Quantum physics and electrical conduction

Wavefunction and Schrödinger's equation

Wave-particle duality

- All particles can sometimes behave as waves (diffract, interfere) ↳ wavepackets.
- Equating the total energy $E=mc^2$ w/ the Planck energy of a quantum $E=hf$, we get

$$E=mc^2=hf$$

For a wave travelling at speed c, $c=f\lambda$, so

$$mc\cdot(f\lambda)=hf$$

$$\boxed{\lambda=\frac{h}{p}}$$

where $p=mc$ is the momentum of the particle.

- The momentum p can be written as

$$\boxed{p=\frac{h}{\lambda}=\frac{h}{2\pi}\cdot\frac{2\pi}{\lambda}=\hbar k}$$

The energy E can be written as

$$\boxed{E=hf=\frac{h}{2\pi}\cdot 2\pi f=\hbar\omega}$$

where $\hbar=\frac{h}{2\pi}$ is the reduced Planck's constant ($h=6.63\times10^{-34}$ Js)

- Wave effects of particles are significant when the distance of obstacles is comparable to its wavelength

The wavefunction ψ.

- The wavefunction ψ contains information on all the measurable parameters of the particle.
- In the 1D case, the wavefunction is $\psi=\psi(x,t)$.
- The probability density function of finding the particle is

$$f(x)=\psi\cdot\psi^*$$

For real $\psi(x,t)$, this becomes

$$f(x)=|\psi(x,t)|^2$$

- The probability of finding a particle btwn $x=a$ and $x=b$ is

$$p(a\le x\le b)=\int|\psi|^2\,dx\ldots$$

- We can operate on the wavefunction ψ to obtain the parameters of the wave:
  - ↳ Momentum operator $\boxed{\hat{p}_x=\frac{\hbar}{i}\frac{\partial}{\partial x}}$. so $\hat{p}_x(\psi)=\frac{\hbar}{i}\frac{\partial\psi}{\partial x}=p_x\cdot\psi$.
  - ↳ Energy operator $\boxed{\hat{E}_x=-\frac{\hbar}{i}\frac{\partial}{\partial t}}$ so $\hat{E}_x(\psi)=-\frac{\hbar}{i}\frac{\partial\psi}{\partial t}=E_x\psi$

- eg: Consider a plane wave $\psi=Ae^{i(kx-\omega t)}$. Find its momentum and energy.

Applying momentum operator $\hat{p}_x$ on the wavefunction ψ,

$$\hat{p}_x(\psi)=\frac{\hbar}{i}\frac{\partial\psi}{\partial x}=\frac{\hbar}{i}ikAe^{i(kx-\omega t)}=\hbar k\psi=\frac{h}{2\pi}\cdot\frac{2\pi}{\lambda}\psi=p_x\psi.$$

Applying energy operator $\hat{E}_x$ on the wavefunction ψ,

$$\hat{E}_x(\psi)=-\frac{\hbar}{i}\frac{\partial\psi}{\partial t}=\frac{\hbar}{i}(i\omega)Ae^{i(kx-\omega t)}=\hbar\omega\psi=\frac{h}{2\pi}\cdot 2\pi f\psi=E_x\psi.$$

The schrödinger's equation.

- The kinetic energy of a particle is given by $T = \frac{P_x^2}{2m}$, so

$$T\psi = \frac{P_x^2\psi}{2m} = \frac{1}{2m}\cdot\frac{\hbar}{i}\frac{\partial}{\partial x}\left(\frac{\hbar}{i}\frac{\partial\psi}{\partial x}\right) = -\frac{\hbar^2}{2m}\frac{\partial^2\psi}{\partial x^2}$$

(The kinetic energy operator $\boxed{\hat{T} = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}}$ so $\hat{T}(\psi) = -\frac{\hbar^2}{2m}\frac{\partial^2\psi}{\partial x^2} = T\psi$)

- The potential energy of a particle is given by V.

- The schrödinger's eqn is a simple statement of $T+V=E$, i.e.

kinetic energy + potential energy = total energy

$$\boxed{-\frac{\hbar^2}{2m}\frac{\partial^2\psi}{\partial x^2} + V\psi = -\frac{\hbar}{i}\frac{\partial\psi}{\partial t}}$$

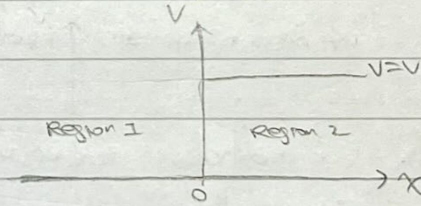
If the total energy is constant, i.e. $-\frac{\hbar}{i}\frac{\partial\psi}{\partial t} = E\psi = \text{const}$, then we have the time independent schrödinger's eqn. (TISE)

$$\boxed{-\frac{\hbar^2}{2m}\frac{\partial^2\psi}{\partial x^2} + V\psi = E\psi}$$

- To solve for the arbitrary consts in the general soln for schrödinger's eqn, we use the following properties of the wavefunction ψ:
  - ↳ ψ must be continuous (since $|\psi|^2$ is a probability density, which must be continuous)
  - ↳ $\frac{d\psi}{dx}$ must be continuous (o/w, we have infinite $\frac{d^2\psi}{dx^2}$ → infinite energy, as $T\psi = -\frac{\hbar^2}{2m}\frac{\partial^2\psi}{\partial x^2}$)
  - ↳ $\boxed{\int_{-\infty}^{\infty}|\psi|^2 dx = 1}$ (probability of finding the particle somewhere is 1)

Particle and a potential barrier (infinite width)

- consider the potential energy profile $V(x) = \begin{cases} 0 & x\le 0 \ \text{(region I)} \\ V_0 & x>0 \ \text{(region II)} \end{cases}$



The TISE is:

Region I $\quad -\frac{\hbar^2}{2m}\frac{\partial^2\psi}{\partial x^2} = E\psi$

Region II $\quad -\frac{\hbar^2}{2m}\frac{\partial^2\psi}{\partial x^2} + V_0\psi = E\psi$

- case 1 : $E > V_0$.

