

PN Junction(PN 结)

半导体制造(Semiconductor Manufacturing Process)



Figure 1: 半导体制作过程

一些名词的解释

- Oxidation: 为了保护晶片(wafer)
- Photolithography(光刻):在晶片上画电路
- Etching(蚀刻):完成电路形状的刻画
- Deposition & Ion Implantation(离子注入与沉积):改变电学性质
- Metal Wiring:加用金属连接
- EDS:测试是否达到标准

pn 结的形成

浓度梯度所造成的漂移

由于电子和空穴分别在 n 和 p 区域内浓度大, 因此向相反的类型区域飘逸, 与相反电荷的载流子结合(recombination), 形成了 Depletion Region, 或者叫做 Space Charge, 这两个术语是一个意思, 可以认为这个区域内没有载流子. 同是, 参杂的离子比如 P^- 或者 B^+ 不能动的, 没有载流子了, 就单独在那里, n 区的 P^- 还有 p 区的 B^+ , 所以耗尽区又叫做空间电荷区(Space Charge region)

Build-in voltage

由于漂移过去的载流子, 形成了内建电场.

$$V_{bi} = V_t \ln \left(\frac{N_A N_D}{n_i^2} \right) \quad 1.$$

$$V_t = \frac{kT}{q} (\text{热电压}) \quad 2.$$

- k:玻尔兹曼常数
- T:热力学温度
- q:元电荷
- N_A :P 型半导体的参杂浓度

- N_B : N 型半导体的参杂浓度
- Intrinsic Carrier Concentration: 本征载流子浓度(没有参杂的情况下载流子的浓度)

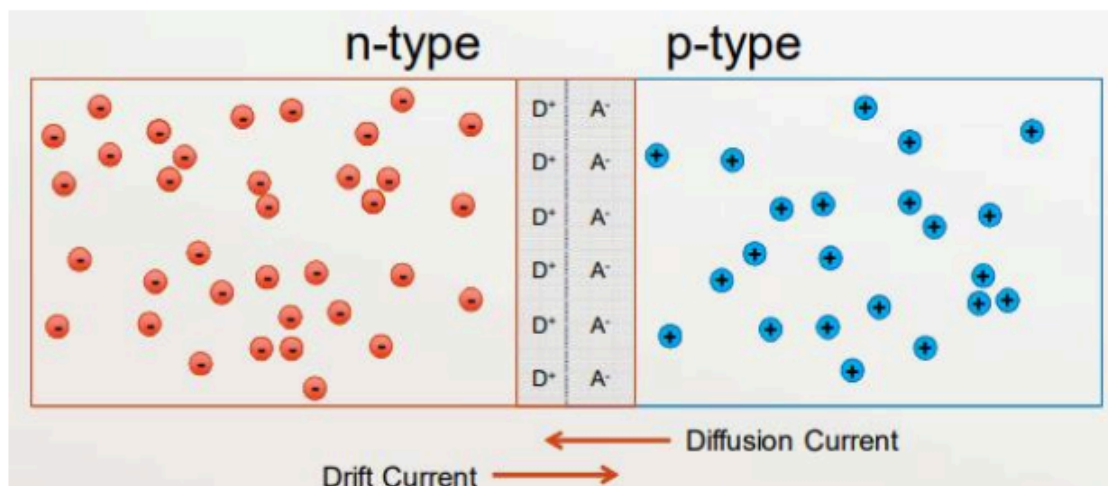


Figure 2: 内建电场

耗尽区的宽度

$$W = \sqrt{\frac{2\epsilon_0\epsilon_r V_{bi}}{q} \left(\frac{1}{N_A} + \frac{1}{N_D} \right)} \quad 3.$$

$\epsilon_0\epsilon_r$ 是主体材料的介电常数, $\epsilon_0 = 8.854 \times 10^{-14} F/m$, 是真空中介电常数.

可以用两边的电荷平衡得到 $N_A W_p = N_D W_n$, 也就是说浓度大的一边耗尽区的宽度比较小, 浓度更高的区域又叫做 heavily doped regions, p^+ 表示 p 型重掺杂区, n^+ 表示 n 型重掺杂区.

pn 结的偏置(biasing)

偏置就是外加电压, 效果是改变了耗尽区的厚度

- zero bias: 没有外接电压
- Forward bias: 正电压接 p 区域, 负电压接 n 区域, 会使得耗尽区的厚度变小
- Reverse bias: 反接, 会使得耗尽区的厚度变大

偏置相当于在作用在内建电压上面, 所以

$$W = \sqrt{\frac{2\epsilon_0\epsilon_r(V_{bi} - V_{bias})}{q} \left(\frac{1}{N_A} + \frac{1}{N_D} \right)} \quad 4.$$

电流公式

$$I = I_s \left(e^{\frac{qV}{nkT}} - 1 \right) = I_s \left(e^{\frac{V}{nT}} - 1 \right) \quad 5.$$

- I_s : saturation current(饱和电流)
- q :元电荷
- V :加在两端的偏置电压
- n 理想因子, 理想的情况下是 1
- k :玻尔兹曼常数
- T :热力学温度

