

# **Energibånd i faste stoffer med anvendelser**

## **Prosjektrapport i FY1013**

**Einar Stiansen  
Eivind Tore Hansen  
Morten Klungervik  
Anders Langseth**

## **Innholdsfortegnelse**

- s.3 Abstract
- s.4 Innledende ord
- s.4 Kvantefysisk beskrivelse av elektronet
- s.4 Hydrogenatomet
- s.4 Kvantisering
- s.5 Atomer med flere elektroner
- s.6 Eksklusjonsprinsippet
- s.6 Energibånd
- s.8 Kronig-Penney modellen
- s.10 Elektrisk ledningsevne
- s.12 Halvleder
- s.12 Doping; n-type og p-type
- s.14 P-n diode
- s.16 Innebygd potensial
- s.16 Energibånddiagram for en p-n overgang
- s.18 Bias over en p-n diode
- s.19 Matematisk beskrivelse av en ideell diode
- s.21 Grenseflaten som en kapasitans
- s.22 LED-diode
- s.23. Fotodetektorer og solceller
- s.24. Transistorer; bipolare
- s.26. Strøm i en transistor
- s.27. Energibånd i en transistor
- s.28. Operasjonsmodi
- s.29 Anvendelser av transistor (forsterker og bryter)
- s.31 Konklusjon
- s.32 Referanseliste

## Abstract

I klassisk fysikk blir elektroner beskrevet som punktladninger som går i bane rundt atomkjerner. I en kvantefysisk beskrivelse betraktes elektroner derimot som bølger beskrevet ved en bølgefunksjon  $\Psi$ . I kvantefysisk teori stilles det spesielle krav til denne bølgefunksjonen, noe som fører til kvantisering. For å kunne beskrive elektronet må ha kjennskap til elektronets fire såkalte kvantetall. I følge eksklusjonsprinsippet kan ikke to elektroner i samme system ha samme kvantetall, slik at tilstander utover grunntilstanden blir påtvunget.

Resultatet av de kvantefysiske prinsippene er at elektroner i et atom kun kan inneha visse energiverdier og man kan snakke om tillatte og forbudte energitilstander. Når atomer går sammen i et fast stoff vil man få splitting av tillatte energinivåer, og dermed energibånd. Mellom energibåndene vil man få båndgap der det ikke finnes tillatte energinivåer. Størrelsen på dette båndgapet og fordelingen av elektroner for et bestemt stoff bestemmer stoffets ledningsevne og dermed om det kan klassifiseres som leder, halvleder eller isolator.

Halvlederens ledningsevne kan forsterkes vesentlig ved å forurense halvlederen med andre typer atomer. Dette kalles doping, og man skiller mellom p-type og n-type dopede halvledere. Når man doper to ulike deler av samme halvlederstykket med ulik type doping får man det man kaller pn-overganger. Slike pn-overganger viser seg å ha interessante egenskaper som utnyttes i elektriske komponenter som dioder og transistorer.

Transistorer og dioder brukes på mange ulike områder og inngår i så å si alt av moderne elektriske artikler. Dioder brukes for eksempel som likerettere i blant annet batteriladere, og enkelte dioder har den egenskapen at de kan brukes som lyskilder. Transistorer har egenskaper som gjør at de kan brukes i forsterkere og i alle slags logiske kretser.

## Innledende ord

Omtrent halvparten av alle fysikere i verden er faststoff fysikere. Nyvinninger innen nettopp dette feltet har revolusjonert det teknologiske samfunnet, oppfinnelser som for eksempel transistoren er grunnmuren i dagens IT hverdag. Vi ser i dette prosjektet på hvordan vi får såkalte bånd av tillatte energier for et elektron i et fast stoff, og hvilke implikasjoner dette har.

## Kvantefysisk beskrivelse av elektronet

Båndteori for faste stoffer er et produkt av en kvantemekanisk beskrivelse av elektroner. Det faller derfor naturlig å introdusere grunnleggende kvantemekaniske prinsipper som er nødvendige for å forstå disse elektronenes natur.

I klassisk fysikk ser man på elektronet som en punktladning som følger bestemte baner rundt atomets kjerne. I kvantefysikken beskrives elektronene som bølger, gitt ved en

bølgefunksjon;  $\Psi(\mathbf{r}, t)$ . Der  $|\Psi^2|$  er sannsynlighetstettheten for å finne elektronet i en bestemt

posisjon ved tiden  $t$ . Som en konsekvens av denne materiebølgeteorien sier Heisenbergs usikkerhetsrelasjon at posisjon og impuls er omvendt proporsjonale størrelser, dermed lar det seg ikke gjøre å gi en skarp beskrivelse av et elektrons posisjon og impuls samtidig. For en

retning i rommet, si x-aksen har vi altså;  $\Delta x \Delta p_x \geq \frac{\hbar}{2}$ , dermed har man ingen mulighet til å

bestemme posisjonen til elektronet, langt mindre kan vi snakke om bestemte baner.

## Hydrogenatomet

Vi betrakter i første omgang et nøytralt, enkeltstående hydrogenatom. I dette tilfellet beveger elektronet seg i potensialfeltet (Coulombfeltet) fra kjernen. Verdien av dette er veldefinert, Schrödingers likning lar seg derfor løse analytisk for dette spesialtilfellet.

Den tidsuavhengige Schrödingers likning i tre dimensjoner;

$$-\frac{\hbar}{2m} \nabla^2 \Psi(\mathbf{r}) + U(\mathbf{r})\Psi(\mathbf{r}) = E\Psi(\mathbf{r}) \quad (1)$$

Løsningen av denne partielle differensiallikningen kan finnes ved å separere variable og er på formen  $\Psi(r, \mathbf{q}, \mathbf{f}) = R(r)\Theta(\mathbf{q})\Phi(\mathbf{f})$ , altså som en kombinasjon av tre likninger der hver er avhengig kun av en variabel i det sfæriske koordinatsystemet. Vi skal se at en kvantisering av  $\Psi$  fremkommer naturlig som en konsekvens visse krav man må stille løsningene.

## Kvantisering

Den semiklassiske Bohrteorien innførte kvantetallet  $n$  for å beskrive energinivåene i ett atom. Den bølgemekaniske beskrivelsen av elektronet fører med seg ytterligere kvantisering av beskrivelsen av elektronet.

- N-kvantetallet fra Bohr-teorien er å finne igjen i den moderne kvantefysikken, og gir tillatte energitilstander for elektronet. For eksempel vil energinivåene for hydrogen være gitt ved;

$$E = -\frac{m_r e^4}{2(4\pi\epsilon_0)^2 \hbar^2} \left(\frac{1}{n^2}\right), n=1,2,3\dots$$

Legg merke til at  $m_r$  ikke er elektronmassen, men den reduserte massen.

- I Bohrteori gir hver verdi for banedreieimpulsen en tilhørende verdi for energi. Men Bohr-teorien begrenser elektronets bevegelse til sirkelbevegelse, straks denne begrensningen er fjernet kan man vise at hver  $n$  kan ha flere verdier av banedreieimpuls. Disse verdiene er gitt ved;

$$L = \sqrt{l(l+1)}\hbar, \quad l = 0, 1, 2, \dots, (n-1)$$

Verdiene fremkommer ved å kreve at  $\Theta(\mathbf{q})$  tar endelige verdier ved  $\mathbf{q} = 0$  og  $\mathbf{q} = \mathbf{p}$  ( $\mathbf{y}$  må være normaliserbar).

Av historiske årsaker er det vanlig å tilegne en bokstav til de første verdiene av  $l$ :

**Tabell 1.** Tilordning av bokstav for banedreieimpuls

$l$	0	1	2	3	4
Bokstav	s	p	d	f	g

- I tillegg til at størrelsen på banedreieimpulsen er kvantisert har vi også romlig kvantisering det vil si at også retningen til  $\mathbf{L}$  kvantisert. Dette fremkommer på grunn av kravet om at  $\Phi(\mathbf{f}) = \Phi(\mathbf{f} + 2\mathbf{p})$ .  $\mathbf{L}$  kan dermed kun ta visse orienteringer i rommet. Hvis vi ser på de tillatte verdiene av  $\mathbf{L}$  langs z-aksen er de gitt ved;

$$L_z = m_l \hbar, \quad m_l = 0, \pm 1, \dots, \pm l$$

Vi ser nå at mens den semiklassiske Bohrteorien gav et kvantetall, gir den bølgemekaniske beskrivelsen av elektronet tre. Ved å kreve sfærisk harmoniske bølgefunksjoner tvinger ytterligere kvantisering seg fram.

- Gjennom eksperimenter er det vist at elektronet selv også har et "spinn". Dette spinn er konstant og er en egenskap ved elektronets natur. Retningen er derimot kvantisert i nøyaktig to retninger, "spinn-opp" og "spinn-ned". Likningen for størrelsen på spinn har samme form som for banedreieimpulsen;  $S = \sqrt{s(s+1)}\hbar$ , og siden banedreieimpulsen gav  $2l+1$  ulike tillatte komponenter og vi her krever nøyaktig to løsninger må vi ha  $s = \frac{1}{2}$ .

$$\text{Altså } S = \frac{\sqrt{3}}{2} \hbar$$

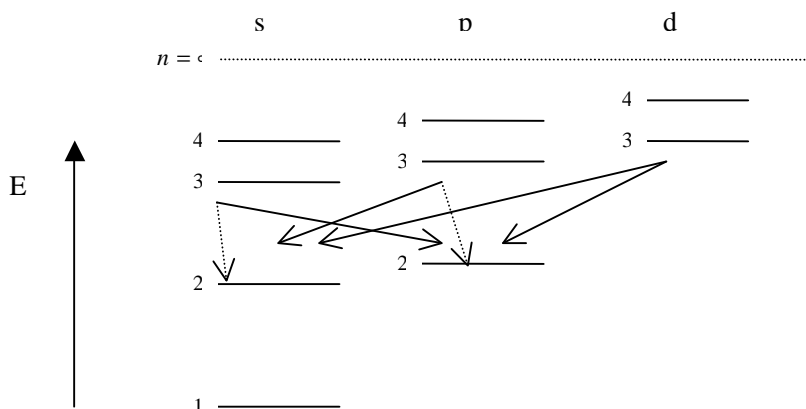
$$\text{Retningen er kvantisert i henhold til; } S_z = m_s \hbar, \quad m_s = \pm \frac{1}{2}.$$

Vi har altså hele fire kvantetall  $n, l, m_s, m_l$  som må oppgis for å beskrive et elektron i et hydrogenatom.

### Atomer med flere elektroner

Vi har så langt bare sett på beskrivelsen av elektronet i hydrogenatomet. I ett atom med flere enn ett atom beveger ikke lenger elektronet seg i et sentralfelt, men også i feltet induisert av de andre elektronene. En nøyaktig og analytisk løsning av Schrödingers ligning for slike tilfeller er umulig. Men, ved å gjøre den såkalte sentralfelttilnærmingen kan vi igjen basere oss på en "to-legeme-problemstilling". I denne tilnærmingen ser vi på hvordan elektronet beveger seg i potensialet fra kjernen og gjennomsnittsbidraget til potensialet fra alle andre elektroner. Like fullt vil totalenergien til elektronet i et mangeelektronatom være avhengig av

både  $n$  og  $l$ . Uavhengigheten av bandedreimpulsen var en konsekvens av Coulombkraftfeltet og forsvinner når vi introduserer flere elektroner. På denne måten for man en finndeling av tillatte energinivåer selv for en gitt verdi av hovedkvantetallet  $n$ . På figuren er det visst at et elektron som deeksiterer fra  $n=3$  til  $n=2$  ikke bare har en mulig energidifferanse, men hele seks. Deeksitasjonene med  $\Delta l = 0$  er sagt å være ulovlige, selv om de opptrer en sjelden gang. Regelen er at spranget må tilfredsstillte  $\Delta l = \pm 1$ .