Region I : $\alpha_1 = \sqrt{\frac{2m}{\hbar^2}E}$ is real → $\psi = A_1e^{i\alpha_1 x} + B_1e^{-i\alpha_1 x}$

Region II : $\alpha_2 = \sqrt{\frac{2m}{\hbar^2}(E-V_0)}$ is real → $\psi = A_2e^{i\alpha_2 x} + \underline{B_2e^{-i\alpha_2 x}}$

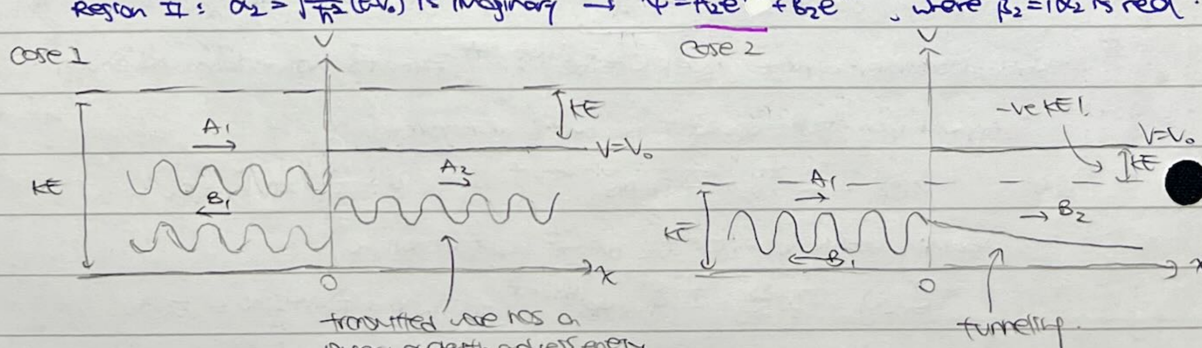
leftward transmitted wave is zero as there is no physical reason the wave reflects at $x>0$.

- case 2: $E<V_0$

Region I: $\alpha_1 = \sqrt{\frac{2m}{\hbar^2}E}$ is real → $\psi = A_1e^{i\alpha_1 x} + B_1e^{-i\alpha_1 x}$

Region II: $\alpha_2 = \sqrt{\frac{2m}{\hbar^2}(E-V_0)}$ is imaginary → $\psi = \underline{A_2e^{\beta_2 x}} + B_2e^{-\beta_2 x}$, where $\beta_2 = i\alpha_2$ is real.

exponential growth term is zero, o/w we have infinite ψ → infinite $|\psi|^2$



Particle and a potential barrier (finite width)

- Consider the potential energy profile $V(x) = \begin{cases} 0 & x \le 0 \quad \text{(region I)} \\ V_0 & 0 < x \le b \quad \text{(region II)} \\ 0 & x > b \quad \text{(region III)} \end{cases}$, and the case $E < V_0$

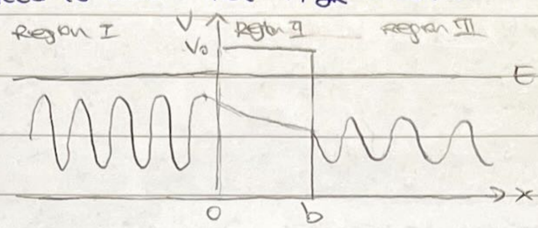
The TISE is:

Region I $\quad -\frac{\hbar^2}{2m}\frac{\partial^2 \psi}{\partial x^2} = E\psi \rightarrow \alpha_1 = \sqrt{\frac{2m}{\hbar^2}E}$ is real $\rightarrow \psi = A_1 e^{i\alpha_1 x} + B_1 e^{-i\alpha_1 x}$

Region II $\quad -\frac{\hbar^2}{2m}\frac{\partial^2 \psi}{\partial x^2} + V_0\psi = E\psi \rightarrow \alpha_2 = \sqrt{\frac{2m}{\hbar^2}(E - V_0)}$ is imaginary $\rightarrow \psi = A_2 e^{\beta_2 x} + B_2 e^{-\beta_2 x}$ $[\beta_2 = i\alpha_2]$

Region III $\quad -\frac{\hbar^2}{2m}\frac{\partial^2 \psi}{\partial x^2} = E\psi \rightarrow \alpha_3 = \sqrt{\frac{2m}{\hbar^2}E}$ is real $\rightarrow \psi = A_3 e^{i\alpha_3 x} + B_3 e^{-i\alpha_3 x}$ terms are zero due to same reasons as in previous section

- In general, region II intercepts the decaying exp. wave before it reaches 0, and allows it to continue as a travelling wave in region III. → wave can tunnel through finite barriers.
- To solve for ψ_i, we need to ensure that $\psi, \frac{\partial \psi}{\partial x}$ is continuous at $x=0, x=b$.



Particle in a box

- Consider the potential energy profile $V(x) = \begin{cases} \infty & x \le 0 \quad \text{(region I)} \\ 0 & 0 < x \le L \quad \text{(region II)} \\ \infty & x > L \quad \text{(region III)} \end{cases}$

The TISE is

Region I $\quad$ N/A (particle-wave cannot exist in infinite barrier)

Region II $\quad -\frac{\hbar^2}{2m}\frac{\partial^2 \psi}{\partial x^2} = E\psi \rightarrow \alpha = \sqrt{\frac{2m}{\hbar^2}E}$ is real $\rightarrow \psi = A' e^{i\alpha x} + B' e^{-i\alpha x} = A\sin\alpha x + B\cos\alpha x$

Region III $\quad$ N/A (particle-wave cannot exist in infinite barrier)

- ψ must be continuous, so BC are $\psi(x=0) = 0$ and $\psi(x=L) = 0$.