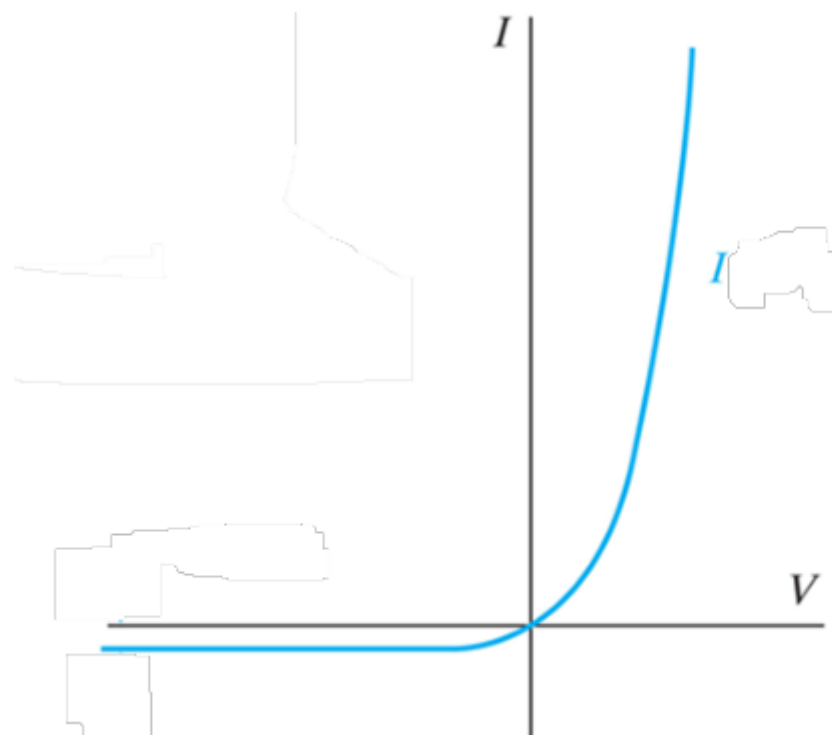


Figure 3: I_V 图

Junction Capacitance(结电容)

$$C = \frac{\epsilon A}{W} \quad 6.$$

W 是厚度, A 是 p-n 结的有效面积, $\epsilon = \epsilon_0\epsilon_r$,是主体材料的介电常数.

载流子的迁移模式

- Drift: 由于电场导致的: 电场产生的速度和载流子自由移动的速度叠加, 在导体中可以认为又一个固定的速度 V_d
- Diffusion: 由于浓度梯度导致的/由于温度梯度导致的.

Drift:

没有阻力的时候:

$$F = qE \quad 7.$$

$$F = ma \quad 8.$$

Drift 的速度: $v_d = \mu E$, 其中 μ 叫做“载流子移动能力 (mobility of the charge carriers)”. 电子和空穴的分别表示为 μ_e, μ_h . 由温度, 参杂浓度和其他因素所决定.

Drift 电流的密度 (J : A/m^2)

$$J = qnv_d = qn\mu E \quad 9.$$

n 是载流子的体积密度, 如果是正载流子, 也可以记作 p . p-n 结中, 两种载流子都有, 要分别算 J 然后相加:

$$J = J_n + J_p \quad 10.$$

$$\sigma(\text{conductivity}) = \frac{l}{s} \frac{I}{El} = q(n\mu_n + p\mu_p) \quad 11.$$

$$\rho = \frac{1}{\sigma} \quad 12.$$

Diffusion

$$J_n = -qD_n \frac{dn}{dx} \quad 13.$$

$$J_p = qD_p \frac{dp}{dx} \quad 14.$$

D_n, D_p : diffusion coefficients

$\frac{dn}{dx}$ 代表了浓度的梯度 (对距离微分)

$$D_n = \mu_n \frac{kT}{q} \quad 15.$$

$$D_p = \mu_p \frac{kT}{q} \quad 16.$$

$$J = J_n + J_p = J_{n,drift} + J_{n,diffusion} + J_{p,drift} + J_{p,diffusion}$$

Figure 4: 总体电流密度

电子和空穴的密度计算

要涉及到之前的 density of states, probability of a state being occupied, 对于 n 型, 在导带上积分, 对于 p 型, 在价带积分.

$$n_n = \int_{E_c}^{E_c+\chi} g_{cb}(E) f(E) dE \quad 17.$$

用能带理论解释 PN 结的性质

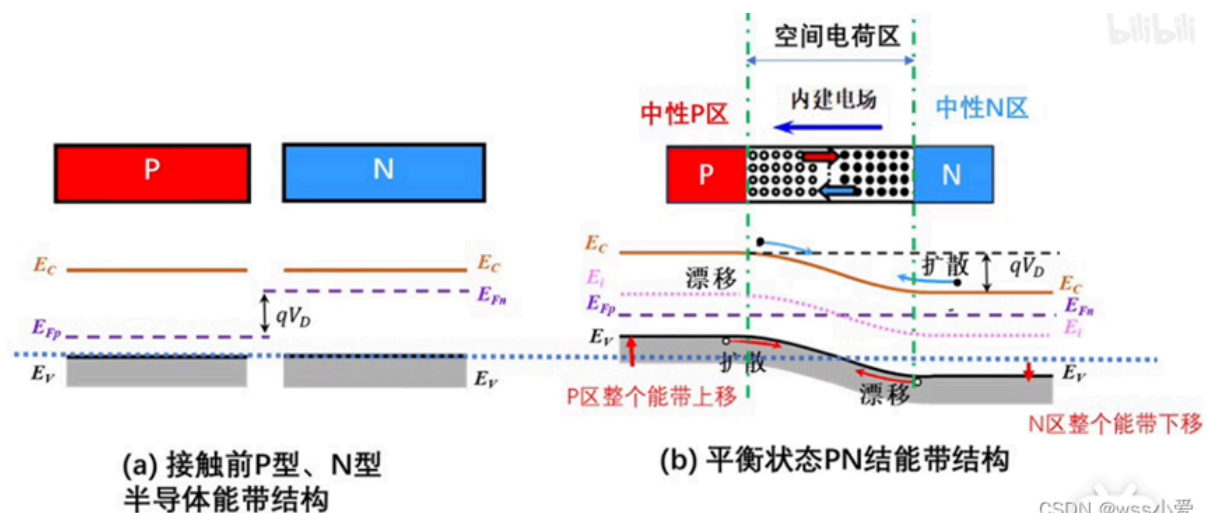


Figure 5: 能带理论

由于 n 型的费米能级更加靠近导带, 可以认为 n 区能量高, 在 n 区导带中的电子会流向 p 区的导带, 流过来的电子由于能量下降, 很快与价带上的空穴复合, 同时 P 区能带上移, N 区能带下降, 当两者的费米能级达到同一水平时, N 区和 P 区之间形成势垒, 势垒区内存在内建电场, N 区电子由于能量不足以跨过势垒, 停止移动。

在反向偏置的情况下, 结的厚度变大, 但是有少子还在移动, 产生了几乎为常数的 反向饱和电流.

