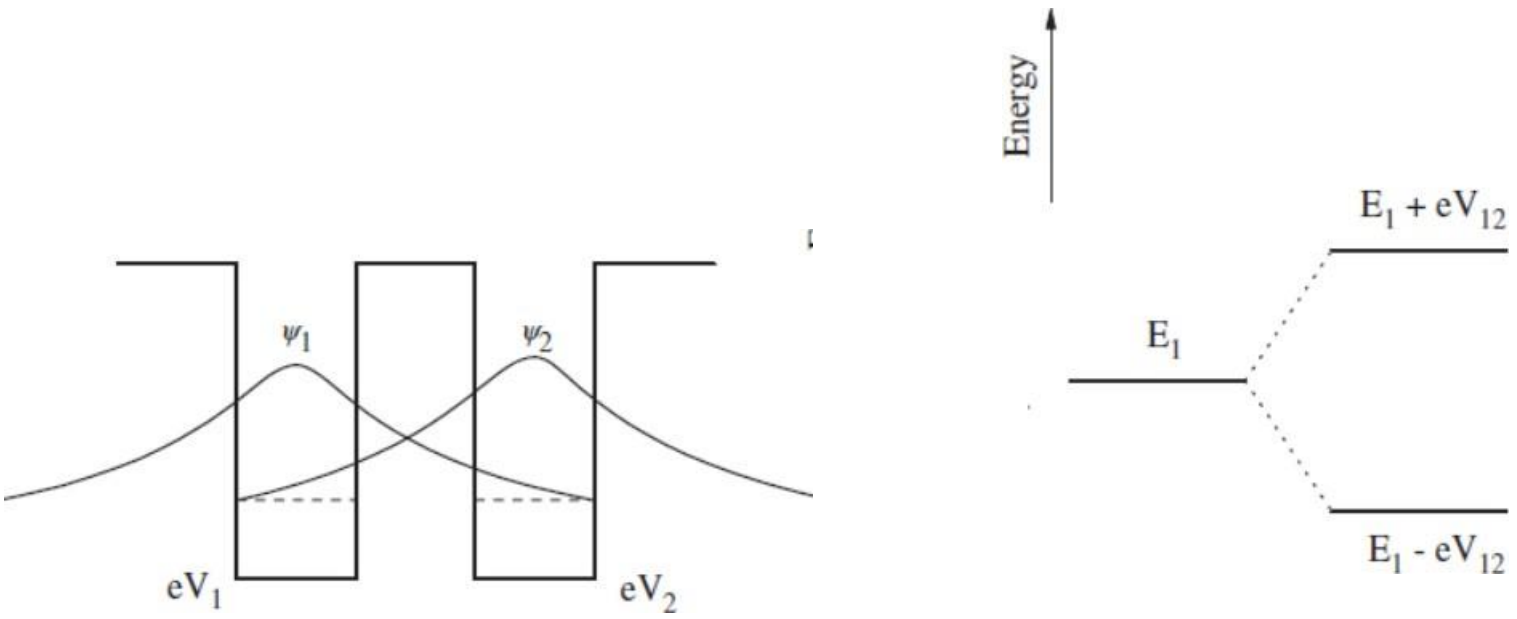
- As the barriers become thinner, the electron wave functions corresponding to the wells overlap due to the tunnelling effect.

- As a consequence, the discrete energy levels of the wells broaden and produce energy bands.

- Superlattice have periodicity *d* in the material, which is equal to the breadth of the well *a*, plus the thickness of the barrier *b.*

- Typical thicknesses for *a* and *b* could be 4 and 2 nm, respectively.

- let us consider first the overlapping between the electron states for a simple two-well system.



- Above figure shows two neighboring identical quantum wells and corresponding wave functions of what is known as the double coupled quantum well system.