**Figur 1,**  
deeksitasjon i et mangeelektronato  
 $m$ . Større verdi for  
 $l$  gir en høyere  
totalenergi.  
"Forbudte"  
sprang med stippet  
pil

Elektronene fordeler seg altså i hovednivåer bestemt av  $n$ , og undernivåer av disse bestemt av  $l$ . Men naturen tilstreber gjerne tilstander som minimaliserer energien, hvordan kan det da ha seg at ikke alle elektronene legger seg i grunntilstanden? Svaret ble formulert av Wolfgang Pauli, og har senere blitt kalt eksklusjonsprinsippet.

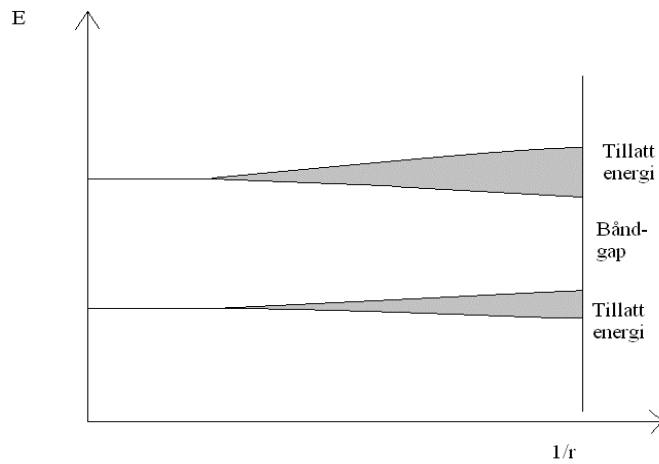
### Eksklusjonsprinsippet

Grunnen til at elektroner blir tvunget opp i høyere energinivåer enn grunntilstanden er at innenfor et system kan to elektroner ikke ha samme sett kvantetall. Slik blir nå antallet elektroner på et bestemt energinivå begrenset. For en gitt verdi av  $n$  og  $l$  vil det som vi har sett være tilgjengelig  $2l+1$  tilstander som følge av variasjon av  $m_l$ . I hver av disse tilstandene kan det være to elektroner som følge av to mulige retninger av elektronets egenspinn. Totalt har vi dermed  $N=2(2l+1)$  tilgjengelige tilstander på hvert gitte  $nl$ -orbital ( $l \leq n-1$ ).

### Energibånd

Vi ser på atomer i en krystallstruktur med et stort antall atomer. Hvis vi nå antar at vi kunne "strekke ut" gitteravstanden, slik at atomene ikke påvirker hverandre på noen slags måte, da ville energinivåene i atomene tilsvare de diskrete nivåene beskrevet for enkeltatomet. Men ettersom denne gitteravstanden minskes vil en vekselvirkning mellom atomene føre til en splitting av energinivåene. Vi snakker om et stort antall atomer ( $\square 10^{22}$ ) og hvert av disse vil gi opphav til et nivå innenfor det vi nå ser framstår som et

bånd av tillatte energitilstander. Bredden på båndet  $\Delta E$  er bestemt av antallet atomer som er nær nok til å sterkt vekselvirke. Dette er et lite antall sett i forhold til antallet i systemet som helhet, dermed vil ikke båndbredden øke med antall atomer, men snarere vil alle de tillatte nivåene legge seg innenfor intervallet  $\Delta E$ . Nivåene vil ligge så tett at vi essensielt kan se på det som en kontinuerlig fordeling av tilstander.



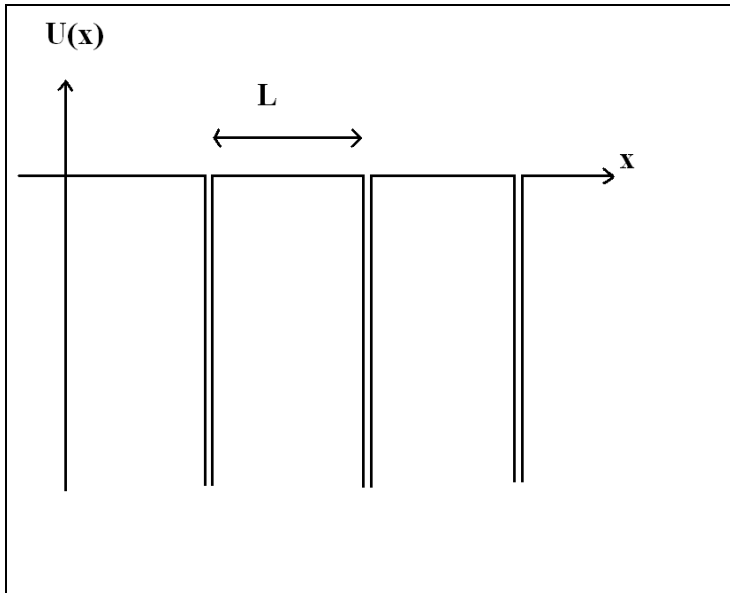
**Figur 2.** Ved å la gitteravstanden bli mindre og mindre, vil de diskrete energinivåene breseg ut i bånd

Merk området mellom de tillatte energinivåene, som vi skal se er bredden på dette båndgapet essensielt for å forstå et materiales elektriske egenskaper. Som vi har sett blir gjerne elektronet beskrevet ved en bølgefunksjon i kvantefysikken. Vi skal se at dette gir oss en mer matematisk beskrivelse av hvorfor energinivåsplittingen oppstår.

## Kronig-Penny modellen

(-drøftingen baserer seg på referanse 5 og 6.)

Modellen tar utgangspunkt i at elektroner beveger seg i et krystallgitter og opplever derfor et periodisk potensial. For enkelhets skyld betrakter vi ett 1-dimensjonalt tilfelle.



**Figur 3.** Periodisk potensial for en 1-dimensjonal gitterstruktur med  $L$  som atomær avstand.

For å lage oss et analytisk, relativt enkelt løsbart problem lar vi potensialbrønnene i ved de positive kjernene være beskrevet ved en deltafunksjon:

$$U(x) = -U_0 L \sum_{N=-\infty}^{\infty} \delta(x - NL)$$

Vi ser altså på en krystall av uendelig utstrekning for å se bort fra effekter fra randa til krystallen.

Det at potensialet og elektronets potensielle energi er periodisk betyr at en translasjon langs  $x$ -aksen i en avstand  $L$  gir oss samme situasjon vi startet med, den kvadrerte bølgefunksjon som gir oss sannsynlighetsdistribusjonen må derfor ha samme periodisitet som gitteret selv;

$$|\Psi(x + L)|^2 = |\Psi|^2$$

Noe som innebærer at bølgefunksjonen, i henhold til Blochs teorem, bare kan variere med en reell konstant fasefaktor  $K$ ;

$$\Psi(x + L) = e^{iKL} \Psi(x)$$

Den generelle løsningen av Schrödingers likning (1) i områdene med konstant null potensial

er:  $\Psi = A \sin kx + B \cos kx$  med  $k = \frac{\sqrt{2mE}}{\hbar}$ . Ved å kreve kontinuitet i grenseområdene,  $x=nL$ ,

og at diskontinuiteten for den deriverte i samme området er proporsjonal med styrken på

deltafunksjonen ( $\Delta(\frac{\partial \Psi}{\partial x}) = -\frac{2mL}{\hbar^2} U_0 \Psi(nL)$ ) kan man med *mye* algebra eliminere vekk

konstantene  $A$  og  $B$ . Konklusjonen for Kronig-Penny modellen er at den gir legger begrensninger på  $k$  for at vi skal få analytisk løsning. Løsning eksisterer hvis  $k$  oppfyller følgende relasjon:

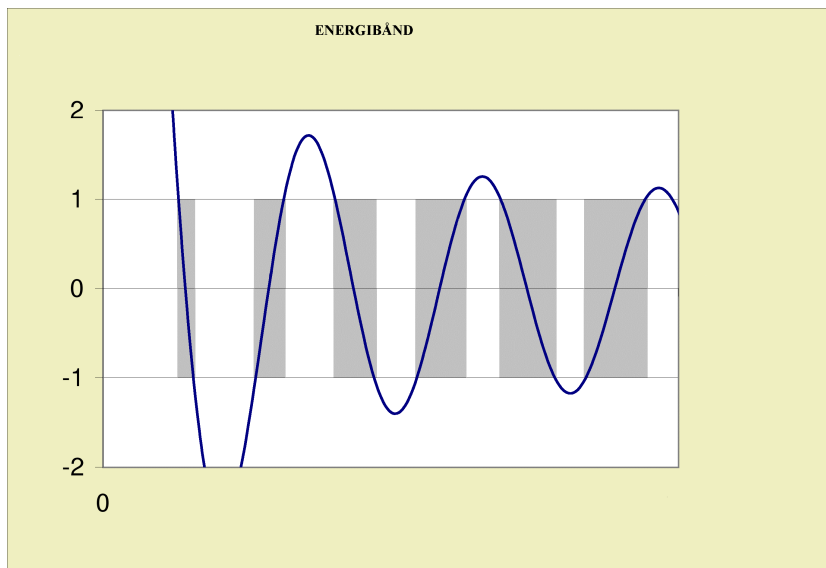


$$\cos KL = \cos kL - \frac{mLU_0}{\hbar^2 k} \sin kL \quad (2)$$

Vi ser nå at siden  $\cos KL$  bare tar verdier fra -1 til 1 vil det være områder hvor likningen ikke er oppfylt. Mens det er områder, bånd, av  $k$ -verdier der det er tillatte løsninger og dermed tillatte energier.

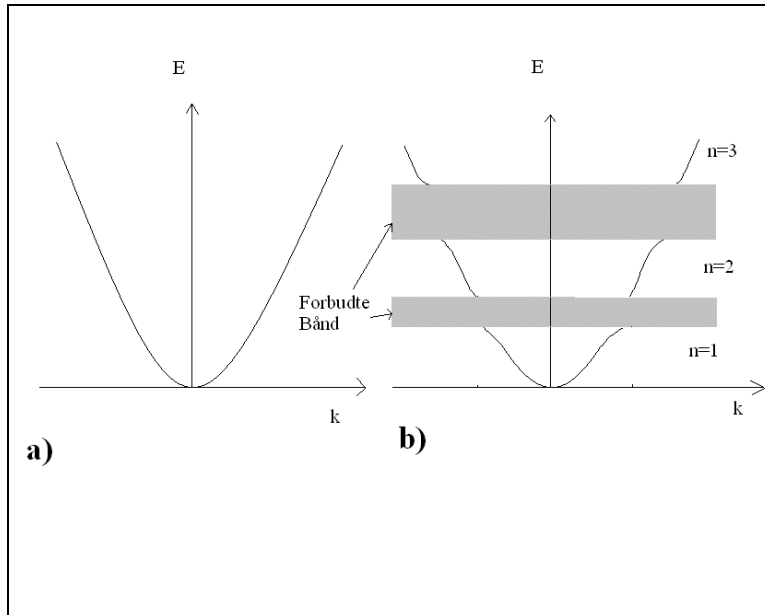
Vi tenker oss at  $x$ -aksen er "bøyd rundt i en ring" slik at vi kan sette randkravene,  $\Psi(x + NL) = \Psi(x)$  vi får da videre fra Blochs teorem at  $\Psi(x + NL) = \Psi(x) = e^{iKLN} \Psi(x)$ . Fordi  $\Psi$  må ta entydige verdier må  $e^{iKLN}$  være røtter av 1.

Dermed;  $e^{iKLN} = 1 \Rightarrow KL = \frac{2\pi n}{N}$  der  $n$  er heltall ( $n = 0, \pm 1, \dots, (N-1)$ ), og  $N$  er antallet potensialbrønner. Da sistnevnte antallet typisk er meget stort, vil en heltallig variasjon i  $n$  kun gi en neglisjerbar differanse. Dermed har vi innenfor de tillatte båndene kontinuerlig tillatt energi i grensen  $N \rightarrow \infty$ .