Thermal generation

只要温度大于 0k, 就一定有价电子被激发到导带, 留下空穴在价带, 产生了电子-空穴对.

击穿(Reverse breakdown)

击穿有两种

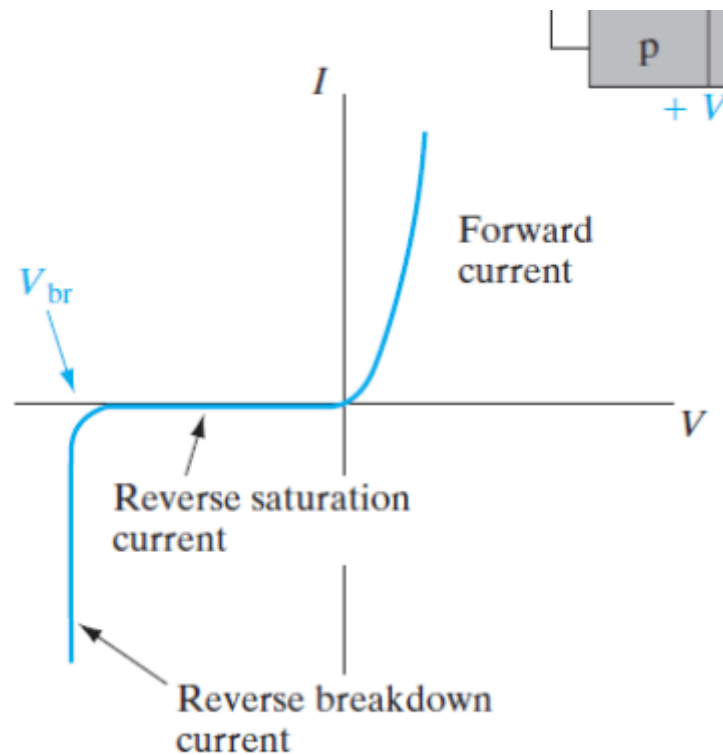


Figure 6: 完整的 i-v 图

- Avalanche(雪崩)

由于在耗尽区的电场特别大, 导致 Drift 电子得到足够的动能让硅原子离子化, 破坏 Si-Si 共价键, 是一种轰击(bombardment)的过程, 这种离子化又叫做: impact ionization(冲击离子化)

这句话描述的是**冲击电离 (Impact Ionization)** 过程中的**雪崩效应 (Avalanche Effect)**。

具体解释如下：

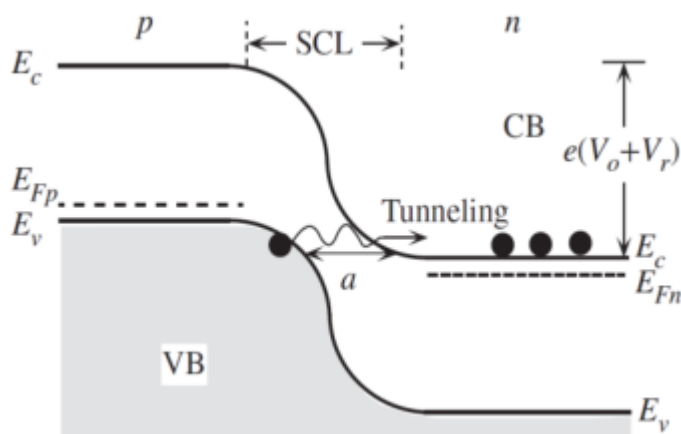
1. **EHPs (Electron-Hole Pairs, 电子-空穴对)**：电子-空穴对是半导体中由于外部激发（如光照、热激发或电场作用）产生的电子和空穴的配对。电子通常会逃离原来的位置，而空穴则是缺失的电子的位置。
2. **冲击电离 (Impact Ionization)**：当电子在电场中加速到足够高的能量时，它们可能会与晶格中的原子或分子发生碰撞，从而将这些原子的电子打出，形成新的电子-空穴对。这就是所谓的冲击电离。
3. **雪崩效应 (Avalanche Effect)**：在冲击电离过程中，产生的电子（EHPs）会被电场加速到更高的能量。这些高速电子与其他原子发生碰撞，进一步产生更多的电子-空穴对。这个过程是自增强的，也就是说，每次碰撞都会生成更多的EHPs，形成一个类似雪崩一样迅速扩展的过程。

所以，整句话的意思是：**冲击电离产生的电子-空穴对会在电场的作用下被加速，并通过碰撞产生更多的电子-空穴对，从而形成一个持续增长的雪崩效应。**

这种现象常见于半导体中的雪崩二极管 (Avalanche Diodes) 和某些光电探测器等器件中，通常会在强电场或高能量条件下发生。

• Zener(齐纳)

由于反向偏置电压, 肯会导致 n 区导带的能量界限比 p 区导带的能量界限低. 这样, p 区价带的电子可以快速的迁移到 n 区的导带. 这也是齐纳二极管的原理.



如何计算饱和电流 I_s

$$I_s = qA \left(\frac{D_p}{L_p} p_n \frac{D_n}{L_n} n_p \right) \quad 18.$$

D_p, D_n : Equation 16, 指的是空穴和电子的扩散速度

q : 元电荷

A : PN 结的面积

L_p, L_n : 空穴和电子的扩散长度

p_n, n_p : N 区空穴浓度, P 区电子浓度(少子的浓度)

在半导体平衡状态下, 满足多子浓度 \times 少子浓度 $= n_i^2$ 对于硅来说,

$n_i = 1.5 \times 10^{10}$

例子

An abrupt Si p-n junction ($A = 10^{-4} \text{ cm}^2$) has the following properties at 300 K:

<i>p side</i>	<i>n side</i>
$N_a = 10^{17} \text{ cm}^{-3}$	$N_d = 10^{15}$
$\tau_n = 0.1 \mu\text{s}$	$\tau_p = 10 \mu\text{s}$
$\mu_p = 200 \text{ cm}^2/\text{V}\cdot\text{s}$	$\mu_n = 1300$
$\mu_n = 700$	$\mu_p = 450$

The junction is forward biased by 0.5 V. What is the forward current? What is the current at a reverse bias of -0.5 V?