$\psi(x=0) = 0: \quad B = 0$

$\psi(x=L) = 0: \quad \alpha L = n\pi \rightarrow \alpha = \frac{n\pi}{L}, \; n \in \mathbb{Z}^+$

- $\int_{-\infty}^{\infty} |\psi|^2 dx = 1$, so

$$\int_0^L \left[A\sin\left(\frac{n\pi}{L}\right)x\right]^2 dx = 1 \rightarrow A = \sqrt{\frac{2}{L}}$$

$$\therefore \psi = \sqrt{\frac{2}{L}}\sin\left(\left(\frac{n\pi}{L}\right)x\right)$$

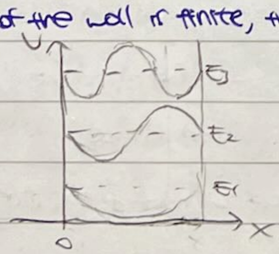
- Applying the kinetic energy operator, we find that

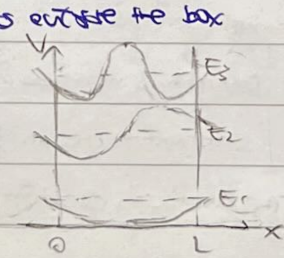
$$T_n\psi = \hat{T}(\psi_n) = -\frac{\hbar^2}{2m}\frac{\partial^2 \psi}{\partial x^2} = -\frac{\hbar^2}{2m}\left(-\frac{n^2\pi^2}{L^2}\right)\psi$$

$$\therefore \boxed{T_n = \frac{n^2\pi^2\hbar^2}{2mL^2}}$$ → energy levels are quantised

(In general, for a wave function w/ wave no. k, its kinetic energy is given by $\boxed{T = \frac{\hbar^2 k^2}{2m} = \frac{p^2}{2m}}$)

- If the height of the wall is finite, the wavefunction ψ leaks outside the box





Heisenberg's uncertainty principle.

- Heisenberg's uncertainty principle states that we cannot measure both position and momentum absolutely accurately. If Δx is the error in determining position, Δp is the error in determining momentum

$$\boxed{\Delta p \Delta x \geq \frac{\hbar}{2}}$$

Equivalently, we cannot measure both energy and time absolutely accurately. If ΔE is the error in determining energy, Δt the error in determining time.

$$\Delta x = v\Delta t \quad ; \quad \Delta E = \frac{dE}{dp}\Delta p = \frac{d}{dp}\left(\frac{p^2}{2m}\right)\Delta p = \frac{2p}{2m}\Delta p = v\Delta p.$$

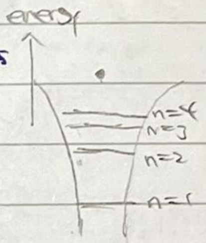
$$\therefore \left(\frac{\Delta E}{v}\right)(v\Delta t) \geq \frac{\hbar}{2}$$

$$\boxed{\Delta E \Delta t \geq \frac{\hbar}{2}}$$

Atomic structure

Atoms

- In an atom, the potential energy term accounts for $e^- - p^+$ and $e^- - e^-$ interactions
- Consider the case of a H atom. The e^- is bound by a coulombic potential well (particle in a box, but the box is not rectangular)



$$F = \frac{mv^2}{r} = \frac{q^2}{4\pi\varepsilon_0 r^2} \rightarrow mv^2 = \frac{q^2}{4\pi\varepsilon_0 r}$$

$$\therefore E = KE + PE = \frac{1}{2}mv^2 - \frac{q^2}{4\pi\varepsilon_0 r} = -\frac{q^2}{8\pi\varepsilon_0 r}$$

The waves must interfere constructively around the orbit

$$2\pi r = n\lambda$$

$$2\pi r = n\frac{h}{p}$$

angular momentum

$$mvr = n\hbar$$

Previously, we find that $mv^2 = \frac{q^2}{4\pi\varepsilon_0 r}$, so $v = \left(\frac{q^2}{4\pi\varepsilon_0 r m}\right)^{1/2}$

$$\therefore mr\left(\frac{q^2}{4\pi\varepsilon_0 r m}\right)^{1/2} = \left(\frac{mq^2}{4\pi\varepsilon_0}\right)^{1/2} r^{1/2} = n\hbar$$

$$r = \frac{n^2\hbar^2}{\left(\frac{mq^2}{4\pi\varepsilon_0}\right)}$$

$$\rightarrow \boxed{E = -\frac{q^2}{8\pi\varepsilon_0 r} = -\frac{mq^4}{32\pi^2\varepsilon_0^2(n\hbar)^2} \propto \frac{1}{n^2}}$$

Chemical bonding.

- C, Si, Ge have 4 valence e^-. All form diamond structure solids. Each atom forms 4 bonds w/ its 4 valence e^-.
- We can also form compounds w/ an average of 4 e^- per atom. For example, in GaAs, Ga has 3 valence e^-, As has 5 valence e^- → same structure as diamond, w/ alternate atoms of Ga and As → Zincblende lattice

* The diamond cubic structure is a FCC lattice w/ half of its interstitial holes filled

Crystal planes

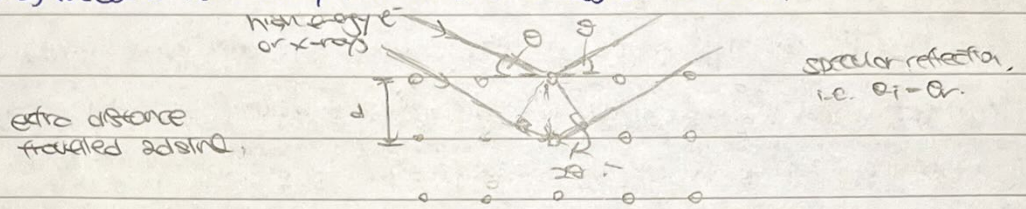
- Crystal planes are defined using Miller indices.
- If the Miller indices are (h,k,l), the plane will cut the x,y,z axis at an integer multiple of $(\frac{1}{h},\frac{1}{k},\frac{1}{l})$ (h,k,l are integers).
- e.g: A plane cuts the x,y,z axis at 2,4,1 respectively.

  Taking reciprocals and multiplying by the LCM of the denominators, we get $\frac{1}{2},\frac{1}{4},1 \xrightarrow{\times 4} 2,1,4$.