**Figur 4.**  $\cos kL + \frac{mLV_0}{\hbar^2 k} \sin kL$  plottet i mørkt, tillatte  $k$ -verdier befinner seg i intervallene farget i grått der  $\cos KL$  ligger mellom -1 og 1.

Ved å plote det tilhørende  $E$ - $k$  kurven med diskontinuiteten i henhold til de tillatte verdiene i figuren over kan dette se noe ut som i figuren under:



**Figur 5.** Energi som funksjon av  $k$  i a) for et fritt elektron og i b) for elektron som opplever et periodisk potensial

Som nevnt vil de kvantetilstandene som har lavest assosiert energi fylles opp først, deretter tvinges elektronene opp i høyere energinivå i samsvar med Paulis eksklusjonsprinsipp. Det høyeste energibåndet som ved det absolutte nullpunkt ( $T=0$  K) er fullstendig fylt kalles valensbåndet. Et elektron som deltar i ledningsevnen til et fast stoff befinner seg i ledingsbåndet som er det ”neste” båndet over valensbåndet. Som vi har sett er de tillatte energiverdiene til et elektron under påvirkning av krefter begrenset til bånd. Men det holder ikke å vite at disse båndene eksisterer, vi må også vite om tilstandene er besatt av elektroner. Sannsynligheten for at en tilstand er besatt av et elektron er beskrevet av Fermi-Dirac distribusjonen;

$$f(E) = \frac{1}{e^{(E-E_F)/k_B T} + 1} \quad (3)$$

som i korte trekk sier at sannsynligheten for at et elektron har nok energi til å ”hoppe” fra valensbåndet til ledningsbåndet (eg. sannsynligheten for at et elektron har en gitt verdi relativt til en referanseenergi; Fermienergien,  $E_F$ ) avhenger av temperaturen. Vi ser at

$$f(E_F) = \frac{1}{2},$$

noe som betyr at sannsynligheten for at tilstand med fermienergien er besatt

med et elektron er  $1/2$ . Størrelsen på dette spranget bestemmer også i hvor stor grad elektroner vil befinne seg i valens- eller ledningsbåndet ved aktuelle temperaturer. Vi ser at sannsynligheten for å finne elektroner i energinivåer over fermienergien faller eksponentielt. Derfor er sannsynligheten for å finne et elektron i ledningsbåndet forsvinnende liten hvis båndgapet er stort.

### Elektrisk ledningsevne

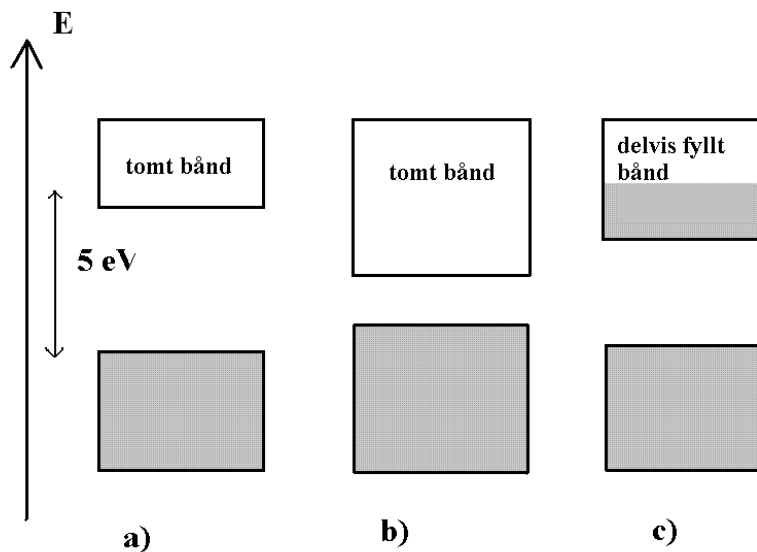
Det er vanlig å kategorisere faste stoffer etter deres evne til å lede strøm, det være seg elektronenes mulighet til å vandre i stoffet under påvirkning av et (svakt) elektrisk felt. Hvilke forhold må så være oppfylt for at et elektron skal delta i den elektriske ledningsevnen? Et energibånd som er tomt for elektroner kan selvsagt ikke bidra til ledningsevnen da det nettopp ikke har noen ladningsbærere. At et fullt bånd ikke heller bidrar til elektrisk ledning er derimot ikke like intuitivt klart. Hvis et elektron skulle bidra til ledning ville det drive mot det påtrykte feltet og således ha mottatt energi fra

feltet. Men dette kan ikke skje fordi alle tilstander i det tillatte båndet er besatt. Skal elektronet eksiteres til et energinivå der det kan delta i en elektrisk ledning må den tilførte energien tilsvare minst energigapet,  $E_g$ , det forbudte båndet representerer. Ved en gitt temperatur er det, som vi så ovenfor, en sannsynlighet for at et elektron får tilført tilstrekkelig energi fra gittervibrasjoner til å eksitere til ledningsbåndet.

Det som karakteriserer en *isolator* er nettopp at valensbåndet er fullt og båndgapet er stort, for stort til at den aktuelle termiske energien vanligvis er tilgjengelig. Et eksempel på isolator er diamant med et båndgap på 6eV. Et elektron kan derimot "presses" over energibarrieren ved å bli eksitert av et foton, høy temperatur eller å påtrykke et sterkt elektrisk felt. Sistnevnte er det vi kjenner som et dielektrisk sammenbrudd. Dette gjør seg derimot bare gjeldene kun for meget kraftige elektriske felt.

Akkurat som for en isolator har også *halvlederne* et fullt okkupert valensbånd ved  $T=0$  K, og et tomt ledningsbånd. Forskjellen er her at båndgapet er av en mye mindre størrelsesorden. Slik at den termiske eksiteringen av elektroner bidrar i større grad til å skape mobile ladningsbærere. Vi skjønner at den elektriske ledningsevnen til halvledere er sterkt avhengig av temperatur. Hvis temperaturen blir hevet i tilstrekkelig grad vil vanligvis isolatorer oppnå samme egenskaper som en halvleder, og motsatt; ved lave temperaturer vil halvledere oppføre seg som isolatorer.

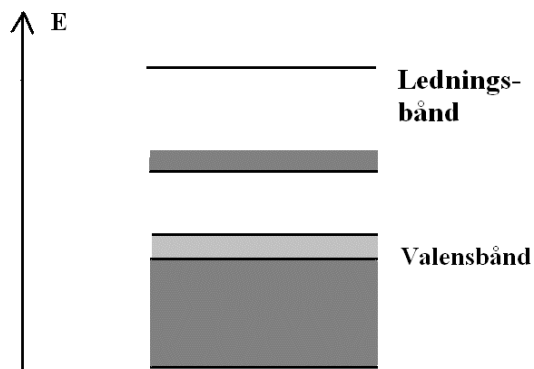
For metaller er bildet et annet, her er valensbåndet fullt, men også ledningsbåndet er delvis fylt ved  $T=0$  K. Et elektron kan dermed lett eksiteres til et litt høyere energinivå som respons på et påtrykt felt.



**Figur 6.** skjematisk fremstilling av energibånd for a) isolator, b) halvleder og c) metall (leder). Alle ved  $T=0$  K. Legg merke til det lille båndgapet for halvledere.

## Halvleder

I en halvleder er atomene bundet sammen i middels sterke bindinger. At bindingene ikke er sterkere enn de er innebærer at det selv ved romtemperatur er nok termisk energi tilstede til at et betydelig antall bindinger brytes og elektroner blir frigjort til ledningsbåndet. Et eksempel på halvleder er silisium med et båndgap på 1,12eV. Når et elektron eksiteres til ledningsbåndet og forlater bindingen vil den etterlate en ledig posisjon. Et valenselektron fra et naboatom kan ta denne plassen og dermed flytte den ledige bindingen. På denne måten vil et slikt "hull" vandre gjennom en halvleder under påvirkning av et felt. Elektronene vandrer mot feltet slik at hullet vandrer med feltet og oppfører seg som en positivt ladet partikkel. I en ren (intrinsikk) halvleder vil hver eksitasjon medføre et hull i valensbåndet, vi har dermed to typer ladningsbærere og bidragsytere til elektrisk ledningsevne.



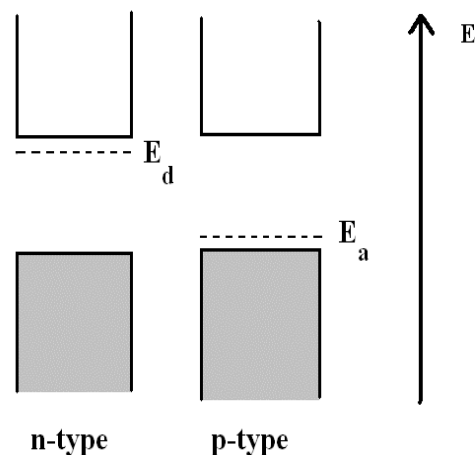
**Figur 7.** halvleder ved  $T > 0$  K. Legg merke til øverste lag i valensbåndet i lyst er tomt for elektroner. De er eksitert til ledningsbåndet og har etterlatt seg hull.

## Doping

En intrinsikk halvleder opptrer sjelden, selv små urenheter i krystallen i form av gitterfeil eller forurensningsatomer vil påvirke ledningsevnen i stor grad. Det å tilføre en halvleder en mengde forurensningsatomer kalles doping. For gruppe IV halvledere vil de viktigste dopingsatomene være type III og type V atomer. Vi ser først på type V doping som gir opphav til det vi kaller n-type doping

### N-type

Ved å tilføre en type V atom til en type IV halvleder vil fire av de fem valenselektronene til type V atomet delta i kovalente bindinger med omkringliggende type IV atomer. Det resterende elektronet vil således være løst bundet til positive kjernen. La oss si vi tilførte fosfor til en silisiumkrystall, da kan det vises eksperimentelt at ioniseringsenergien til fosforatomet er 0,05eV. Det vil si at det kun kreves 0,05eV for å eksitere dette løst bundet elektronet til ledningsbåndet. Halvlederen har nå et overskudd av elektroner, og sies nå å være n-type (negativ-type) halvleder. Fosforatomet gav fra seg et elektron og kalles et donatoratom. På grunn av den lille energidifferansen nødvendig for å eksitere elektronet er det vanlig å si at alle donatoratomene er ionisert når man skal beregne ledningsevne. Ved en slik doping skal det kun meget små konsentrasjoner med forurensningsatomer til for å øke ledningsevnen betraktelig. Elektroner blir den desidert største bidragsyteren til ledning, faktisk så dominerende at vi kan neglisjere ledning pga hull.