Figure 7: Example 1

$$I_s = qA \left(\frac{D_p}{L_p} p_n + \frac{D_n}{L_n} n_p \right) \quad 19.$$

- 先算少子的浓度 N_a, N_d 是掺杂浓度, 可以近似为多子的浓度
- $p_n = \frac{(1.5 \times 10^{10})^2}{10^{15}} = 2.25 \times 10^5 \text{ cm}^{-3}$
- $n_p = \frac{(1.5 \times 10^{10})^2}{10^{17}} = 2.25 \times 10^3 \text{ cm}^{-3}$
- $D_p = \frac{kT}{q} \mu_p = 11.66 \text{ cm}^2/\text{s}$
- $D_n = \frac{kT}{q} \mu_n = 18.13 \text{ cm}^2/\text{s}$
- τ_n 指 p 区少子(电子)的寿命, τ_p 指 n 区少子(空穴)的寿命
- $L_p = \sqrt{D_p \tau_p} = 1.08 \times 10^{-2} \text{ cm}$
- $L_n = \sqrt{D_n \tau_n} = 1.35 \times 10^{-3} \text{ cm}$
- $I_0 = 4.370 \times 10^{-15} \text{ A}$
- $I_+ = I_0 \left(e^{\frac{V}{V_T}} - 1 \right) = 1.058 \times 10^{-6} \text{ A}$
- $I_- = -I_0 = -4.370 \times 10^{-15} \text{ A}$ (直接用饱和电流近似反向偏置电流)

An abrupt Si p-n junction has $N_a = 10^{18} \text{ cm}^{-3}$ on one side and $N_d = 5 \times 10^{15} \text{ cm}^{-3}$ on the other.

- Calculate the Fermi level positions at 300 K in the p and n regions.
- Draw an equilibrium band diagram for the junction and determine the contact potential V_0 from the diagram.
- Compare the results of part (b) with V_0 as calculated from Eq. 20.1

Figure 8: Example 2

(a) 求 **Fermi level position**:指的是当前费米能级和没有参杂的(本征费米能级)的差

在 pn 结中, p 型的能量更高:

$$E_{Fp} - E_i = kT \ln\left(\frac{N_a}{n_i}\right) = 0.467 \text{ eV} \quad 20.$$

$$E_{Fn} - E_i = -kT \ln\left(\frac{N_d}{n_i}\right) = -0.329 \text{ eV} \quad 21.$$

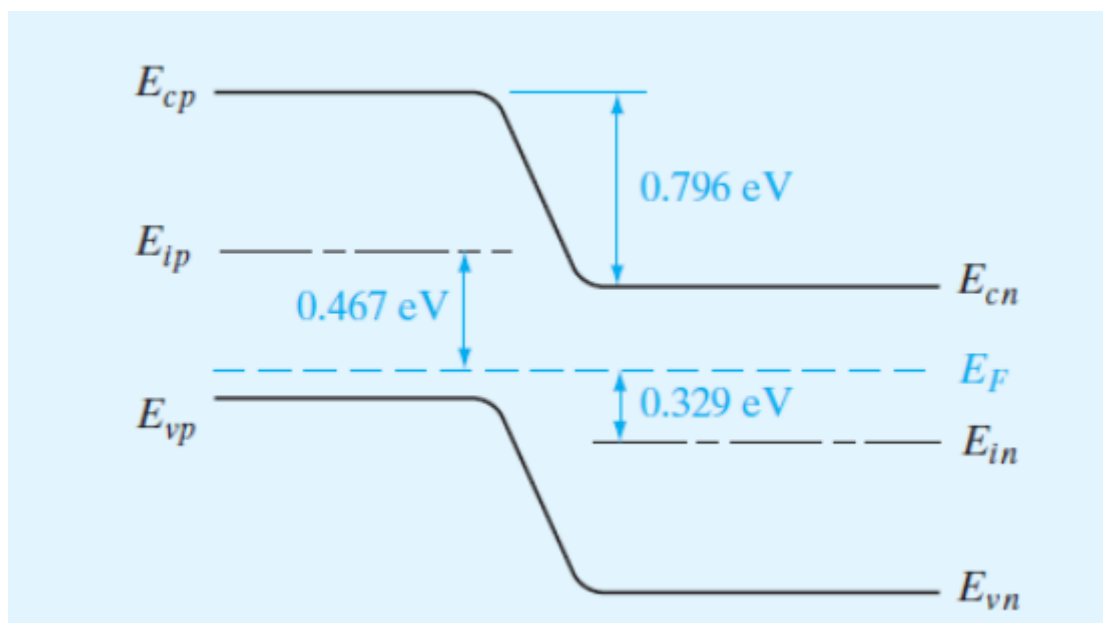


Figure 9: pn 结能级图

注意:带间隙 $E_c - E_v$ 是由主体材料本身决定的, 和参杂浓度无关

$$0.329 = E_F - \left(\frac{E_{cn} + E_{vn}}{2}\right) \quad 22.$$

$$0.467 = \frac{E_{cp} + E_{vp}}{2} - E_F \quad 23.$$

$0.329 + 0.467 = \frac{E_{cp} - E_{cn} + E_{vp} - E_{vn}}{2} = E_{cp} - E_{cn} = 0.769\text{eV} = qV_0$, 这个值又叫做接触电势(contact potential)

把两个式子加起来可以得到 $qV_0 = kT \frac{N_a N_d}{n_i^2} = 0.796 \text{ eV}$

A Si sample is doped with 10^{17} As atoms/cm³. What is the equilibrium hole concentration p_0 at 300 K? Where is E_F relative to E_i ?