  $\therefore$ The plane has Miller indices $(2,1,4)$

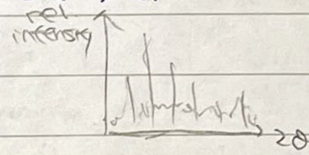
Bragg's Law — electron diffraction

- The diffraction of e^- can be used to identify properties of a crystal:
  - ↳ 1) impinge a beam of high energy e^-/x-rays on a crystal at an angle θ.
  - ↳ 2) measure the intensity pattern of e^- strikes on a screen.



There is constructive interference when the extra distance travelled is $n\lambda$, i.e.

$$\boxed{n\lambda = 2d\sin\theta} \quad \text{[Bragg's Law]}$$

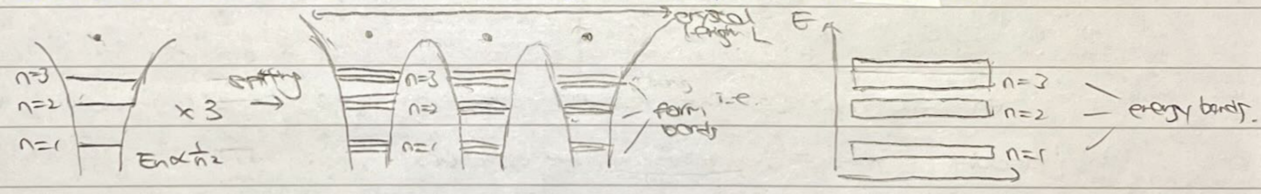


The diffraction pattern has peaks at certain angles → identify d

Conduction in metals and semi-metals

Formation of energy bands

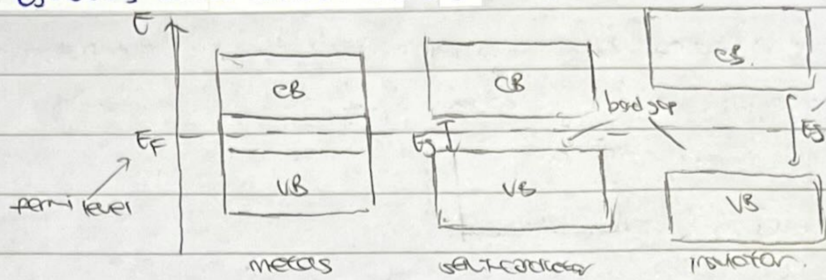
- As shown previously, atoms have discrete energy levels. If we have n identical atoms together, due to Pauli's exclusion principle, each energy level is split into n sub-levels.



- As n becomes very large, the sub-levels become very close to each other → formation of energy bands.

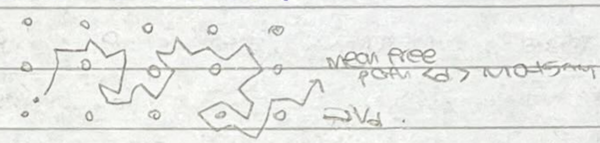
Metals, insulators and semiconductors.

- The Fermi level is the thermodynamic work req. to add one e^- to the solid.
- The valence band (VB) and conduction band (CB) are the bands closest to the Fermi level. It is the energy gap E_g btwn VB and CB that determines the electrical conductivity of the solid.
- Metals: $E_g = 0eV$; semiconductors: $0eV < E_g < 2eV$; insulators: $E_g > 2eV$



Conduction in metal

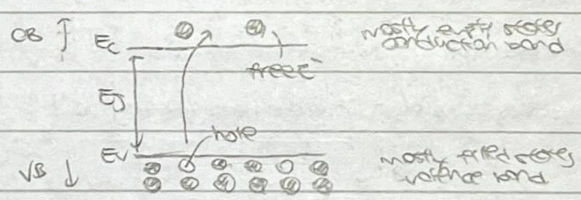
- A metal's outer e^-s are shared btwn bonds and run around btwn all atoms as free e^-s. The free e^-s are available for conduction while core e^- stay fixed to their atoms.
- The free e^- are like a gas confined to the volume of material, and move randomly around, colliding w/ vibrating host atoms w/ their net +ve charge.



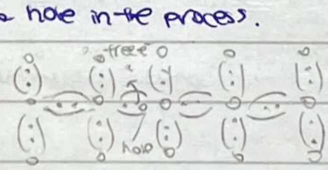
- The conductivity of the metal is determined by the e^- conc. and drift velocity of e^-

Conduction in semiconductors

- Consider the band diagram of a semiconductor



When $T \neq 0K$, some e^- have sufficient thermal energy ($>E_g$) to get promoted from the VB to the CB, becoming a free e^- and leaving a hole in the process.



- The conductivity of the semiconductor is determined by the e^-/h^+ conc (depends on thermal energy and doping) and drift velocity of e^-/h^+

Carrier concentration in a semiconductor

- For an intrinsic semiconductor (pure, undoped) at thermal eqm, the no. of free carriers per unit volume available for conduction is the intrinsic carrier conc. n_i
- Thermal eqm. implies the material is kept in the dark and no electrical bias is applied, i.e. the only source of energy for e^- is thermal due to ambient temp.
- The intrinsic carrier conc. n_i varies w/ temp. as follows:

$$\boxed{n_i \propto T^{3/2} e^{-\frac{E_g}{2k_BT}}}$$

obeys Arrhenius law $\exp(-\frac{Q}{RT})$
→ an activation process

In Si at 300K, $n_i \sim 1.05 \times 10^{10}\ cm^{-3}$.