**Figur 8.** Introduserte tillatte energinivåer som følge av doping.

## **P-type**

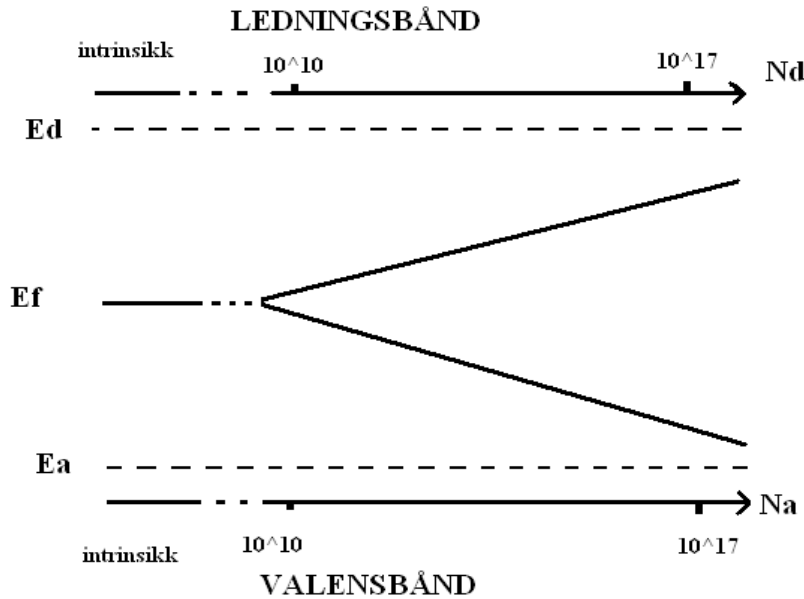
Man kan også tilføre et atom som har færre valenselektroner enn det som er nødvendig for å danne en fullstendig kovalent krystallstruktur. Ved å tilføre en type III atom f. eks. bor til silisium vil det kun dannes tre kovalente bindinger i stedet for fire. Denne ”mangelen” på et elektron som deltar i bindinger framstår som en ledig tilstand rett over toppen på valensbåndet. Det kreves således lite energi for at et valenselektron i et annet atom eksiteres til denne tilstanden. Et elektron som gjør dette vil etterlate seg et hull (ledig tilstand) i valensbåndet som kan delta i ledningsevnen. Vi ser at boratomet har mottatt et elektron og er således en akseptor. Silisiumet har en overvekt hull som bidrar til ledningsevnen og kalles en p-type (positiv-type) halvleder. For at det sistnevnte skal gjelde må antall akseptoratomer være større enn antall elektroner i ledningsbåndet. Det er viktig å legge merke til at mens en eksitasjon i en intrinsikk halvleder gav opphav til to ladningsbærere, ett elektron i ledningsbåndet og ett hull i valensbåndet, gir doping kun en ladningsbærer for hvert forurensningsatom.

Et annet viktig aspekt ved doping av halvledere er det såkalte np-produktet. Ved en gitt temperatur er produktet av konsentrasjonen av elektroner i ledningsbåndet og konsentrasjonen av hull en konstant.

$$n \cdot p = K(T)$$

Dvs. at ved å dope en halvleder til å f. eks. øke konsentrasjonen av elektroner med en faktor 10, vil konsentrasjonen av hull bli redusert med samme faktor. På denne måten vil ledningsegenskapene til en n-type halvleder være hovedsakelig av elektroner og kun neglisjerbart av hull. Det er derfor mulig å styre ledningsevnen til halvlederen i meget stor grad ved å tilføre forurensningsatomer i varierende mengde.

Doping medfører også at ferminivået forandrer posisjon. For en intrinsikk halvleder befinner dette nivået seg midt i båndgapet, men den skyves opp mot ledningsbåndet eller ned mot valensbåndet avhengig om doping er henholdsvis n-type eller p-type. Som vi har sett beskriver dette nivået den energien som har 0,5 sannsynlighet for å være besatt. Dersom dette nivået ligger nærme ledningsbåndet vil sannsynligheten for å finne elektroner i ledningsbåndet være stor. Vi vil da ha mange elektroner i ledningsbåndet, og som vi har sett desto færre hull, slik at vi den dominerende ladningsbæreren er elektroner. I motsatt fall ligger ferminivået nærmere valensbåndet slik at få elektroner eksiteres opp til ledningsbånd. Vi har i så fall ledning i dominerende grad av hull og altså en p-type halvleder.



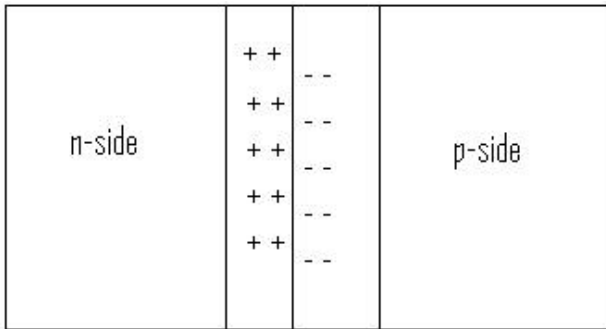
**Figur 9.** Skjematisk fremstilt posisjon av ferminivået for en intrinsikk halvleder (ladningskonsentrasjon  $\square 10^{10}$ ), og som funksjon av dopingskonsentrasjon. Konsentrasjoner oppgitt i antall pr.  $cm^3$

## P-n diode

(Drøftingen baserer seg på referanse 9)

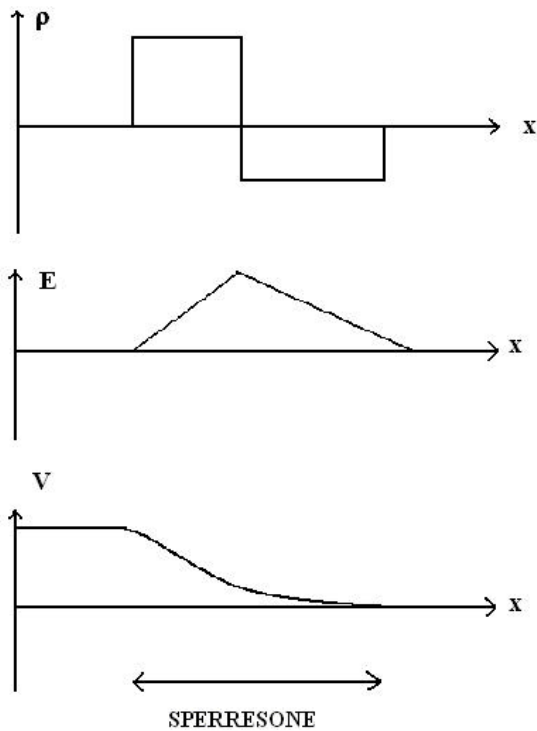
En p-n overgang består av en halvleder som er dopet med akseptorer på den ene siden og donorer på den andre. Dens hovedoppgave er å *likerette* strømmen i kretsen slik at den kun går en vei. Den har også andre anvendelser som bla fotodiode (lysfølsom og brukes i solceller), LED (diode som sender ut lys) og varaktor diode, (brukes som en variabel kapasitans i en krets)

Dersom vi tar utgangspunkt i en halvleder av silisium får vi en p-n diode ved å dope de to ulike sidene med f.eks. fosfor og bor. På hver av de to sidene har vi da ulike konsentrasjoner elektroner og hull som kan bevege seg fritt, og naturlig nok så vil det oppstå en diffusjonsprosess der elektronene og hullene ønsker å jevne ut konsentrasjonen slik at konsentrasjonen av elektroner og hull vil være lik overalt. Men denne prosessen kan ikke vare evig siden hullene vil etterlate seg negativt ladete akseptorkjerner og elektronene vil etterlate seg positivt ladete donorkjerner fordi kjernene ikke kan bevege seg. Det vil nå settes opp et elektrisk felt som hindrer den videre diffusjonsprosessen, og det vil innrette seg en likevekt. Denne likevekten baserer seg på at ladningene som diffunderes vil rekombinere slik at vi fortsatt får at  $n \cdot p = n_i^2$  på begge sider av dioden. Litt enkelt kan vi si at akseptorene ønsker å beholde hullene, mens donorene vil beholde elektronene.



**Figur 10.** P-n kontakt i likevekt. Området med ladningene kalles for sperresone eller depleksjonsområde.

Hvis vi ser på figuren over kan vi begynne å gjette oss til hvordan det elektriske feltet og det innebygde potensialet ser ut. Det elektriske feltet vil være øke der det er positiv ladning og minke der det er negativ ladning, det er også klart at feltet må være null i de nøytrale områdene langt unna grenseflaten. For potensialet har vi at det vil ha motsatt fortegn av det elektriske feltet slik at det vil være konstant i de nøytrale områdene, mens det vil minke i grenseflaten.



**Figur 11.** Egenskaper ved en idealisert pn-kontakt. Øverst er ladningsfordelingen, i midten det elektriske feltet og nederst potensialet.

Om vi snur figuren og tenker på det motsatte tilfellet ser vi lett at vi får et negativt elektrisk felt og dermed en økning i potensialet.

## Innebygd potensial i en p-n diode

For å finne det innebygde potensialet kan vi benytte to relativt enkle metoder. I den ene metoden må vi bruke Poissons ligning og integrere over ladningene to ganger. Dette forutsetter at vi må se på ladningene i sperresonen til å være konstante på begge sider av grenseflaten slik som figur 8 viser. Den andre metoden inneholder ingen integrering, men vi må til ha kjennskap til energibåndteori for en p-n overgang og konsentrasjon av ladningsbærere i ledningsbåndet og valensbåndet.

Ved å bruke Poissons ligning;

$$\frac{d^2V}{dx^2} = -\frac{\mathbf{r}(x)}{\mathbf{e}},$$

der  $\mathbf{r}(x)$  er ladningsfordelingen får vi:

$$E(x) = -\frac{dV}{dx} = -\frac{d}{dx} \int dV = \frac{1}{\mathbf{e}} \int p(x) dx$$

$$E_n(x) = \frac{eN_D x}{\mathbf{e}} + K_2 = \frac{eN_D x}{\mathbf{e}} - \frac{eN_D x_n}{\mathbf{e}} = \frac{eN_D}{\mathbf{e}} (x - x_n)$$

$$E_p(x) = -\frac{eN_A x}{\mathbf{e}} + K_1 = -\frac{eN_A x}{\mathbf{e}} - \frac{eN_A x_p}{\mathbf{e}} = -\frac{eN_A}{\mathbf{e}} (x + x_p)$$

Der  $E_n$  og  $E_p$  er det elektriske feltet på henholdsvis n-siden og p-siden. Dette gir for det innebygde potensialet:

$$V_{ib} = -\int_{-x_p}^{x_n} E(x) dx = \frac{eN_D}{\mathbf{e}} \left( \frac{x_n^2}{2} \right) + \frac{eN_A}{\mathbf{e}} \left( \frac{-x_p^2}{2} \right)$$

Når vi vet at ladningene på hver side av grenseflaten er like store, dvs.:  $eN_A x_p = eN_D x_n$  får vi:

$$V_{ib} = \frac{eN_D^2 x_n^2}{2\mathbf{e}} \left( \frac{N_A + N_D}{N_A N_D} \right) = \frac{eN_A^2 x_p^2}{2\mathbf{e}} \left( \frac{N_A + N_D}{N_A N_D} \right)$$

Den andre metoden forutsetter at vi vet at konsentrasjonen av hull i valensbåndet er

$p = n_i \exp\left(\frac{E_i - E_F}{kT}\right)$  på p-siden og konsentrasjonen av elektroner i ledningsbåndet på n-siden

er  $n = n_i \exp\left(\frac{E_F - E_i}{kT}\right)$ . Fra energibåndteori om p-n overgangen skal vi snart vise at

$|V_{ib}| = |V_p| + |V_n|$  der  $qV_p = -(E_i - E_F)$  og  $qV_n = -(E_i - E_F)$ . Ved å substituere for  $E_i - E_F$  og

sette  $q=e$  får vi:  $|V_{ib}| = \frac{kT}{e} \ln \frac{N_A N_D}{n_i^2}$

## Energibånddiagram for en p-n overgang

Når vi skal tegne opp et energibånddiagram for en p-n overgang er det viktig å huske på at det kjemiske potensialet (fermienergien) er konstant i hele dioden. Som vi har sett i figur 9. avhenger fermienergien av dopingskonsentrasjonen. Vi vil da få forskjellige energinivåer for valensbåndet og ledningsbåndet på hver side av sperreområdet.