Figure 10: Example 3

近似认为多子浓度就是参杂浓度:

$$p_0 = \frac{n_i^2}{n_0} = 2.25 \times 10^3 \quad 24.$$

$$E_F - E_i = kT \ln \frac{n_0}{n_i} = 0.407V \quad 25.$$

Appendix

首先要澄清一个可能存在的误会：在半导体能带结构中，导带顶(Conduction band minimum, E_c)通常在高能量处，而价带顶(Valence band maximum, E_v)则在较低的能量处。能量轴一般是往上为能量增加的方向，换句话说， E_c 在上， E_v 在下，中间是禁带。费米能级(E_F)的位置决定了载流子(电子和空穴)的分布特性。

****对于未掺杂(本征)半导体：****

费米能级 E_F 接近能隙中点(E_i)，即在 E_c 与 E_v 之间的中间位置。

****当对半导体进行 p 型掺杂(掺入受主杂质)：****

1. p 型掺杂会引入受主能级(接近价带顶 E_v 附近)，这些受主能级容易捕获电子，从而留下空穴(正电荷载流子)在价带中。结果就是半导体中空穴浓度大幅上升。

2. 空穴是价带中缺失的电子，因此费米能级会向价带方向移动，以反映价带上态被部分填满、空穴浓度增多的状况。

由于价带在能量坐标中比导带更低 ($E_v < E_c$)，费米能级向价带移动意味着费米能级在能量图中实际上是“下降”了(即更接近较低能量的 E_v ，而不是 E_c)。所以正确的说法应该是：****p 型掺杂会使费米能级向较低能量处移动，更接近价带顶 E_v ，而非导带****。这通常被描述为“费米能级下移”。从本征态 E_i 出发， E_F 会从中间位置向下靠

近 E_v 。

****为何有时会产生“费米能级提高”的误解？****

可能出现误解的原因有以下几种：

1. **参考能量基准不同：**

有时某些图表或文献可能以价带顶 E_v 作为参考零点。当你以 E_v 为参照点(将 E_v 设为 0 eV)，本征费米能级 E_i 可能是在某个负的相对值(因为 $E_c > E_v$ 而 E_i 在中间)，此时如果掺入 p 型杂质使 E_F 进一步“接近” E_v ，那么相对于 E_v 参考来说， E_F 数值从“更负”变得“接近零”或略高(若极重 p 型使 E_F 进入价带)，表面上看是“提高”了数值，但本质上 E_F 是从高能量的 E_c 方向往下移到接近 E_v 。

简而言之：****如果参考点是 E_v ，那么 E_F 接近 E_v 意味着数值上从负值提升到接近零或成为正数。这在数值上看似 E_F 提高，但在实际的绝对能量标尺中， E_F 是向更低的绝对能量移动。****

2. **情境描述不当或曲解：**

或者在阅读教材或论文时，对应图中某些文字描述有歧义，或者作者在比较两种不同的情况(如与 n 型情形比较，p 型的 E_F 自然比 n 型时要低)。有时候描述中提到“费米能级离导带更近或更远”需要谨慎理解上下文，如果材料或参照物不同而混淆，就可能产生误解。

****总结：****

- p 型掺杂会使费米能级更接近价带顶 E_v ，这在绝对能量坐标下是 E_F 降低(往低能量方向走)的过程。
- 如果你以价带顶 E_v 为参考零点，并将 E_c 设为更高的正能量，那么从负值的 E_F 移到接近 0 或正值在数值上好像是“上升”，但实际上是因为改变了参考点导致的错觉。
- 正确理解：p 型掺杂使半导体中的空穴增多， E_F 向 E_v 移动，即费米能级在能量上“下降”到更接近价带的水平。

当讨论 p-n 结能带图时，需要将整个半导体区域从一个较为宏观且统一的角度来考量。在你看到的 p-n 结能带示意图中，常常会注意到 p 区的能带(包括导带与价带)相对于 n 区而言“上下错位”或能量更高/更低。这种能带阶梯式变化是由于内建电场(内建电势)的存在所导致的。

****关键原因在于费米能级对齐及内建电势：****

当 p 型半导体和 n 型半导体紧密接触形成 p-n 结后，两边材料各自原来的费米能级(E_F)不一样。对于 n 型区， E_F 更接近导带；对于 p 型区， E_F 更接近价带。将这两种半导体接触后，为了达到热平衡(不考虑外加电压的情况下)，电子会从一边扩散到另一边，最终整个 p-n 结系统需要形成一个统一的费米能级(全结处于热平衡时费米能级是平的，即统一的)。在此过程中，将形成一个内建电场，使能带发生弯

曲和相对升降。

****为什么 p 区看上去能量更高？****

1. **能带弯曲与电势差：**

原先在单独存在时，p 区和 n 区有各自的费米能级和能带结构。当它们接触后，电子从 n 区(相对电子较多一侧)向 p 区扩散，同时由于空穴的扩散也会发生(相反方向)，最终导致在界面附近形成一个空间电荷区(耗尽区)，这个空间电荷区内存在电场。该电场的存在使得能带在空间上发生弯曲，为满足整个结的费米能级统一，p 区和 n 区的能带边缘(E_c 和 E_v)相对于彼此的“参照值”会错开。

2. **参考点与电势分布：**

在能带图中，通常是将电子能量增加方向向上画，因此电势的上升会使能带图在该区域“向上抬”。p 区侧由于掺杂类型和浓度不同以及形成内建电势，造成 p 区一侧的绝对电势(空间电势)相对于 n 区更高。能带边缘(E_c , E_v)在能量图中会对应这种电势变化而整体抬高或压低。