- At thermal eqm, the no. of e^-/h^+ in the conduction/valence band per unit volume n/p is equal to the intrinsic carrier conc, n_i, i.e.

$$\boxed{n_i = n = p}$$

- e^-/h^+ conc. can be shifted from intrinsic conc. $n/p = n_i$ by doping the semiconductor to get an extrinsic semiconductor. The mass action law states that for any semiconductor,

$$\boxed{np = n_i^2}$$

- N-type doping is adding impurity atoms w/ 5 valence e^- (eg. P) into a Si lattice. 4 e^- participate in bonding and 1 escapes the dopant, and is free to conduct.
- The dopants are donors as they donate e^-. When dopant conc. N_D is much greater than the intrinsic carrier conc n_i ($N_D \gg n_i$), then

$$\boxed{n = n_i + N_D \approx N_D} \rightarrow \boxed{p = \frac{n_i^2}{n} = \frac{n_i^2}{N_D}}$$

- P-type doping is adding impurity atoms w/ 3 valence e^- (eg. B) into a Si lattice. 4 e^- are needed to participate in bonding so 1 is pulled in from the lattice → a hole escapes the dopant and is free to conduct.
- The dopants are acceptors as they accept e^-. When dopant conc. N_A is much greater than the intrinsic carrier conc. n_i ($N_A \gg n_i$), then

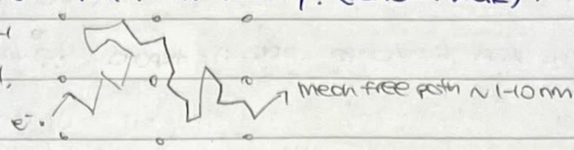
$$\boxed{p = n_i + N_A \approx N_A} \rightarrow \boxed{n = \frac{n_i^2}{p} = \frac{n_i^2}{N_A}}$$

* Typically, at room temp, $n_i \sim 10^{10}\ cm^{-3}$, $N_A/N_D \sim 10^{15}\ cm^{-3}$,

Velocity of carriers

- When an electric field E is applied, e^-/h^+ start moving to the +ve/-ve electrode.
- In a semiconductor/metal, the charge carriers cannot accelerate indefinitely, but rather accelerate then immediately scatter (collide w/ lattice / defects / each other). → the motion is intermittent bursts of speed and fall to low velocity. (random walk).

$\langle v \rangle \sim 10^6 ms^{-1}$
$v_d \sim 10^{-4} ms^{-1}$



- At low electric fields E, the drift velocity v_d is prop. to the electric field E

$$\boxed{v_d = \mu E}$$

where μ is the mobility.

* Note that increasing dopant conc. N_A/N_D increases the no. of defects, hence effects of scattering, thus reducing the mean free path and drift velocity v_d.

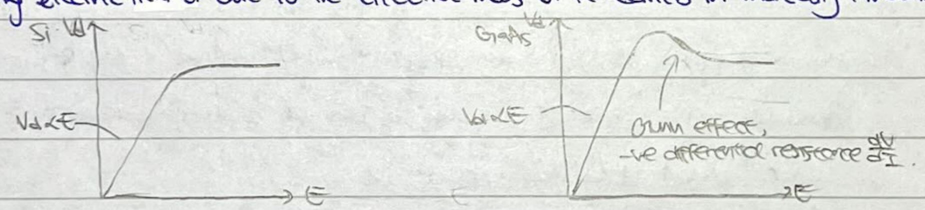
- At low electric fields E, the drift velocity of e^-/h^+, v_{dn}/v_{dp} is given by

$$\boxed{v_{dn} = \mu_n E} \qquad \boxed{v_{dp} = \mu_p E}$$

- Typically, in Si, $\mu_n = 1500 cm^2 V^{-1} s^{-1}$; $\mu_p = 500 cm^2 V^{-1} s^{-1}$.

- At high electric fields, E, the drift velocity depends on the semiconductor. In most semiconductors inc. Si, the velocity saturates at high fields.

- However, in GaAs, we observe the Gunn effect — the velocity initially increases, then reduces w/ increasing electric field E due to the effective mass of the carriers m^* increasing & mobility μ decreasing.
  due to inter-valley transfer in E-k diagram



Mobility and effective mass.

- In a crystal lattice, the charge carrier interacts w/ the lattice — it polarises the neighbouring atoms and moves through the lattice. To use Newton's laws to model the charge carrier, we must use its effective mass m^* instead of its rest mass m.

- consider a charge carrier moving through a crystal lattice due to an electric field.

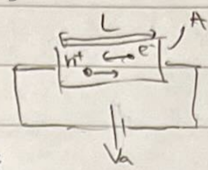
$$\frac{dp}{dt} = \frac{m^* v_d}{\tau} = qE = F$$

$$v_d = \frac{q\tau}{m} E \rightarrow \boxed{\mu = \frac{q\tau}{m}}$$

where τ is the mean time btwn collisions

- The effective mass m^* is related to the curvature of the band on a E-k plot.

$$E = \frac{p^2}{2m^*} = \frac{\hbar^2 k^2}{2m^*} \rightarrow \frac{\partial^2 E}{\partial k^2} = \frac{\hbar^2}{m^*} \rightarrow \boxed{m^* = \frac{\hbar^2}{\frac{\partial^2 E}{\partial k^2}}}$$



Drift current in a semiconductor

- The drift current density J_{dr} is given by

$$J_{dr} = qnv_{dn} + qpv_{dp} = qn\mu_n E + qp\mu_p E = q(n\mu_n + p\mu_p)E$$

By defn of current density,

$$I = J_{dr}A = q(n\mu_n + p\mu_p)AE$$

Relating E and V using $E = \frac{dV}{dx}$,

$$I = q(n\mu_n + p\mu_p)A\frac{dV}{dx}$$

$$\int_0^L I\,dx = \int_0^{V_a} q(n\mu_n + p\mu_p)A\,dV$$

$$V_a = \frac{1}{q(n\mu_n + p\mu_p)}\frac{L}{A}I$$

$$\therefore \boxed{R = \frac{V_a}{I} = \frac{1}{q(n\mu_n + p\mu_p)}\frac{L}{A}} \rightarrow \boxed{\rho = \frac{A}{L}R = \frac{1}{q(n\mu_n + p\mu_p)}}$$

- If the semiconductor is intrinsic, $n_i = n = p$,

$$\boxed{\rho = \frac{1}{q(n\mu_n + p\mu_p)} = \frac{1}{qn_i(\mu_n + \mu_p)}}$$

If the semiconductor is doped w/ donor dopants of conc. N_D, $N_D \gg n_i \rightarrow n \approx N_D$, $p \approx \frac{n_i^2}{N_D}$ $(n \gg p)$

$$\boxed{\rho = \frac{1}{q(n\mu_n + p\mu_p)} \approx \frac{1}{q(n\mu_n)}}$$

If the semiconductor is doped w/ acceptor dopants of conc. N_A, $N_A \gg n_i \rightarrow p \approx N_A$, $n \approx \frac{n_i^2}{N_A}$ $(p \gg n)$

$$\boxed{\rho = \frac{1}{q(n\mu_n + p\mu_p)} \approx \frac{1}{q(p\mu_p)}}$$

MOSFETs, TFTs and displays

Electrostatics in semiconductors

Poisson's equation.