- Each original energy level, say $E1$, of the isolated wells splits into two, with energies

$$E = E1 \pm |V12|$$

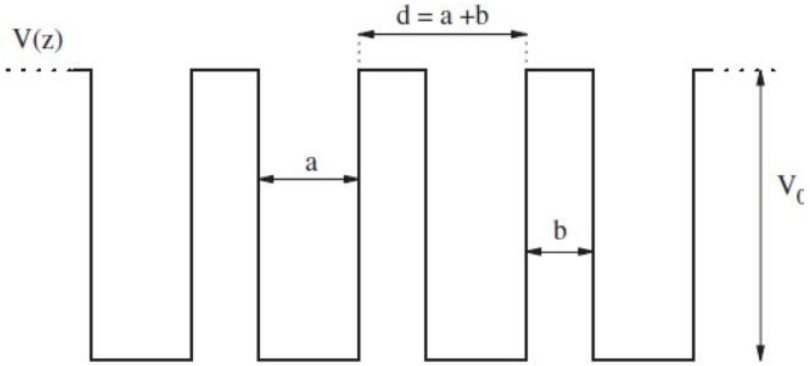
$V12$: overlap integral.

- where the magnitude of $V12$ is an indication of how much one well can influence the energy states of the neighboring one, hence the name overlap integral.

$$V_{12} = \int_{-\infty}^{+\infty} \psi_1^* V(z) \psi_2 dz$$

Kronig–Penney model of a superlattice.

- As per Kronig and Penney model, periodic potential seen by the electrons was precisely that of the square type for a superlattice potential.



Scheme of the periodic potential of a superlattice.

- This periodic one-dimensional potential is characterized by the following parameters: well thickness *a*, barrier thickness *b*, and barrier height *V*0.

- The spatial periodicity is $d = a + b$

Wave function in Kronig–Penney model

- In the well region ($0 < z < a$), $V = 0$

$$\psi(z) = Ae^{ik_0 z} + B\mathrm{e}^{-ik_0 z} \qquad k_0^2 = \frac{2mE}{\hbar^2}$$

- In the barrier region $-b < z < 0$, V= $V0$

$$\psi(z) = Ce^{qz} + De^{-qz}$$

A, *B*, *C*, and *D are amplitudes of the wave*

- Wave vector and the energy are related by

$$V_0 - E = \frac{\hbar^2 q^2}{2m}$$

- Kronig and Penney suggested that
  A+ B= C+D

  $ik0(A - B) = q(C - D)$

- Applying this for the above case

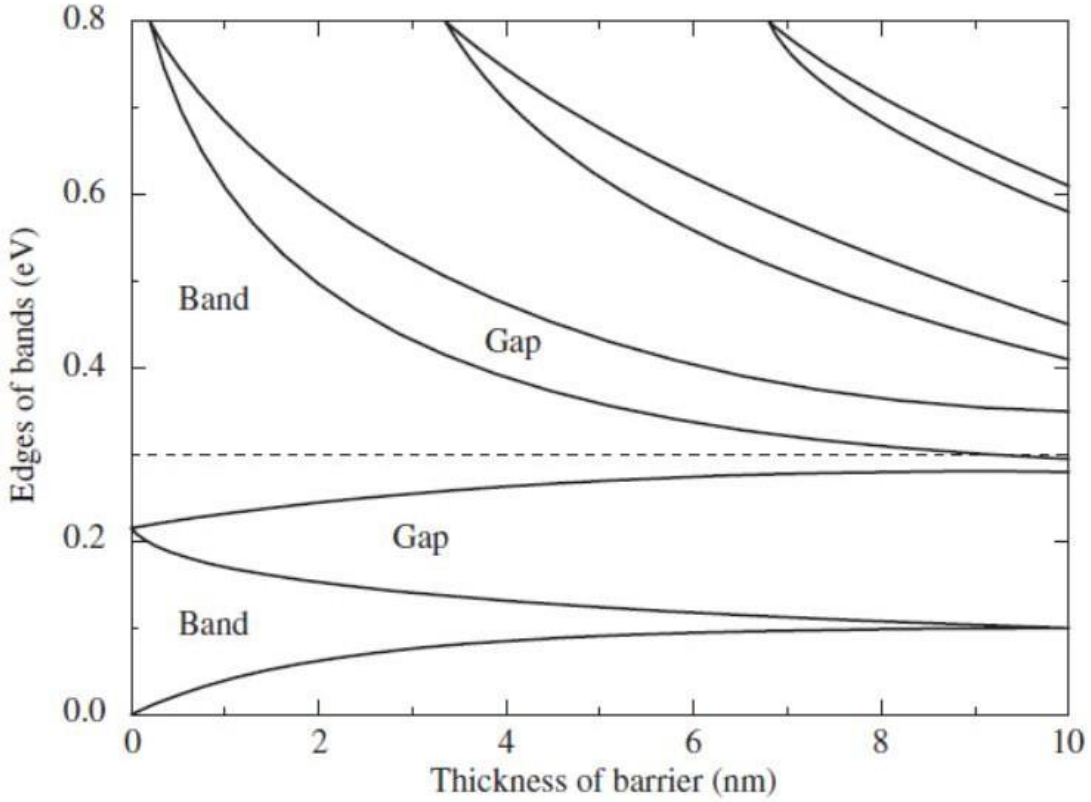
$$Ae^{ik_0a} + Be^{-ik_0a} = (Ce^{-qb} + De^{qb})e^{ik_0(a+b)}$$