简而言之，当将两个区域连接起来统一费米能级，p-n 结会在平衡下形成一个内建势垒，这个势垒在图中体现为从 n 区朝 p 区方向的 E_c 与 E_v 整体上升(或说 p 区侧相对于 n 区侧的能量参考值更高)。这种升高是相对于 n 区来说的，因为要让费米能级线对齐，p 区必须在能带图上产生一定的势垒关系。

3. **统一费米能级的结果：**

在平衡状态下，费米能级贯穿整个 p-n 结应当是连续且统一的。为了达到这一点，如果以 n 区为参考，那么 p 区的能带需要上下移动(主要是通过空间电荷层内建电场所引起的“弯曲”)以使得 EF 对齐。这就导致在能量图中，你会看到 p 区看起来像是“被抬高”了，即 p 区的价带与导带在图中位置比 n 区的相应带稍微上移。在实际物理中，这代表 p 区具有较高的电势(对电子而言是一个势垒)，电子从 n 区流向 p 区会遇到潜在障碍，这有助于维持平衡状态下的载流子分布。

****总结：****

- p 型与 n 型单独存在时各有自己的费米能级位置。
- 接触形成 p-n 结后，为了达到热平衡，费米能级在整个结处统一。
- 内建电势的存在导致能带弯曲，使得在能带图上 p 区的能带相对于 n 区出现“位能升高”的情况。
- 这并不是说 p 区掺杂本身就提高了能带，而是由于结的整体平衡与内建电场的存在，使得从 n 区看向 p 区时，p 区的能带相对于 n 区“显得更高”。

Recombination & Depletion Capacitance

复合

τ_p : p 区中电子生存时间

τ_n : n 区中电子生存时间

diffusion length:

$$L_p = \sqrt{D_p \tau_p} \quad 26.$$

$$L_n = \sqrt{D_n \tau_n} \quad 27.$$

关于费米能级的问题

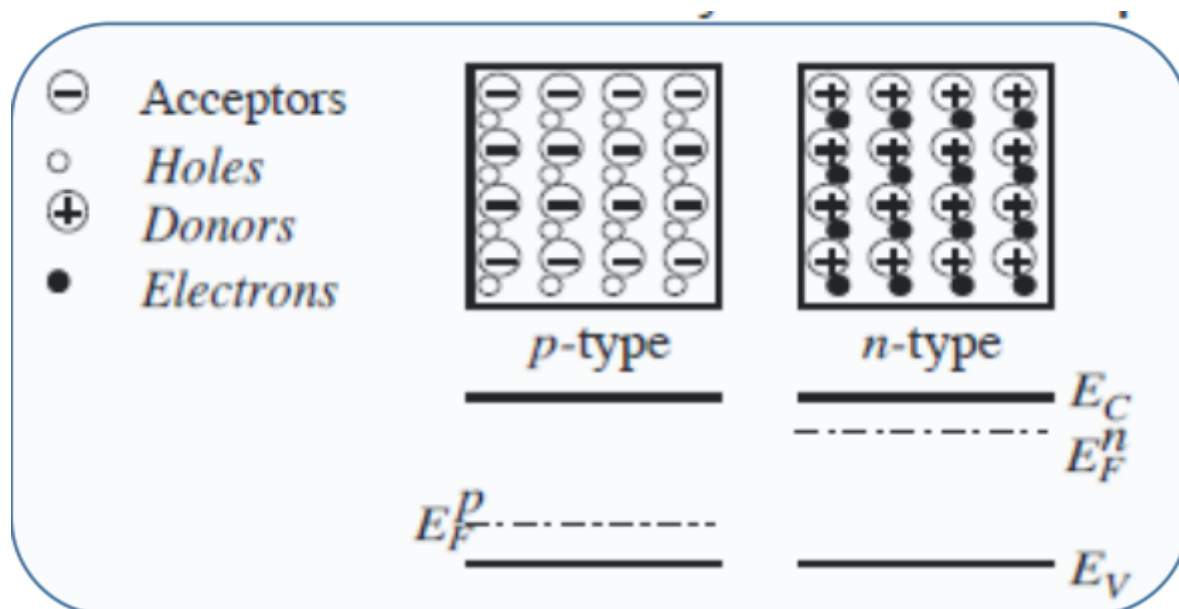


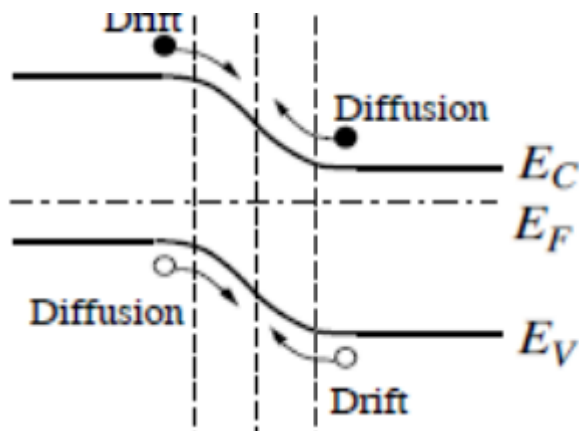
Figure 11: 分开的情况

当分开的时候, 确实 p 型的费米能级接近价带, n 型的费米能级解决导带.