- starting from the differential form of Gauss' law for electrostatics,

$$\nabla \cdot E = \frac{\rho}{\varepsilon}$$

Denoting the potential as ϕ, and using $E = -\nabla\phi$, we have

$$\nabla^2 \phi = -\frac{\rho}{\varepsilon}$$

In 1D, this simplifies to

$$\boxed{\frac{d^2\phi}{dx^2} = -\frac{\rho}{\varepsilon}}$$

Electrostatics in insulators.

- Consider an insulator in a // plate capacitor. The insulator has no charges, one end ($x=0$) is at potential V_a while the other end ($x=L$) is at 0 potential.

Gauss law: $\frac{dE}{dx} = \frac{\rho}{\varepsilon} = 0 \rightarrow E = C_0 = \text{const}$

$-\frac{d\phi}{dx} = E \rightarrow \phi = C_0 x + C_1$

BC: $\phi(x=0) = V_a$, $\phi(x=L) = 0 \rightarrow E = \frac{V_a}{L}$, $\phi = V_a(1 - \frac{x}{L})$

Electrostatics in semiconductors

- Generally in semiconductors, there are two types of charges.
  - ↳ 1) mobile charges — e^-, h^+.
    They will respond to any E field and move about. Therefore their conc. will depend on the potential in the semiconductor
  - ↳ 2) fixed charges — dopant ions donor$^+$ / acceptor$^-$
    The n-type donor / p-type acceptor dopant has given up its e^-/h^+ and therefore is +ve/-ve charges.
    donor → donor$^+$ + e^- / acceptor → acceptor$^-$ + h^+
- The semiconductor is still neutral after doping.
- Sometimes, a region in a semiconductor might be depleted — it is nearly void of carriers. If the depletion region is doped, the only charges present will be the dopant ions.
- consider a p-doped semiconductor w/ acceptor dopant conc. N_A. The region from $x=0$ to $x=d$ in the semiconductor is depleted.

depletion zone $n = p = 0$

charge density: $\rho = q(\cancel{N_D}^{0} + \cancel{p}^{0} - N_A - \cancel{n}^{0}) = -qN_A$

not n-type

Gauss law: $\frac{dE}{dx} = \frac{\rho}{\varepsilon} = -\frac{qN_A}{\varepsilon} \rightarrow E = -\frac{qN_A}{\varepsilon}x + C_0$

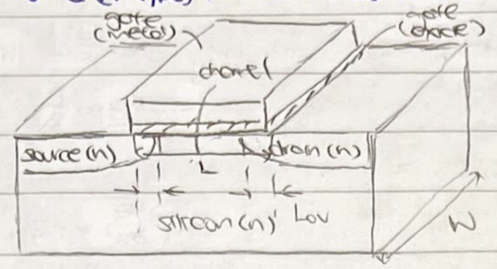
$-\frac{d\phi}{dx} = E \rightarrow \phi = \frac{qN_A}{\varepsilon}\frac{x^2}{2} - C_0 x + C_1$ $\quad d \propto \sqrt{\phi}$

BC: $E(x=d) = 0 \rightarrow E = \frac{qN_A}{\varepsilon}(d-x)$, $\phi_{0 \to d} = -\frac{qN_A}{\varepsilon}\frac{d^2}{2}$

MOSFETs

MOSFET

- A MOSFET has a gate (G), source (S), drain (D) and body contact (connects to the bulk Si)
- The physical construction of a (n-type) MOSFET is as follows



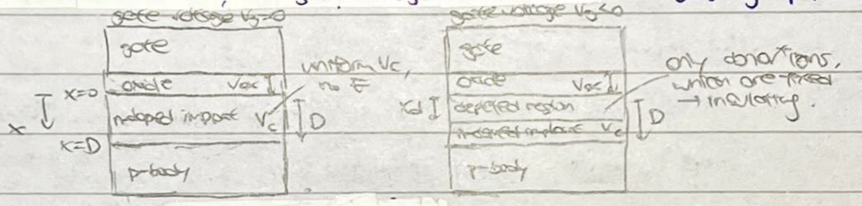
L is the channel length, W is the channel length, L_{ov} is the overlap length, $\frac{W}{L}$ is the aspect ratio

- The gate voltage V_g controls the resistance of the channel via field effect. The transistor is ON/OFF when the channel has low/high resistance.
- For depletion mode MOSFET, a -ve V_T is req. at G to turn the transistor OFF (o/w it is ON)

  For enhancement mode MOSFET, a +ve V_T is req. at G to turn the transistor ON (o/w it is OFF)

Depletion-mode MOSFET calculation.

- Consider a n-type depletion-mode MOSFET. Assume the entire n-doped implant is at potential V_c, and a voltage V_{ox} is dropped across the oxide. Let the thickness of the depletion region be x_d. When full depletion occurs ($x_d = D$), the gate voltage is the pinchoff voltage, $V_g = V_p$.



Consider the depletion region

charge density: $\rho = q(N_D + \cancel{p}^0 - \cancel{N_A}^0 - \cancel{n}^0) = qN_D$

Gauss Law: $\frac{dE}{dx} = \frac{\rho}{\varepsilon} = \frac{qN_D}{\varepsilon} \rightarrow E = \frac{qN_D}{\varepsilon}x + C_0$

$$-\frac{d\phi}{dx} = E \rightarrow \phi = -\frac{qN_D}{\varepsilon}\frac{x^2}{2} - C_0x + C_1$$

BC: $E(x=x_d) = 0 \rightarrow E = \frac{qN_D}{\varepsilon}(x - x_d)$, $\phi_{0\to d} = \frac{qN_D}{\varepsilon}\frac{x_d^2}{2}$

BC: $\phi(x=x_d) = V_c \rightarrow \phi(x=0) = \phi(x=x_d) - \phi_{0\to d} = V_c - \frac{qN_D}{\varepsilon}\frac{x_d^2}{2}$

$$\therefore \phi(x=0) = \boxed{V_g - V_{ox} = V_c - \frac{qN_D}{\varepsilon}\frac{x_d^2}{2}}$$

Full depletion when $x_d = D$, $V_g = V_p$

$$\therefore \boxed{V_p = V_c + V_{ox} - \frac{qN_D}{\varepsilon}\frac{D^2}{2}}$$

* Actually, the thickness of the depletion region is not uniform over the channel length since $V_{gd} < V_{sd}$ (calculation beyond the scope of this course).