$$ik_0(Ae^{ik_0a} - Be^{-ik_0a}) = q(Ce^{-qb} - De^{-qb})e^{ik_0(a+b)}$$

- On solving above two expressions

$$\frac{q - k_0^2}{2qk_0} \sin k_0a \sinh qb + \cos qa \cosh qb = \cos q(a + b)$$

Energy band diagram of a superlattice

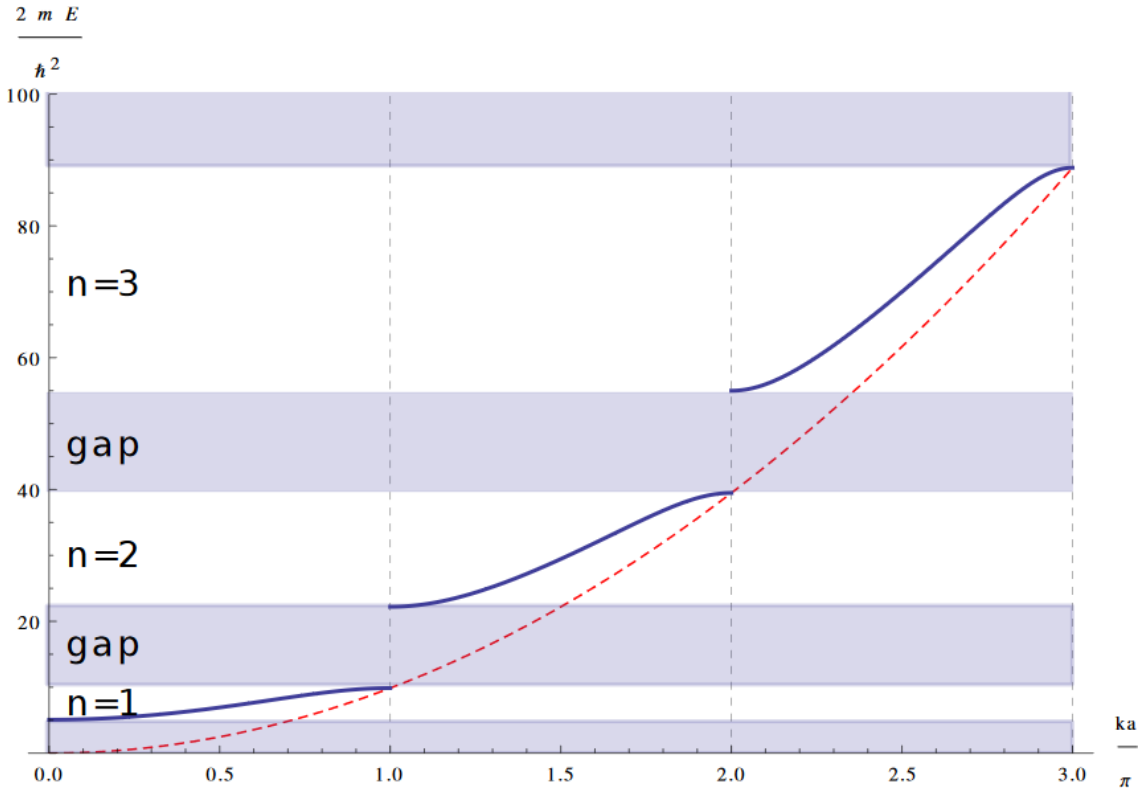


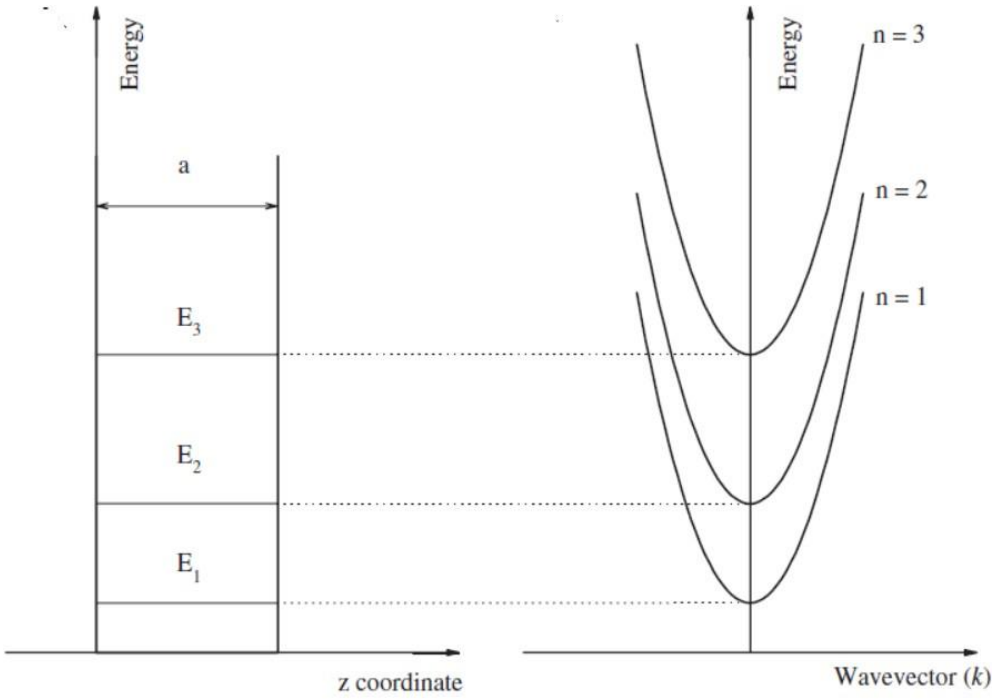
Energy band diagram of a superlattice with the width of the barrier equal to the width of the wells.

- When barrier thickness is greater than 10nm, each wells are isolated and will acts as quantum wells.
- Thus energy states are quantized and there will not be any separation in bands.
- For values of *a* larger than about 10 nm, the electron energies are well defined and correspond to the individual quantum wells.
- When barrier thickness is lesser than 6nm, wells are connected together and hence the structure become a normal material. This will separate quantized energy levels into different bands and forbidden states.
- When the barrier width is smaller than about 6 nm, bands as well as forbidden zones arise.

Energy and formation of bands

- In super lattice the free electron parabola, breaks down into several bands and gaps.