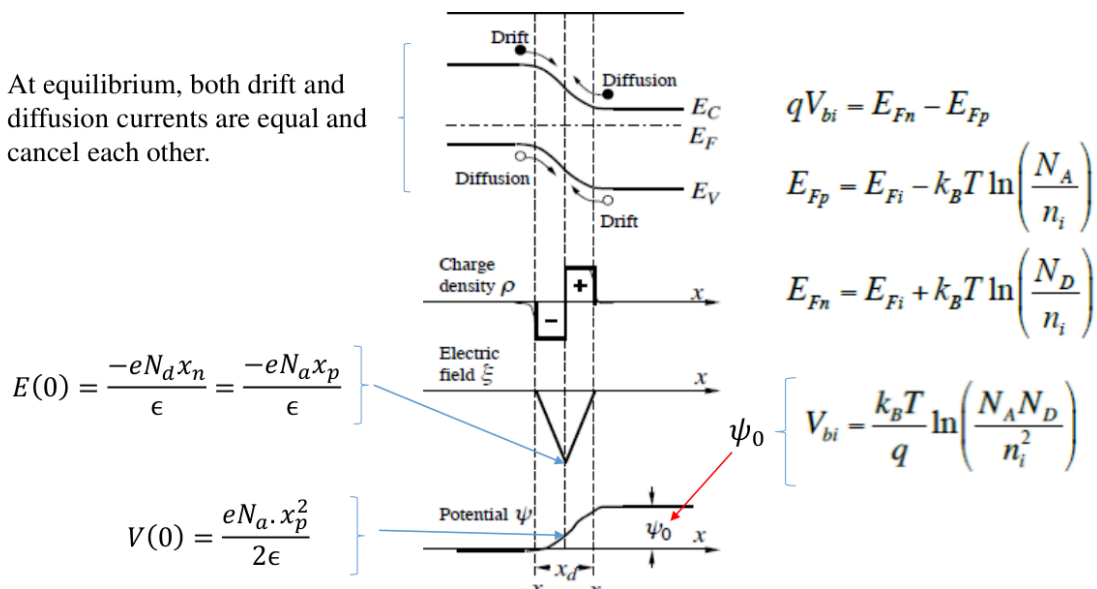
- 计算多子: $n \approx N_D = n_i \exp\left(\frac{E_F^n - E_i}{k_B T}\right)$, $p \approx N_A = n_i \exp\left(\frac{E_i - E_F^p}{k_B T}\right)$

$$\begin{aligned}
 E_F^n - E_F^p &= (E_F^n - E_i) + (E_i - E_F^p) \\
 &= k_B T \ln(N_D / n_i) + k_B T \ln(N_A / n_i) \\
 &= k_B T \ln\left(\frac{N_D N_A}{n_i^2}\right)
 \end{aligned}$$

当 p 型半导体和 n 型半导体连接在一起的时候, 发生了变化



At equilibrium, both drift and diffusion currents are equal and cancel each other.



contact potential 就是 V_{bi}

$$x_n = \left(\frac{2\varepsilon_0\varepsilon_S(\psi_0 + V)N_A}{qN_D(N_A + N_D)} \right)^{1/2}, \quad x_p = \left(\frac{2\varepsilon_0\varepsilon_S(\psi_0 + V)N_D}{qN_A(N_D + N_A)} \right)^{1/2}$$

$$w = x_p + x_n = \left(\frac{2\varepsilon_0\varepsilon_S(\psi_0 + V)(N_D + N_A)}{qN_A N_D} \right)^{1/2}$$

φ_0 就是 V_{bi}

- Depletion Capacitance(耗尽电容)

当 PN 结被加上一个小信号的时候会有一定的电容效应

The small signal capacitance of the space charge layer is given by $C \equiv \frac{dQ}{dV_R}$
where Q is the unipolar charge in the depletion region

$$Q = qAN_D x_n = qAN_A x_p = A \left(\frac{2q\varepsilon_0\varepsilon_S(\psi_0 + V_R)N_A N_D}{(N_A + N_D)} \right)^{1/2}$$

Therefore

$$C = A \left(\frac{q\varepsilon_0\varepsilon_S N_D N_A}{2(N_A + N_D)(\psi_0 + V_R)} \right)^{1/2}$$

单极性电荷 (unipolar charge): 表示耗尽区中某一侧的总电荷量

细节: 电流的 **I-V** 组成关系

- 平衡状态下 drift and diffusion 是相反而且相互抵消的
- 正向偏置下, 由于内建电场导致的 drift 电流减小了, 电子被注入 p 区, 空穴被注入 n 区

Carrier distribution in a forward biased **pn** junction. It is assumed that the **p region** is more heavily doped than the **n region**: $N_A \gg N_D$.

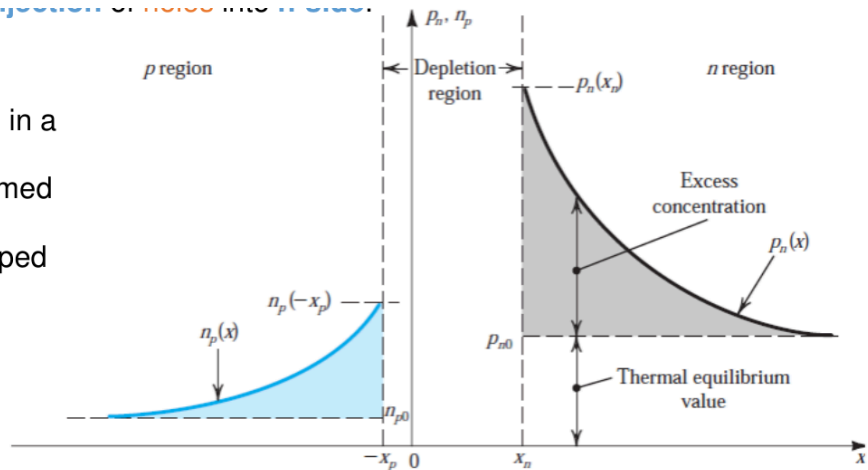
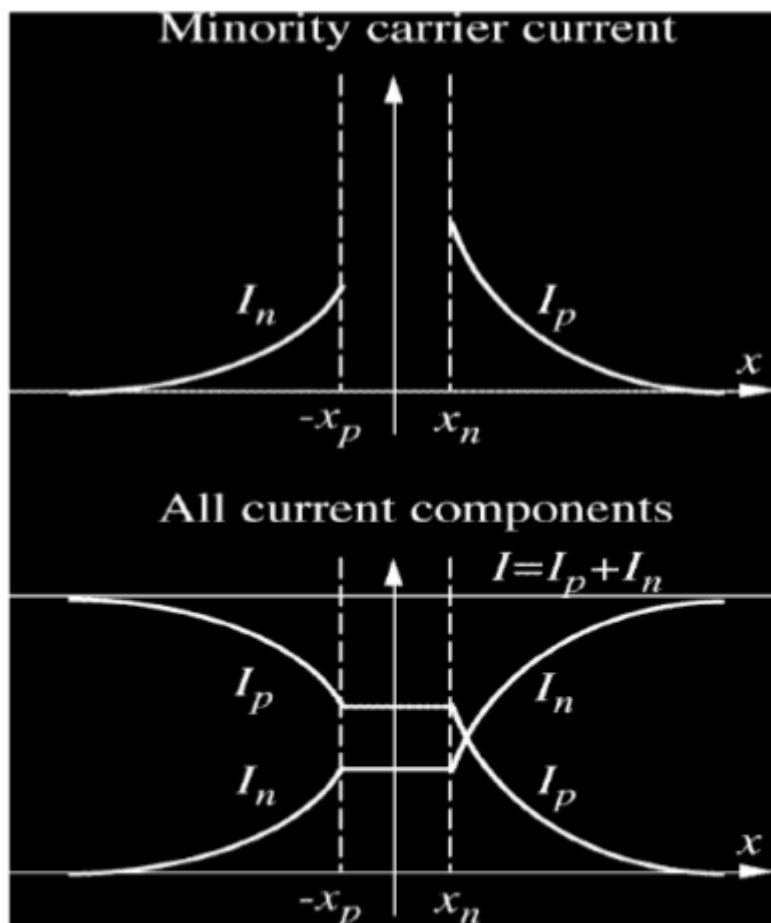