Hvis vi ser på hva som skjer matematisk er det viktig å huske på at det har innstilt seg en dynamisk likevekt, og at prosessene ikke har stoppet opp. I tillegg til diffusjonsstrømmen vil



vi ha en driftsstrøm. Denne driftsstrømmen oppstår fordi det termisk genereres hull nær kanten av sperresonen på n-siden og elektroner på p-siden. På grunn av det elektriske feltet vil disse ladningene bli transportert til den andre siden, og vi får altså to nye strømmer i motsatt retning av de to diffusjonsstrømmene. Altså vil vi ha at summen av diffusjonsstrømmen og driftsstrømmen vil være null. Hvis vi setter opp de aktuelle ligningene for de positive ladningsbærerne får vi:

$$J_p = J_p(\text{Drift}) + J_p(\text{diffusjon}) = 0$$

Her er  $J_p(\text{diffusjon})$  knyttet til strømmen fra diffusjonsprosessen mellom hull og elektroner, mens  $J_p(\text{drift})$  er knyttet til den termiske genereringen av elektroner og hull nær kantene av sperresonene. For å løse denne ligningen må vi bruke uttrykk for driftsstrømmen og diffusjonsstrømmen, og en del relasjoner mellom disse.

$$J_p = q\mathbf{m}_p p E_x - qD_p \frac{dp}{dx} = 0$$

$$J_p = \mathbf{m}_p \left( p \frac{dE_i}{dx} - kT \frac{dp}{dx} \right) = 0$$

Her er  $\mathbf{m}_p$  hullmobiliteten, og  $D_p$  diffusjonskoeffisienten.

Ved å substituere for  $dp/dx$  får vi når:

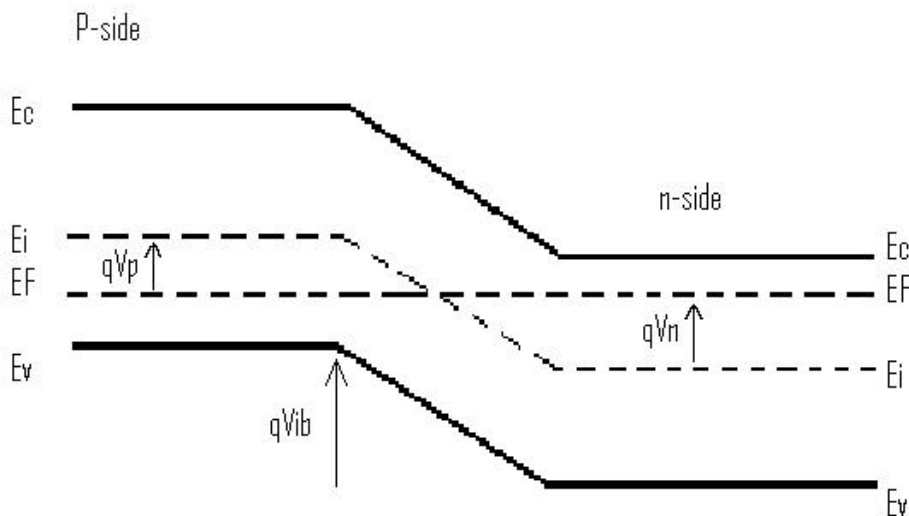
$$p = n_i \exp\left(\frac{E_i - E_F}{kT}\right)$$

$$\frac{dp}{dx} = \frac{p}{kT} \left( \frac{dE_i}{dx} - \frac{dE_F}{dx} \right)$$

Som da gir oss:

$$J_p = \mathbf{m}_p \left( p \frac{dE_i}{dx} - kT \cdot \frac{p \left( \frac{dE_i}{dx} - \frac{dE_F}{dx} \right)}{kT} \right) = \mathbf{m}_p p \frac{dE_F}{dx} = 0$$

Siden vi nødvendigvis vil få de samme ligningene for de negative ladningsbærerne, vil vi uavhengig av fortegn få at  $\frac{dE_F}{dx} = 0$ , nemlig at den kjemiske energien må være konstant gjennom hele dioden.

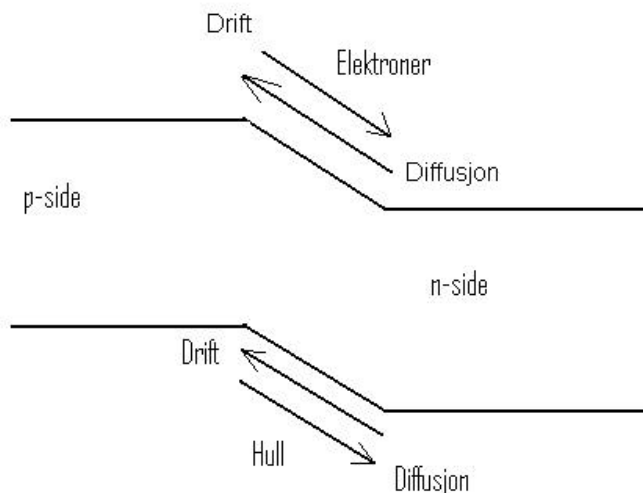


**Fig 12.** Energibåndfigur med potensialforskjeller. Vi ser  $E_F$  er konstant gjennom hele dioden, og at p-siden og n-siden ligger på forskjellig energetisk nivå

### Bias over en p-n diode

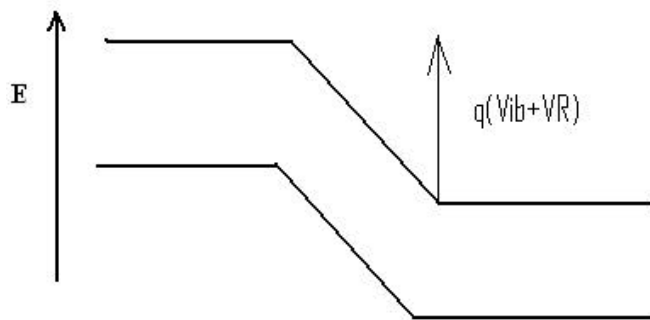
Når vi setter spenning over en diode kaller vi den for bias. Den viktigste egenskapen til en p-n diode er som sagt å være en likeretter i en krets. Dette ser vi ganske lett ved å studere ett *strøm-spenning* diagram. Der ser vi at strømmen går lett i den ene retningen, og at det knapt går noen strøm når vi snur biasen.

For å forklare hva som skjer med strømmen når vi anvender en bias på dioden må vi se på hva som skjer med de to diffusjonsstrømmene og de to driftsstrømmene. Siden driftsstrømmene kommer fra ladningsbærerne som genereres termisk nær kanten av sperresonen vil det være naturlig å anta at denne prosessen vil være den samme som når vi ikke anvendte bias på dioden. Dette vil vi også anta ved å se på et energidiagram, fordi driftsstrømmene ikke har noen potensialbarriere å passere vil ikke en endring av denne barrieren ha noen å si for driftsstrømmene.



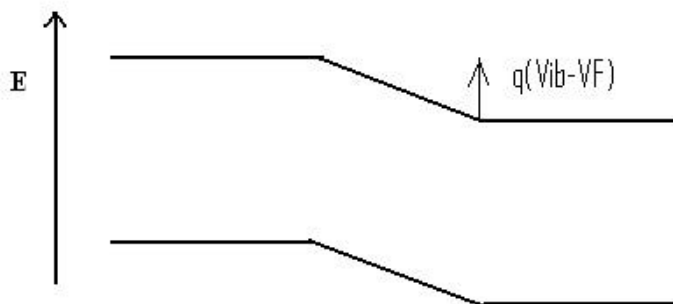
**Figur 13.** energibåndmodell med de fire strømmene indikert

En interessant egenskap oppstår når vi setter på en bakover bias. Den påtrykte spenningen vil da sette opp ett elektrisk felt som har samme retning som feltet inne i dioden. Siden det er dette feltet som hindrer diffusjonsstrømmen, vil vi intuitivt anta at dette vil redusere strømmen gjennom dioden når det vil innstille seg en ny likevekt. Dersom vi ser på hva som skjer med et energibånddiagram når setter på en bakover bias, ser vi at potensialforskjellen mellom ledningsbåndene øker. Dette gjør at det vil bli vanskeligere for ladningsbærerne å bevege seg gjennom grenseflaten, og det vil da gå mindre strøm gjennom dioden.



**Figur 14.** Pn-kontakt med påtrykt bakover-bias. VR representerer den påtrykte spenningen.

Når vi setter på en forover bias vil den påtrykte spenningen sette opp et elektrisk felt som er motsatt rettet av det som allerede eksisterer i dioden. Dette vil da bety at vi svekker det som hindrer diffusjonsstrømmen, og vi vil da anta at strømmen gjennom en diode vil øke med økende påtrykt spenning. Når vi ser på ett energibånddiagram som viser hva som skjer med en forover bias, vil vi få at potensialforskjellen ladningsbærerne må passere i grenseflaten vil bli mindre, og dermed mulighet for at flere ladningsbærere vil bevege seg gjennom dioden.



**Figur 15.** Pn-kontakt med påtrykt forover-bias

### Matematisk beskrivelse av en ideell diode

For å avgjøre formen på *strøm-spenning* diagrammet må vi vite litt mer om hvordan diffusjonsstrømmene oppfører seg på hver side av sperresonen. Diffusjonsprosessen er et forsøk på å få samme konsentrasjonen av henholdsvis hull og elektroner på hver side av grenseflaten. Om vi ser på likevektskonsentrasjonene av hull og elektroner på hver side av grenseflaten får vi:

$n_{n0}$  konsentrasjon av elektroner på n-siden

$n_{p0}$  konsentrasjon av elektroner på p-siden

$p_{n0}$  konsentrasjon av hull på n-siden

$p_{p0}$  konsentrasjon av hull på p-siden

Siden vi vet at  $n \cdot p = n_i^2$  på hver side av grenseflaten, får vi det innebygde potensialet:

$$V_{ib} = \frac{kT}{q} \ln \frac{n_{n0}}{n_{p0}} = \frac{kT}{q} \ln \frac{p_{n0}}{p_{p0}}$$

Som kan skrives om slik at vi får:

$$n_{n0} = n_{p0} \exp\left(\frac{qV_{ib}}{kT}\right) \text{ og } p_{p0} = p_{n0} \exp\left(\frac{qV_{ib}}{kT}\right)$$

Hvis vi ser på hva som skjer om vi anvender en bias  $V$  på dioden, og antar at forholdene mellom elektronkonsentrasjonene er de samme som under likevekt får vi:

$$n_n = n_p \exp\left(q \frac{V_{ib} - V}{kT}\right)$$

Hvis vi i tillegg kun ser på når en liten bias blir anvendt, har vi at:  $n_n \approx n_{n0}$  som gir:

$$n_n = n_p \exp\left(q \frac{V_{ib} - V}{kT}\right) = n_{n0} = n_{p0} \exp\left(\frac{qV_{ib}}{kT}\right)$$

$$n_p - n_{p0} = n_{p0} \left[ \exp\left(\frac{qV}{kT}\right) - 1 \right]$$

Nå må vi i tillegg huske på at ladningene som diffunderer gjennom grenseflaten vil rekombinere med de motsatte ladningene på den andre siden. Det viser seg at for elektroner har vi at konsentrasjonen et stykke  $l_p$  fra kanten på sperresonen inn i p-siden er:

$$dn(l_p) = dn(0) \exp\left(-\frac{l_p}{L_n}\right)$$

Der  $dn = n_p - n_{p0}$ ,  $dn(0) = n_{p0} \left[ \exp\left(\frac{qV}{kT}\right) - 1 \right]$  og  $L_n$  er diffusjonslengden til elektronet. Nå

bruker vi igjen at  $J_n = qD_n \frac{dn}{dx}$  for diffusjonsstrømmen. Ved å se på strømtettheten ved kanten på sperresonen får vi:

$$J_n(l_p = 0) = -\frac{qD_n n_{p0}}{L_n} \left[ \exp\left(\frac{qV}{kT}\right) - 1 \right]$$

En tilsvarende analyse for hullene gir:

$$J_p(l_n = 0) = -\frac{qD_p p_{n0}}{L_p} \left[ \exp\left(\frac{qV}{kT}\right) - 1 \right]$$

Dette gir en total strømtetthet:

$$J = J_0 \left[ \exp\left(\frac{qV}{kT}\right) - 1 \right]$$

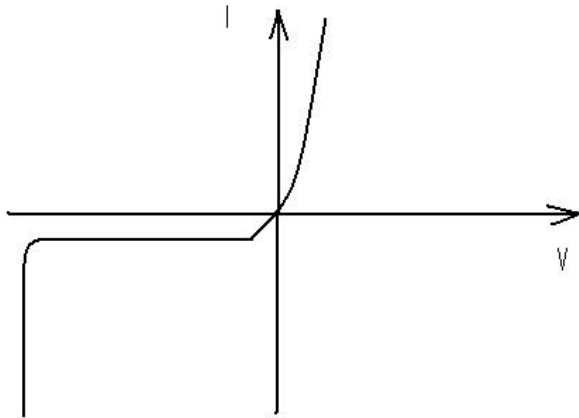
Og for strømmen får vi da den ideelle diode ligningen:

$$I = I_0 \left[ \exp\left(\frac{qV}{kT}\right) - 1 \right]$$

Vi ser at dersom vi anvender en så stor negativ bias at eksponentialleddet går mot null vil vi få en konstant strøm på  $-I_0$ . I følge det vi tidligere har skrevet må dette da representere driftsstrømmen, noe som igjen betyr at eksponentialleddet representerer diffusjonsstrømmen. Altså:

$$I_{diff} = I_0 \exp\left(\frac{qV}{kT}\right)$$

$$I_{drift} = -I_0$$



**Figur 16.** Strøm-spenning forhold for ideell diode. Kun foroverbias gir stor strøm, men ved stor bakoverbias kan et dielektrisk sammenbrudd inntreffe. Dette er indikert med stor negativ strøm ved stor negativ bias på figuren.