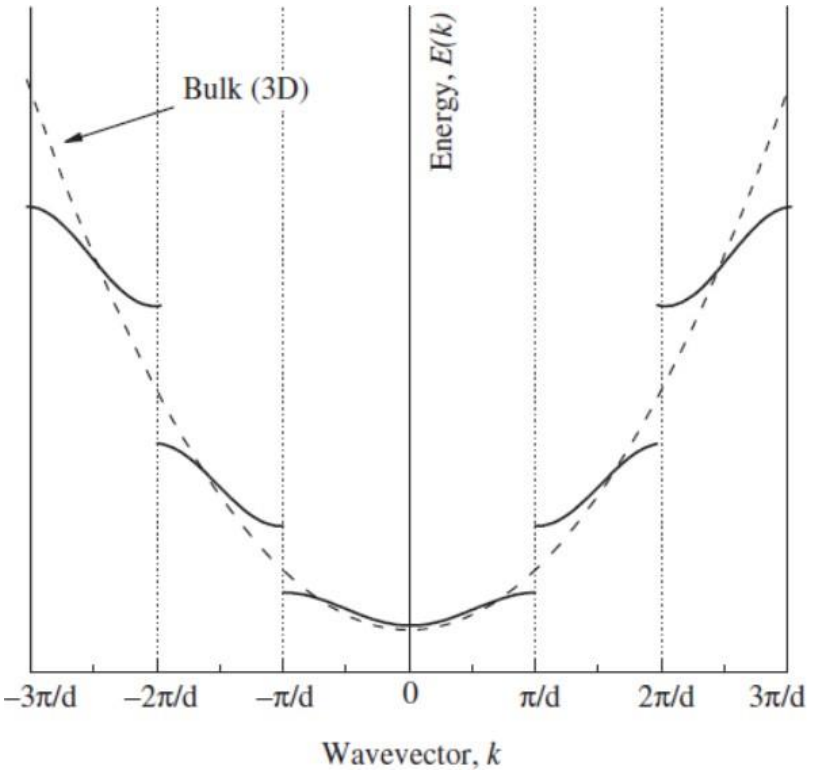




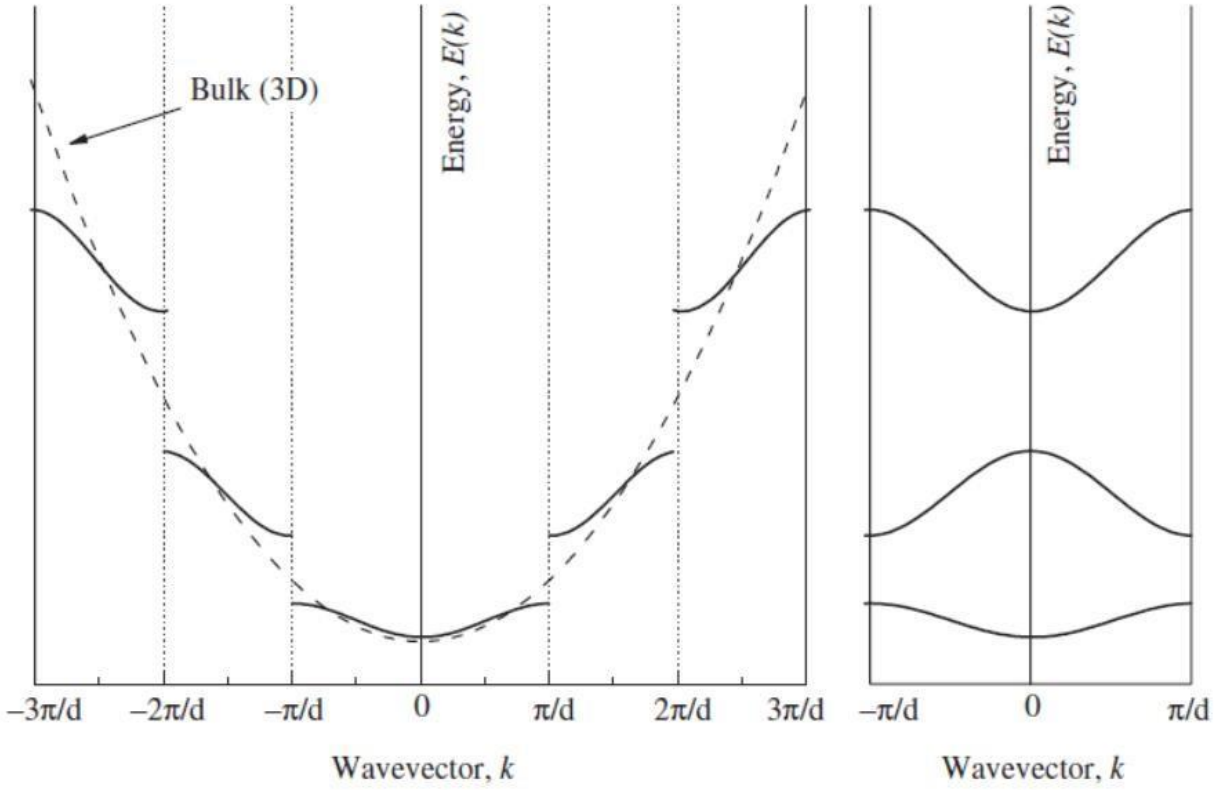
Energy levels in quantum well

Energy subbands

- The dashed line is the parabola corresponding to free electrons.($k_0^2 = \frac{2mE}{\hbar^2}$)



- The portions of the bands can be translated to the reduced zone.



- This band folding procedure is typical of superlattices and is called *zone folding* since it implies that the pieces of the band in the extended representation are zone-folded into the smaller zone.