Figure 12: PN 结少数载流子的分布, 被注入的边缘浓度最高

- 总体上, 电流是一个不随区域变化的恒定值, 在从 P 到 N 的方向, 可以认为是少数载流子的 diffusion 和多子的 drift 组成电流, 距离 PN 结越远, 少数载流子的 diffusion 越少, 多数子的 drift 越多



二极管 I-V 公式的导出

Hole diffusion into n-region:

$$p_n(x) = p_{n0} + p_{n0}(e^{V/V_T} - 1)e^{-(x-x_n)/L_p}$$

$$J_p(x) = -qD_p \frac{dp_n(x)}{dx}$$

$$J_p(x) = q\left(\frac{D_p}{L_p}\right)p_{n0}(e^{V/V_T} - 1)e^{-(x-x_n)/L_p}$$

$$J_p(x_n) = q\left(\frac{D_p}{L_p}\right)p_{n0}(e^{V/V_T} - 1) \xrightarrow{\text{NB for electron diffusion into p-region}} J_n(-x_p) = q\left(\frac{D_n}{L_n}\right)n_{p0}(e^{V/V_T} - 1)$$

$$I = A(J_p + J_n)$$

$$I = Aq\left(\frac{D_p}{L_p}p_{n0} + \frac{D_n}{L_n}n_{p0}\right)(e^{V/V_T} - 1)$$

Substituting for $p_{n0} = n_i^2/N_D$ and for $n_{p0} = n_i^2/N_A$ gives

$$I = Aqn_i^2\left(\frac{D_p}{L_pN_D} + \frac{D_n}{L_nN_A}\right)(e^{V/V_T} - 1) \longrightarrow I_S = Aqn_i^2\left(\frac{D_p}{L_pN_D} + \frac{D_n}{L_nN_A}\right)$$

$$I = I_S(e^{V/V_T} - 1)$$

Famous **SHOCKLEY** equation

Where I_S is also known as the “Reverse Bias Saturation” Current

左上角的是空穴扩散到 n 区的公式, 同理推导出电子扩散到 p 区的公式’

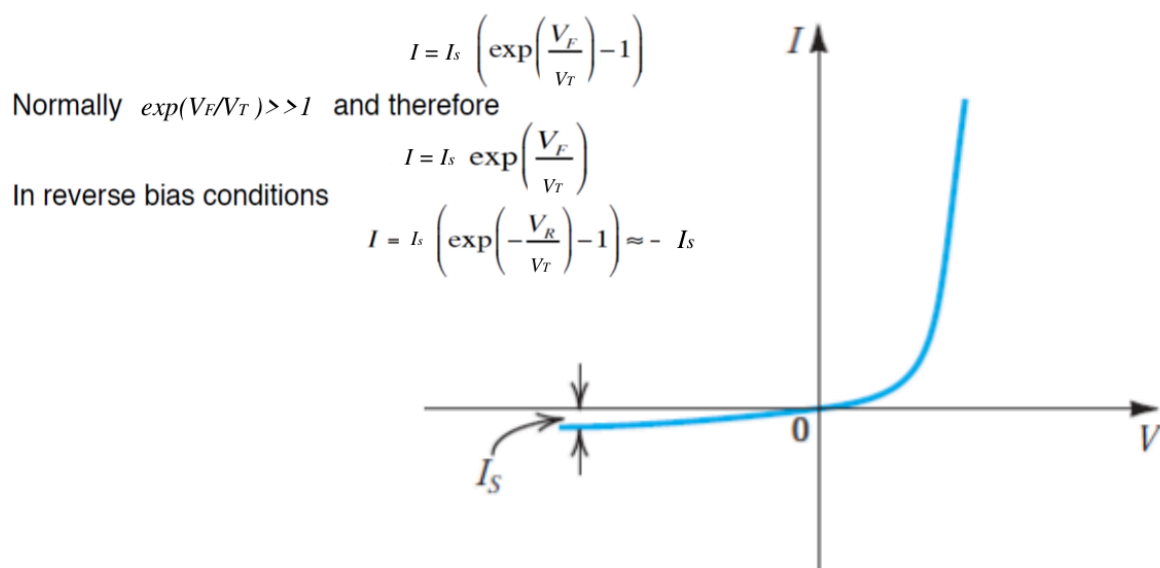


Figure 13: 一般使用的 I-V 公式