Som indikert i figur 16 vil dioden få et sammenbrudd ved stor bakoverbias. Dette kommer av at den elektriske feltstyrken i sperresonen kommer over en viss verdi (ca  $10^5 \text{ V/cm}$ ) og vi får et dielektrisk sammenbrudd. Men vi vil også få problemer med stor foroverbias. Dersom vi anvender en foroverbias som er større enn det innebygde potensialet, vil vi også ødelegge dioden. Dette illustreres enkelt ved å tenke på hvordan et energibånddiagram ville sett ut om vi anvender en foroverbias sterkere enn det innebygde potensialet. Det er ganske innlysende at en slik situasjon må føre til en ødeleggelse av dioden siden potensialforskjellene gjør at strømmene ønsker å skifte retning, noe som er i kontrast med egenskapene til n og p siden. Eventuelt kan vi tenke på sperresonen som i en slik situasjon da vil få negativ utstrekning.

### Grenseflaten som en kapasitans

Så langt har vi avdekket at en diode har en sperresone som inneholder en viss mengde ladning, og at ladningen i sperresonen varierer med spenningen over diode. Hvis vi sammenligner med en vanlig kapasitans ser vi helt klare likhetstrekk, bare at vi nå vil ha en kapasitans som varierer med spenningen. En diode som blir laget for å utnytte dette kalles for en varaktor.

Det viser seg at kapasitansen til en slik diode er gitt ved:

$$C = \frac{e_{si}}{W}$$

Som er nøyaktig det samme som for en plate kondensator. Hvis vi ser på en diode der den ene siden har en høy dopingkonsentrasjon i forhold til den andre, en **abrupt one-side junction**, har vi at sperresonen kun ligger på den ene siden av grenseflaten. Der har vi at:

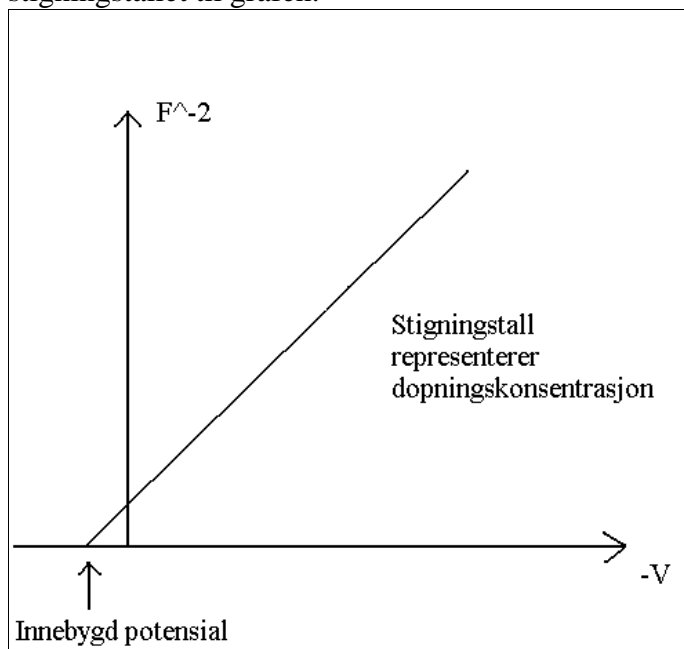
$$W = \left( \frac{2e[V_{ib} - V]}{qN} \right)^{\frac{1}{2}}$$

Der  $V$  er anvendt bias og  $N$  er dopingkonsentrasjonen til siden med lav dopingkonsentrasjon.

Hvis vi nå utfører et forsøk der vi plotter  $\frac{1}{C^2}$  som en funksjon av  $V$ , kan vi finne ut både det innebygde potensialet og dopingkonsentrasjonen.

$$\frac{1}{C^2} = \frac{2(V_{ib} - V)}{qeN}$$

Ut fra formelen over ser vi at  $V_{ib}$  er gitt fra skjæringspunktet med førsteaksen og  $N$  fra stigningstallet til grafen.



**Figur 17.** kapasitans-spenningsdiagram for en varaktor diode

### LED-diode (light emitting diode)

I det vi anvender en forover bias på en p-n diode, vet vi at diffusjonsstrømmen øker. Et resultat av dette er at også rekombinasjonen av elektroner og hull øker. Under denne rekombinasjonen vil det frigjøres energi ved at elektroner faller tilbake til valensbåndet og fyller opp hull. Fra Bohrs teori om atommodeller og tillatte energinivåer vet vi at når et elektron faller til en lavere energitilstand vil det sende ut et foton med en gitt bølgelengde. Derfor vil det sendes ut lys fra en p-n diode når hull og elektroner rekombinerer.

Ved å bruke ulike grunnstoffer og variere mengden av de kan vi fikse størrelsen på båndgapet slik at det vil emitteres lys i ønskede bølgelengder, også infrarøde og ultrafiolette spekteret. Det er derimot umulig å lage en diode som sender ut hvitt lys siden hvitt lys er en samling av alle bølgelengdene i det synlige spekteret. Også enkelte deler av det synlige spekteret har vist seg vanskelig å lage klart lys av. Dette kommer av at det er kvantefysisk større sannsynlig med eksitasjoner over et direkte båndgap, dvs når valensbåndet og ledningsbåndet ikke overlapper hverandre. Spesielt for blått har dette vært et problem, det viser seg vanskelig å lage et direkte båndgap med tilsvarende energi som blått lys. Vanligvis baserer LED-diodene som lyser blått seg på silisium som bare har indirekte båndgap, og som dermed ikke er den mest energetisk gunstige løsningen.

Selv om vi har et indirekte båndgap fins det måter å effektivisere lysemittasjonen på. Dette gjøres best ved å innføre en isoelektrisk kjerne, dvs. å erstatte en dopingkjerne med en annen kjerne fra samme gruppe. En annen metode som igjen er litt mindre effektiv er å innføre andre urenheter i halvlederen. Om det ikke finnes kjente isoelektriske kjerner eller urenheter som gjør lysdioden mer energieffektiv for de ønskede lysfargene, må vi bruke en halvleder med indirekte båndgap.

Et eksempel på en LED-diode er et AlGaAs/GaAs system der vi har ulike materialer på hver side av p-n overgangen, en såkalt heteroovergang. Ved å variere konsentrasjonen av aluminium kan vi bestemme om vi ønsker et indirekte eller direkte båndgap. Så lenge vi ønsker et direkte båndgap i delen som skal emittere lyset må aluminiumskonsentrasjonen være på under 44%, noe som gjør at vi vil få lys med bølgelengder fra 880nm-663nm, altså i det røde og infrarøde spekteret. Som vi ser av *figur 18* vil vi ha en n-side med 70% aluminium og to ulike typer p-sider, der delen med 40% Al vil fungere som lyskilde. Denne n-siden har et større båndgap enn p-siden noe som gjør den til en effektiv injektor av elektroner, mens det samtidig er større sjans for rekombinasjoner i et mindre båndgap. I tillegg vil også fotonene som produseres lettere kunne passere n-siden pga. det store båndgapet. Alt i alt gjør dette at et slikt system vil levere klart rødt lys veldig energieffektivt.

n-side	AlGaAs	70%Al
p-side	AlGaAs	40%Al
p-side	GaAs	

**Figur 18.** hetero pn-overgang

I tillegg til dioder som lyser rødt og blått er også gult, grønt og oransje vanlige farger på lysdioder. Alle disse diodene har vanligvis indirekte båndgap, men bruker en isoelektrisk kjerne for å gjøre lyset klarere og mer energieffektivt.

## Fotodetektorer og solceller

Når vi så på en LED-diode brukte vi at når elektroner faller ned til en lavere energitilstand vil det sende ut et foton. Fra kvantefysikken vet vi at det motsatte også er gyldig, at et foton kan eksitere et elektron gitt at fotonet har nok energi. Det er dette prinsippet som vi bruker i en fotodetektor. I motsetning til en LED-diode trenger en fotodetektor kun riktig båndgap og er ikke avhengig av om det er direkte eller indirekte.

Mekanismen vi bruker *for å detektere* fotoner er at et elektron som eksiteres nær eller i sperresonen vil gi opphav til en strøm som følger av det elektriske feltet i området. Dersom elektronet blir eksitert i de nøytrale områdene av en p-n overgang vil de rekombinere med hull før de når sperresonen, og kan være opphav til en strøm. Derfor vil en god fotodetektor ha en stor sperresone, slik at flest mulig av de eksiterte elektronene vil være opphav til en strøm gjennom dioden. Fra teorien om p-n dioden vet vi at å sende på en stor bakover bias vil vi kunne maksimere sperresonen.

Fra kvantefysikken vet vi at om fotonet ikke har nok energi, vil det ikke være i stand til å eksitere elektroner, men også fotoner med for høy energi vil vi få problemer med å *detektere*. Dette har sammenheng med at de mest energirike fotonene er de som absorberes først i en

halvleder. For å øke effekten til detektoren vil det derfor være naturlig å sørge for en tynt lag på toppen, slik at minst mulig lys absorberes før de kommer til sperresonen.

Dersom vi lar være å anvende en bakover-bias på en fotodetektor og istedenfor plasserer den i en krets der den vil fungere som en spenningskilde vil vi ha en solcelle. Den spenningen en slik diode vil bidra med kalles for fotovoltaisk effekt og den kan vi finne ved å bruke den ideelle diode ligningen.

$$I = I_0(\exp(\frac{qV}{kT}) - 1)$$

Dersom vi løser denne med hensyn på V får vi:

$$V = \frac{kT}{q} \ln(1 + \frac{I}{I_0})$$

Størrelsen av V vet vi må være mindre enn det innebygde potensialet, som vi igjen vet er mindre enn potensialet i båndgapet. Vi kan dermed avgjøre sånn noenlunde hvor mye spenning en slik solcelle på bakgrunn av egenskapene til materialet vi velger å bruke. En standard silikon solcelle vil produsere ca 1 volt og 20mA. Det er derfor vanlig å koble sammen flere celler i serie og parallell for å sørge for en tilfredsstillende energiproduksjon.