Transit time and maximum current

- At low fields, $v_d = \mu E$.

  At high fields, velocity saturates and equals the scattering limited velocity $v_d = v_s = 10^5 ms^{-1}$ (for Si)

- The transit time t is then given by

$$\boxed{t = \frac{L}{v_d}}$$

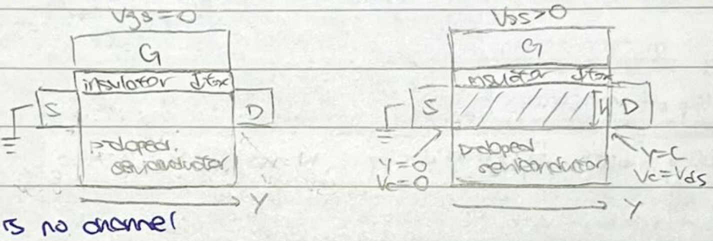
and the max. (saturation current is given by

$$\boxed{I_{sat} = \frac{Q}{t} = \frac{qN_D(WDL)}{t} = qN_D(WD)v_s}.$$

Thin film transistors

Thin film transistors (TFTs)

- TFTs are built w/ depositing thin films on substrates (e.g. glass, plastic), and are the building blocks of large area electronic systems (e.g. displays)
- The I-V characteristics of TFTs are similar to enhancement mode MOSFETs, but w/ some diff. due to the nature of the material. (we will ignore these diff. in this course)



- At $V_{gs}=0$, there is no channel
- At $V_{gs} > V_T$ (threshold voltage), there is a conductive channel of e^- btwn source and drain. when a channel forms, it has a channel potential $V_c(y)$ that varies from source to drain
- If we imagine the channel (sheet of e^-) to have a thickness h, the current through the channel I_{ds} is

$$\frac{I_{ds}}{Wh} = -qnv_{dn} = -qn\mu_n E = qn\mu_n \frac{dV_c}{dy}$$

- If we define the insulator capacitance per unit area as $C_{ox} = \frac{\varepsilon_{ox}}{t_{ox}}$,

$$Q = qn(LWh) = C_{ox}(LW)(V_{gs}-V_T-V_c)$$

$$qn = \frac{C_{ox}(V_{gs}-V_T-V_c)}{h}$$

$$\therefore I_{ds} = qn\mu_n \frac{dV_c}{dy}(Wh) = \mu W C_{ox}(V_{gs}-V_T-V_c)\frac{dV_c}{dy}$$

$$\int_0^L I_{ds}\,dy = \int_0^{V_{ds}} \mu W C_{ox}(V_{gs}-V_T-V_c)\,dV_c$$

$$\boxed{I_{ds} = \frac{\mu W C_{ox}}{L}\left[(V_{gs}-V_T)V_{ds} - \frac{V_{ds}^2}{2}\right]} \quad V_{ds} < V_{gs}-V_T$$

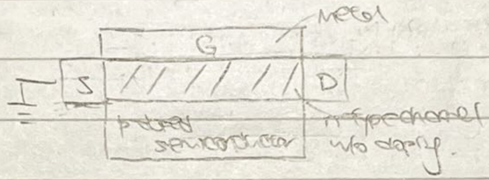
The above operation is only true if $V_{ds} < V_{gs}-V_T$ (linear operation) → resistor

- If $V_{ds} \geq V_{gs}-V_T$, there will be a region near the drain where the channel cannot exist ("pinch off") The current saturates and becomes independent of V_{ds} (saturation operation) → gate-controlled current source.
- To calculate I_{ds} in saturation operation, we substitute $V_{ds} = V_{gs}-V_T$

$$\boxed{I_{ds} = \frac{\mu W C_{ox}}{2L}(V_{gs}-V_T)^2} \quad V_{ds} \geq V_{gs}-V_T$$

Metal semiconductor field effect transistor (MESFET).

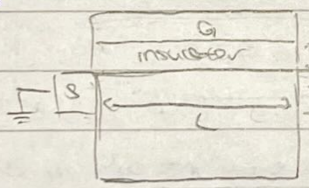
- The key diff. btwn MESFET and MOSFET is that there is no insulator btwn gate metal & semiconductor
- When a metal (M) interfaces w/ a semiconductor (S), there are two types of contact to be made:
  ↳ (i) ohmic contact ↳ (ii) Schottky / rectifying contact.
- An ohmic contact will ensure the M-S junction behaves like a resistor — there is no charge transfer, this junction is used if we want to inject current from metal to semiconductor, and vice versa.
- A Schottky junction forms a diode-like contact — a depletion region forms where both the metal and semiconductor are depleted of free carriers.
- A MESFET uses a gate metal that forms a Schottky contact w/ the semiconductor, the resulting depletion region can be controlled to set the source-drain current.
- The MESFET is often used for high freq. operations (e.g. radar) since there is no doping of semiconductors, which would o/w introduce dopant atoms, thus defects, which lowers the mobility μ.
- A disadvantage of the MESFET is that we can't turn it off (like a MOSFET).