Et av de største problemene med en solcelle er at de ønsker å utnytte sollyset som består av et stort spekter frekvenser og bølgelengder, og som vi har sett er det vanskelig å utnytte et slikt stort spekter. Dette gjør at man får en teoretisk virkningsgrad som ligger langt unna 100 %. Det er vanskelig å angi nøyaktig hva som er den teoretiske virkningsgraden, siden det forskes mye på solceller og man hele tiden kan finne nye og bedre teoretisk metoder. Men de høyeste eksperimentelle verdiene som er oppnådd er på rundt 40 %. Et annet problem er at for å få transportert strøm ut av solcellene. For å gjøre dette trenger vi en ohmsk metallkontakt på begge sider av solcella. En slik kontakt vil hindre sollyset i nå ned til selve dioden. Et kompromiss har vært å lage tynne metallfingre på tvers av solcella og knytte disse til en større metallbit på kanten av cella.

## Transistorer

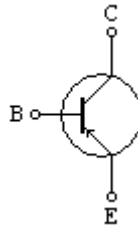
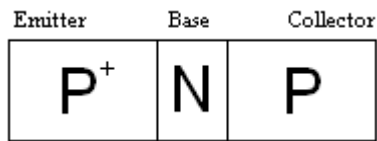
En transistor er en elektrisk komponent som ble utviklet ved Bells Laboratorier i USA. Den første transistoren ble laget i 1948, og tre personer som har fått æren for denne revolusjonerende oppfinnelsen er John Bardeen, Walter Brattain, og William Shockley. Ordet transistor er satt sammen av de engelske ordene transfer og resistor. Direkte oversatt vil ”transfer resistor” bety overførbare elektrisk motstand. Det finnes flere typer transistorer som for eksempel bipolare transistorer (BJT) og felteffekttransistorer (FET).

### Bipolare transistorer

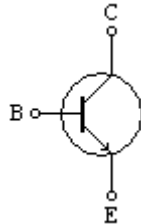
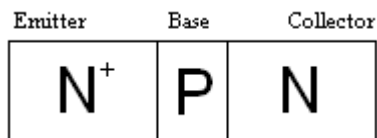
Bipolare transistorer består av to pn-overganger hvor de to overgangene har en felles p eller n region. Denne felles regionen ligger mellom to regioner med motsatt type doping slik at vi enten har en pnp-transistor eller en npn-transistor.

Den midterste regionen, som er vesentlig smalere enn de andre, kalles for transistorens base og de to andre regionene kalles emitter og kollektor. I de aller fleste tilfellene har emitterregionen betydelig større dopingkonsentrasjon enn basen, som igjen har større dopingkonsentrasjon enn kollektorregionen. I elektriske kretser har pnp-transistorer og npn-transistorer hvert sitt skjematisk symbol.





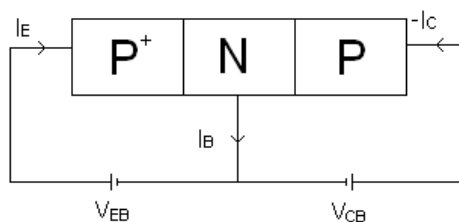
**Figur 19a):** Pnp-transistor og dens symbol i elektriske kretser.



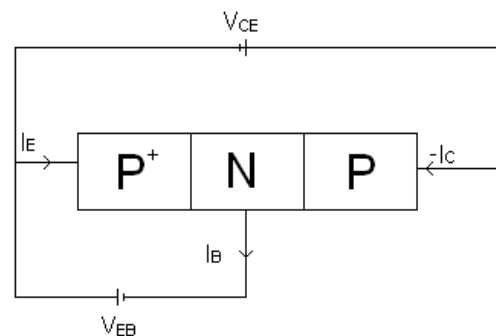
**Figur 19b):** Npn-transistor og dens symbol i elektriske kretser.

### Konfigurasjoner

Når en transistor koples til to forskjellige spenningskilder kan dette gjøres på forskjellige måter og vi skiller blant annet mellom felles base konfigurasjon og felles emitter konfigurasjon. I felles base konfigurasjon koples en spenningskilde over emitter og base, mens en annen spenningskilde koples over base og kollektor. Det eneste felles koplepunkt er altså basen, derav navnet felles base konfigurasjon. I felles emitter konfigurasjon koples da spenningskilder over emitter og base, og over emitter og kollektor.



**Figur 20a):** Felles base konfigurasjon.



**Figur 20b):** Felles emitter konfigurasjon.

## Strøm i en transistor

Vi ser på felles base konfigurasjon av en pnp-transistor. Vi kaller spenningen over emitter og base for  $V_{EB}$  der vi for positiv spenning har høyest potensial på emitter. Spenningen over base og kollektor kaller vi  $V_{CB}$  der vi for positiv spenning har høyest potensial på kollektor.

Strøm i emitteren er sammensatt av strøm av hull fra emitter til base og strøm av elektroner fra base til emitter:  $I_E = I_{Ep} + I_{En}$

Strøm i kollektor er sammensatt av strøm av hull fra base til kollektor og strøm av elektroner fra kollektor til base:  $I_C = I_{Cp} + I_{Cn}$

Kirchoffs knutepunktregel gir oss strøm i basen:  $I_B = I_E - I_C$

Vi definerer felles base strømefaktor:  $\mathbf{a}_0 \equiv \frac{I_{Cp}}{I_E}$

Kollektorstrømmen blir da:  $I_C = \mathbf{a}_0 I_E + I_{Cn} = \mathbf{a}_0 I_E + I_{CB_0}$

$I_{CE0}$  kalles kollektor-base lekkasjestrøm med åpen emitter. Det er altså den strømmen som går mellom kollektor og base når  $I_E = 0$ .

Strømefaktoren kan også skrives  $\mathbf{a}_0 = \frac{I_{Cp}}{I_{Ep} + I_{En}} = \frac{I_{Ep}}{I_{Ep} + I_{En}} \cdot \frac{I_{Cp}}{I_{Ep}} = \mathbf{g} \cdot \mathbf{a}_T$

der vi definerer to nye størrelser:  $\mathbf{g} \equiv \frac{I_{Ep}}{I_E}$  og  $\mathbf{a}_T \equiv \frac{I_{Cp}}{I_{Ep}}$  som er henholdsvis emittereffektivitet

og basetransportfaktor. Emittereffektiviteten forteller oss hvor stor andel av emitterstrømmen som skyldes vandring av hull, mens basetransportfaktoren forteller hvor stor andel av hull som vandrer fra emitter til base, som også vandrer videre til kollektor.

Dersom vi ser på felles emitter konfigurasjon av den samme transistoren vil strøm i emitteren i følge Kirchoffs knutepunktregel bli:  $I_E = I_B + I_C$

Uttrykket for Kollektorstrømmen blir da:  $I_C = \mathbf{a}_0 I_E + I_{CB_0} = \mathbf{a}_0 (I_B + I_C) + I_{CB_0}$

som gir:  $I_C = \frac{\mathbf{a}_0}{1 - \mathbf{a}_0} I_B + \frac{I_{CB_0}}{1 - \mathbf{a}_0} = \mathbf{b}_0 I_B + I_{CE0}$

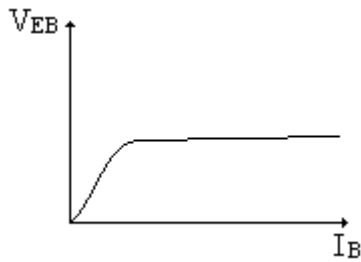
der  $\beta_0$  er en størrelse kalt felles emitter strømforsterkning og som er definert ved:

$$\mathbf{b}_0 \equiv \frac{dI_C}{dI_B} = \frac{\mathbf{a}_0}{1 - \mathbf{a}_0}$$

$I_{CE0}$  kalles kollektor-emitter lekkasjestrøm med åpen base. Det vil si den strømmen som går mellom kollektor og emitter når  $I_B = 0$

Ut i fra felles base konfigurasjonen kan vi konkludere med at  $\alpha_0$  er mindre enn, i "beste fall" nesten lik 1, siden  $I_C$  må være mindre enn  $I_E$ . (Fordi noen av hullene vil rekombinere å bidra til  $I_B$ ) Vi ser at når  $\alpha_0$  går mot 1 så vil  $\beta_0$  være mye større enn 1 noe som vi skal se er avgjørende for transistorens virkning.  $\beta_0$  er typisk i størrelsesorden 10-100.

Med felles emitter konfigurasjon vil vi ha en sammenheng mellom strøm inn på basen og spenningen over emitter og base som skissert nedenfor.

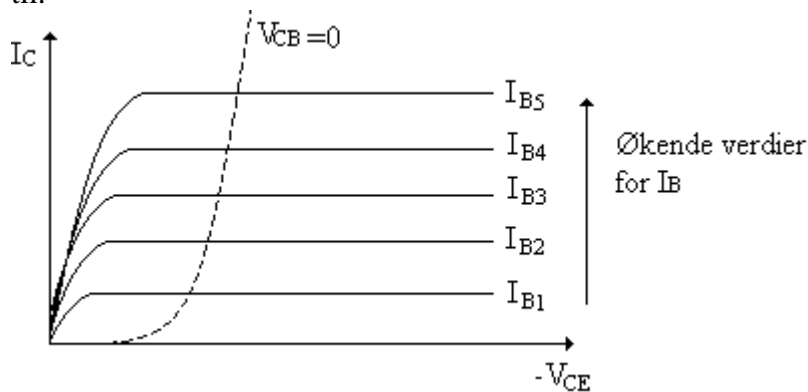


**Figur 21:** Sammenheng mellom basestrøm og emitter-base spenning.

Sammenheng mellom kollektorstrøm og spenningen over kollektor og emitter kan vi tenke oss ved å sammenligne base-kollektorkovergangen med en ideell diode. I det som vi senere vil kalle det aktive område for transistoren, vil vi ha bakover biasert denne pn-overgangen og strøm ut i fra kollektoren. Når vi skisserer strøm-spenning kurven tar vi derfor og speiler strøm-spenning grafen fra den ideelle dioden om begge aksene.

Vi ønsker å skissere hvordan basestrømmen påvirker Kollektorstrømmen og tegner derfor flere kurver i samme koordinatsystem for ulike verdier av  $I_B$ .

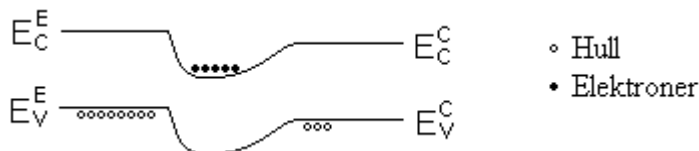
I figuren nedenfor har vi i tillegg tegnet inn grafen for når kollektor-emitter spenningen er like stor som emitter-base spenningen (stiplet linje). Betydningen av dette skal vi komme tilbake til.



**Figur 22:** Sammenheng mellom kollektorstrøm og emitter-kollektor spenning For ulike verdier av  $I_B$ .

### Energibånd for en transistor

Dersom vi betrakter en pnp-transistor som ikke er koplet inn i noen krets, slik at alle tre tilkoplingspunkter ligger på samme potensial, kan vi se på de to pn-overgangene hver for seg og bruke samme argumenter som for den enkle pn-overgangen til å tegne energibåndene for transistoren. Vi må nå ta hensyn til det innebygde potensialet i både emitter-base overgangen og base-kollektor overgangen.



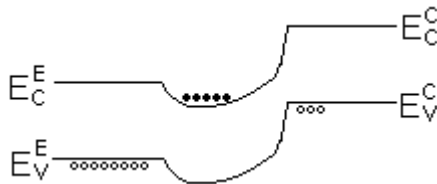
**Figur 23:** Pnp- transistor i likevekt utenfor krets.

## Operasjonsmodi

Vi fortsetter å se på pnp-transistoren, men denne gangen koplet til to spenningskilder i en felles base konfigurasjon.

Vi kan veksle mellom å ha  $V_{EB}$  og  $V_{CB}$  positiv og negativ slik at vi har fire ulike kombinasjonsmuligheter som gir fire ulike operasjonsmodi for transistoren. Ved å bruke det vi vet om de enkelte pn-overgangene kan vi se hvordan de forskjellige kombinasjonene av fortegn på spenningene vil påvirke energibåndene til transistoren og dermed vil vi kunne si noe om hva som karakteriserer de fire operasjonsmodiene.

- Med positiv  $V_{EB}$  og negativ  $V_{CB}$  vil emitter/base-overgangen være forover biasert, mens kollektor/base-overgangen vil være bakover biasert. Dette betyr at vi får en lavere energibarriere for hull mellom emitter og base, slik at hull vil begynne å vandre fra emitter til base. I basen vil vi nå kunne få rekombinasjon av elektroner og hull, men med tilstrekkelig smal base vil mange hull kunne passere basen og ”falle ned” i kollektoren. Det vil altså begynne å gå en relativt stor strøm fra emitter til kollektor. Vi sier nå at transistoren er i aktivt modus.



**Figur 24:** Energibånd for aktiv transistor.

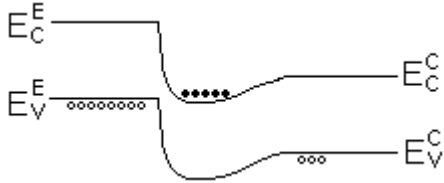
- Med både  $V_{EB}$  og  $V_{CB}$  positiv vil begge pn-overgangene i transistoren være forover biasert. Det vil si at vi får lavere energibarrierer både mellom emitter og base og mellom kollektor og base. En liten økning i spenningene i forhold til likevekt vil føre til store strømmer, så transistoren oppfører seg som et knutepunkt og vi sier at vi har metning.



**Figur25:** Energibånd for transistor ved metning.

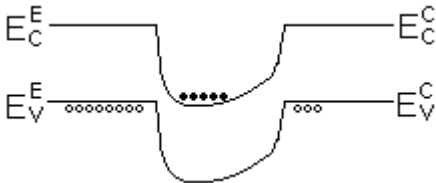
- Med negativ  $V_{EB}$  og positiv  $V_{CB}$  vil emitter/base-overgangen være bakover biasert, mens kollektor/base-overgangen vil være framover biasert. Vi ser at situasjonen er veldig lik situasjonen for en aktiv transistor der vi nå får en lav energibarriere for hull mellom kollektor og base, slik at hull kan begynne å vandre. Forskjellen ligger her i dopingkonsentrasjonen for de ulike lagene. Kollektoren har en vesentlig mindre

konsentrasjon av doping enn emitteren slik at strømmen gjennom transistoren blir betydelig lavere. Vi kan betrakte situasjonen som om vi har koplet en aktiv transistor ”feil” vei, og sier at vi har invertert modus.



**Figur 26:** Energibånd for invertert transistor.

- Med både  $V_{EB}$  og  $V_{CB}$  negativ vil begge pn-overgangene i transistoren være bakover biasert, og vi får høye energibarrierer både mellom emitter og base, og mellom kollektor og base. Resultatet er små strømmer og vi sier at transistoren er i cut-off.



**Figur 27:** Energibånd for transistor i cut-off.

På samme måte vil vi kunne finne ut at for en npn-transistor får vi:

- Med  $V_{EB}$  negativ,  $V_{CB}$  positiv: aktiv transistor
- Med  $V_{EB}$  negativ,  $V_{CB}$  negativ: metning
- Med  $V_{EB}$  positiv,  $V_{CB}$  negativ: invertert transistor
- Med  $V_{EB}$  positiv,  $V_{CB}$  positiv: cut-off

Det samme vil også gjelde om vi ser på en felles emitter konfigurasjon, men da må vi huske at spenningsforskjellene over emitter og base og over kollektor og base, henholdsvis  $V_{EB}^*$  og  $V_{CB}^*$ , er avhengige av fortegnene til både  $V_{EB}$  og  $V_{CE}$  som på figuren representerer spenning fra spenningskildene. I figuren som viser strøm-spenning karakteristikk for transistoren har vi tegnet inn grafen for  $V_{EB}^* = V_{CE}$  eller ekvivalent  $V_{CB}^* = 0$ . Til høyre for denne kurven vil da den aktive regionen for transistoren befinne seg, mens vi til venstre har metning.

## Anvendelser av transistorer

Når en transistor anvendes i en elektrisk krets har den egenskaper som gjør at den kan brukes enten som en bryter, eller som en forsterker av elektriske signal. Den vanligste konfigurasjonen er den vi har kalt felles emitter, og vi skal derfor se på hvilke sammenhenger denne kan benyttes og hvordan bruken av denne samsvarer med de transistoregenskapene vi allerede har nevnt.

Strømmen i basen  $I_B$  kalles i praktiske situasjoner ofte for kontrollstrømmen da det er denne strømmen vi kontrollerer og som bestemmer hva som skal foregå i transistoren når det er en kjent spenning  $V_{CE}$  over emitter og kollektor. At vi kan bruke en strøm til å kontrollere hva som skjer, når det som vi tydelig har sett er spenninger som bestemmer transistorens oppførsel, skyldes at strømmen  $I_B$  vil øke i takt med spenningen  $V_{EB}^*$

### Transistor som forsterker

Da vi så på sammenhengen mellom strømmer i transistoren så vi at vi hadde en sammenheng mellom kontrollstrømmen  $I_B$  og den kontrollerte strømmen  $I_C$  gitt ved:  $I_C = \beta_0 I_B + I_{CE0}$ .

Under visse forutsetninger er  $\beta_0$  en konstant slik at vi har en lineær sammenheng mellom kontrollstrøm og kontrollert strøm. Det vi da kan gjøre er å sende signalet vi ønsker å forsterke som kontrollstrøm. Strømmen gjennom transistoren vil da være en forsterket utgave av kontrollstrømmen der mønsteret i signalet er bevart.

Et eksempel på en situasjon der det kan være ønskelig å forsterke signal er dersom vi ønsker å måle lysintensiteten fra en svak lyskilde ved hjelp av et amperemeter koplet til en solcelle der sammenhengen mellom lysintensitet og generert strøm er kjent. Dersom amperemeteret vi bruker er beregnet for sterkere strømmer vil det være mulig å kople strømmen fra solcellen til basen på en transistor, mens vi lar amperemeteret måle kollektorstrømmen. Amperemeteret vil da gjøre et større utslag. Dersom vi kjenner sammenhengen mellom endring i  $I_B$  og  $I_C$  for transistoren, det vil si at  $\beta_0$  er kjent, kan vi beregne styrken på det opprinnelige signalet og regne om til lysintensitet.

Dersom vi ønsker ytterligere forsterkning av et elektrisk signal, kan man ganske enkelt lede Kollektorstrømmen fra en transistor A inn på basen til en annen transistor B. Samlet strømforsterkning blir da produktet  $\beta_{\text{tot}} = \beta_{0A} * \beta_{0B}$ .

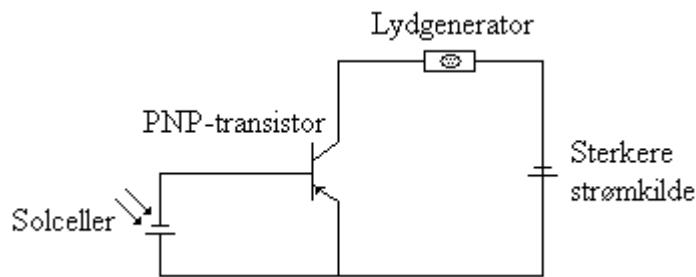
Transistorer brukes, på lignende måte som i det enkle eksempelet, i alle slags forsterkere. Transistorer brukes da riktignok i kombinasjon med flere komponenter i langt mer avanserte kretser for å oppnå ønsket kvalitet på det forsterkede signalet.

### Transistor som bryter

Som vi har sett vil det ikke kunne gå mye strøm gjennom en transistor når den settes i "cut off". Når en transistor brukes som bryter vil derfor denne tilstanden være ekvivalent med at bryter står i av-posisjon. For å skru på "bryteren" må vi gjøre noe med transistoren slik at strøm kan passere gjennom transistoren. Vi kan da sette den i aktivt modus eller i en tilstand med metning. Dersom vi antar at det er en kjent negativ spenning  $V_{CE}$ , kan vi ved å sørge for at  $I_B=0$  (og dermed  $V_{EB}^*=0$ ) skru "bryteren" av. For å skru "bryteren" på er det bare å sende en kontrollstrøm inn på basen. Dersom denne kontrollstrømmen er så stor at  $V_{EB}^*$  blir større enn  $V_{CE}$  vil  $V_{CB}^*$  bli positiv slik at vi får metning.

Strømmen vi ønsker å skru på er typisk betydelig større enn kontrollstrømmen.

Vi kan, i et tenkt eksempel, forestille oss at vi har lyst til å bli vekket av en eller annen elektrisk lydgenerator ved soloppgang. Vi har tilgang til et lite solcellepanel som ikke kan generere nok strøm til å drive denne generatoren, men vi er også i besittelse av en sterkere strømkilde som er i stand til å få generatoren til å virke. Vi kan da kople inn en transistor slik at solcellepanelet styrer strømmen inn på basen, mens den sterkere strømkilden er koplet over kollektor og emitter. Lydgeneratoren kople vi i serie med Kollektorstrømmen. Når sola står opp, vil solcellepanelet generere en liten strøm som åpner for at den sterkere strømmen kan gå gjennom transistoren og få den hjemmelagede vekkerklokka til å ringe.



**Figur 28:** Hjemmelaget vekkerklokke.

De fleste forbinder trolig ordet transistor med datamaskiner. I datamaskiner er det transistorens bryteregenskaper som utnyttes i såkalte logiske kretser. I logiske kretser opereres det med verdiene 0 og 1, representert av transistorer som er skrudd av og på. Transistorer veksler da mellom cut-off og metning.

## Konklusjon

Vi har sett hvordan denne teknologien har sitt utspring i kvantefysisk argumentasjon, hvordan kvantiserte energitilstander for elektroner i atomer fører til den såkalte båndteorien, og igjen sett hvordan dette utnyttes i elektroniske kretser.

Oppfinnelsen av den integrerte kretsen på slutten av 50-tallet er en hendelse som har revolusjonert teknologien og preget hele samfunnet vi lever i. En integrert krets er enkelt sagt en samling av sammenkoplete transistorer, dioder, kondensatorer og elektriske motstander fabrikkert på ett enkelt stykke silikon, også kalt en chip. En slik chip inneholder flere millioner komponenter pr kvadratcentimeter. Integrerte kretser, og dermed også transistorer, brukes over alt rundt oss, som for eksempel i datamaskiner, mobiltelefoner, biler og kameraer. Dersom man klarer å forstå hovedprinsippene bak halvlederkomponenter som transistorer og dioder, vil man derfor samtidig forstå mye av den moderne verden vi lever i.

## Referanseliste

1. L.Solymar, D. Walsh: "Electrical properties of materials"
2. B. Tuck, C. Christopoulos: "Physical electronics"
3. H. Young, R. Freedman: "University Physics"
4. Serway, Moses, Moyer: "Modern Physics"
5. [http://quantummechanics.ucsd.edu/ph130a/130\\_notes/node156.html](http://quantummechanics.ucsd.edu/ph130a/130_notes/node156.html)
6. P.C. Hemmer; "Kvantemekanikk"
7. P.C. Hemmer: "Faste stoffers fysikk"
8. C. Kittel: "Introduction to solid state physics"
9. G. Parker: "Introductory semiconductor device physics"
10. G. Glasford: "Analog electronic circuits"
11. A. Ferendeci "Solid state and electron devices"
12. [www.allaboutcircuits.com](http://www.allaboutcircuits.com)
13. <http://www.britneyspears.ac/lasers.htm>