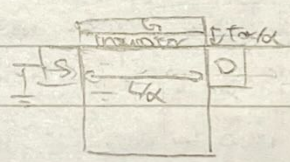


Scaling and material of transistors

Scaling of transistors

- Scaling implies that we are reducing L, W, t_{ox} by a factor of α.
  (Practically, scaling down t_{ox} increases tunneling current → ↑ε_r and keep t_{ox} const)





- There are two methods for scaling:
  ↳ (i) Constant field scaling.
  Keep field const → scale down applied voltage → voltage levels not back-compatible
  ↳ (ii) Constant voltage scaling
  Keep voltage const → field becomes large → may cause breakdown
- The following quantities are scaled as follows (L, W, t_{ox} scaled by $\frac{1}{\alpha}$)

Parameter	field	voltage	gate capacitance	Oxide capacitance	current	power
symbol	$\mathcal{E}$	V	C_G	C_{ox}	I	P
const-field scaling	1	$\frac{1}{\alpha}$	$\frac{1}{\alpha}$	α	$\frac{1}{\alpha}$	$\frac{1}{\alpha^2}$
const-voltage scaling.	α	1	$\frac{1}{\alpha}$	α	α	α

Materials of transistors

- Si : Best for general purpose MOSFET

  It is a v. well established technology → low cost

  Easy to get high purity wafers w/o defects

- GaAs : Very good for high frequency applications → MESFETs

  Gunn effect → negative differential resistance → circuits w/ high Q factors

  crystals are quite defective

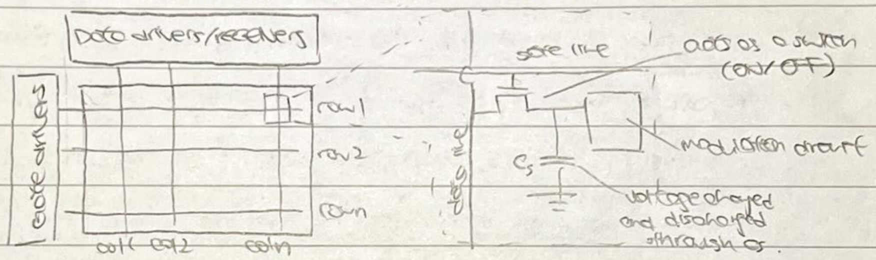
- GaN : very high breakdown voltage → high voltage / high power / high temp applications.

  stable wide band gap semiconductor → optoelectronic materials for short wavelength.

Displays

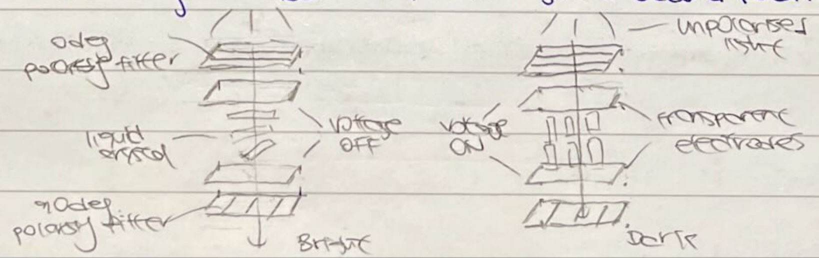
Active matrix displays

- An active matrix display consists of pixel control transistors + light modulators.
- The architecture for an active matrix display works as follows:



Liquid crystals (LC)

- LC is a display element based on modulating ext light (backlight)
- LCs are a phase of matter btwn liquid and crystal.
- When the LC is placed along two "aligning guide" crystals, it will direct itself along the guide. We can use this to obtain a twist btwn two ⊥ polarisers.
- When an E-field is applied, it reorients the LCs to no longer twist the light.
- A backlight is modulated as follows:
  - ↳ Voltage OFF : unpolarised backlight → polarised to 0 deg → slowly twisted by the LC to polarised 90 deg. → passes through the second polariser
  - ↳ Voltage ON : unpolarised backlight → polarised to 0 deg → not twisted by the LC and remains at 0 deg → does not pass through the second polariser



E-ink

- E-ink is a display element based on modulating ext. light (reflected ambient light)
- E-ink consists of +ve/-ve charged white/black pigments enclosed in transparent microcapsules, which are sandwiched btwn transparent conductive electrodes.
- Depending on the field, we have a diff. proportion of white/black particles on the top, which implies poor/good reflection of ambient light.

Light-emitting diodes (LED)

- LED is a display element based on generating light
- LEDs use p-n junction diodes. When the diode is forward biased, the built-in potential barrier is lowered and carriers can diffuse across the junction → diffusion current.

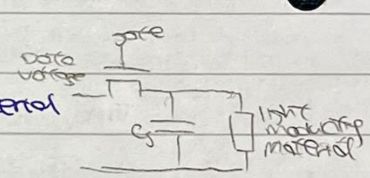
$$I = I_0\left(e^{\frac{qV_B}{k_BT}} - 1\right)$$

(transport of charge due to a conc. gradient, i.e. diffusion)

- The diffusion of e^-/h^+ from n/p to p/n side leads to an increase in recombination of e^- and h^+, producing light → light intensity depends on the current.
- Direct bandgap semiconductors are preferred.

Pixel circuit

① Voltage programmed pixel circuit (LCD, E-ink)



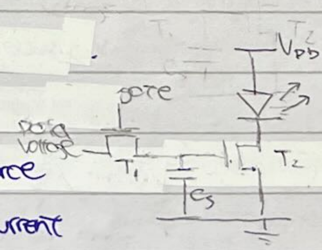
- 1 TFT needed, used as a switch
- The "Data voltage" is the voltage req. to modulate the light-modulating material
- The resistance of the switch is decided by the gate voltage V_s.
- The switch operates in linear bias (resistor) → R_{TFT}

$$\boxed{R_{TFT} = \frac{dV_{ds}}{dI_{ds}} \approx \frac{1}{\mu C_{ox}(W/L)(V_{gs}-V_T)}}$$

in linear mode, $V_{gs}-V_T >> V_{ds}$.

The time const. is therefore $\tau = R_{TFT}C_s$

② Current programmed pixel circuit (LED)



- 2 TFTs needed, T1 used as a switch, T_2 used as a current source
- The "Data voltage" is the gate voltage req. for T_2 to drive the right current
- T1 operates in linear bias (resistor); T_2 operates in saturation mode (voltage controlled current source)

$$I_{ds} = \frac{\mu C_{ox}(W/L)}{2}(V_{gs}-V_T)^2$$