

## Fasta tillståndets fysik

## Günter Grossmann Fasta tillståndets fysik, LTH 2019

Delar av texten har reviderats av Carina Fasth, Dan Hessman och Martin Leijnse.

Omslagsbild av Magnus Heurlin. Bilden föreställer nan<br/>otrådar av indiumfosfid (InP) odlade i ett hexagonalt mönster.

# Innehåll

1	Inle	ledning		
	1.1	Elektronen "upptäcks"1897	10	
	1.2	Elektronens barndom	14	
	1.3	Elektronerna i en metall som en klassisk gas (P Dru- de, 1900)	16	
	1.4	Svartkroppsstrålning – M Planck (1900) $\ldots$	18	
	1.5	Fotoelektriska effekten	21	
	1.6	A. Einstein 1905	22	
	1.7	Röntgenstrålning upptäcks (WC Röntgen 1895)	24	
	1.8	Atomernas struktur	30	
	1.9	Värmekapacitet	34	
<b>2</b>	Elel	ktronen som partikel	51	
	2.1	Elektroner i metaller som en klassisk gas	51	
	2.2	Driftström	53	
	2.3	Diffusion (i en dimension)	57	
	2.4	Halleffekt	58	
	2.5	Värmeledningsförmåga	60	
	2.6	Värmekapacitet	61	
	2.7	Var kommer $E_{kin} = \frac{3}{2}kT$ ifrån?	63	
3	Elel	ktronen som våg	65	
	3.1	Davisson-Germer interferensexperiment	65	
	3.2	Schrödingerekvationen	66	

## INNEHÅLL

	3.3	Stationära lösningar		
	3.4	Fria elektroner		
	3.5	$Vågpaket \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots $		
	3.6	Vågpaketets rörelse		
	3.7	Dispersion		
4	Frie	lektronmodellen 83		
	4.1	Elektron i en lådpotential		
	4.2	Lådpotential i tre dimensioner		
	4.3	Planvågslösningar		
	4.4	Periodiska randvillkor		
	4.5	Tillståndstäthet		
	4.6	Fermi-Dirac fördelningen		
	4.7	Klassisk gräns hos en elektrongas		
	4.8	Värmekapacitet		
	4.9	Konduktivitet - en andra gång		
	4.10	Slutsatser		
	4.11	Tillståndstäthet igen - nu med fixa randvillkor! 110		
	4.12	Utträdesarbete		
	4.13	Övergång mellan två metaller med olika utträdesar- beten 111		
	4.14	Termisk elektronemission		
<b>5</b>	Ban	dstruktur 115		
	5.1	Frielektronmodellen (FEM) - en påminnelse 115		
	5.2	Introduktion till nästan-frielektronmodellen		
		5.2.1 Röntgeninterferenser		
		5.2.2 Tillbaka till elektronen		
	5.3	Nästan-frielektronmodellen		
		5.3.1 Antaganden och ansats		
		5.3.2 Lösning av SE med periodisk potential 121		
		5.3.3 Resultat för olika val av $k$		

4

		5.3.4 Val av vågvektorer $\ldots \ldots \ldots \ldots \ldots \ldots 12$	7	
		5.3.5 Allmän periodisk potential	9	
		5.3.6 Bandstuktur i 3D $\ldots$ 130	0	
	5.4	Rörelseekvationen för elektroner i en kristall 13	3	
	5.5	Metaller - Isolatorer	5	
	5.6	En en-dimensionell kedja av atomer	7	
		5.6.1 Randvillkor: ändlig kedja - stående vågor 138	3	
		5.6.2   Periodiska randvillkor - fortskridande vågor  . 140	0	
		5.6.3 Antal tillstånd per band	2	
	5.7	Bundna tillstånd	3	
	5.8	Modeller för bandstruktur - sammanfattning 140	6	
6	Hal	vledare 14'	7	
U	c 1			
	0.1	Elektroner i bandstrukturen	(	
	6.2	Hål	1	
	6.3	Optiska övergångar	3	
	6.4	Laddningsbärarkoncentrationer	7	
	6.5	Massverkans lag	C	
	6.6	Intrinsiska halvledare	C	
	6.7	Dopning	1	
	6.8	Extrinsiska halvledare	3	
	6.9	Elektron- och hålströmmar i halvledare	9	
	6.10	Spridningsprocesser	9	
		6.10.1 Spridning mot atomära vibrationer 17	1	
		6.10.2 Spridning mot joniserade orenheter 172	2	
	6.11	Generation och rekombination – livstider	3	
	6.12	Djupa störställen	4	
	6.13	Kommentar till bundna tillstånd och bandstruktur . 173	5	

7	pn-	övergången 17		
	7.1	7.1 pn-övergången i jämvikt		
		7.1.1	Ett tankeexperiment	177
		7.1.2	Detaljerad härledning	181
		7.1.3	Ström vid jämvikt?	187
	7.2	pn-övergången med pålagd spänning		
		7.2.1	pn-övergången som likriktare	188
		7.2.2	Var hamnar spänningsfallet?	190
		7.2.3	Bidrar minoritetsladdningsbärarna till ström- men i de neutrala områdena?	191
		7.2.4	Ideala diodekvationen	192
		7.2.5	Backspänning	200
	7.3	Spärrs	skiktskapacitans (backspänning)	202
	7.4	Ström	n-spänningskarakteristik	203
	7.5	7.5 Generation och rekombination i utarmningsområde		208
		7.5.1	Framspänning	208
		7.5.2	Backspänning	209
	7.6	Ekviv	alent småsignalschema	210
	7.7	Genor	mbrott vid backspänning	210
8	Den	ı bipol	lära transistorn	213
	8.1	.1 Vi bygger en strömkälla		
	8.2	Hålens väg genom basen		
	8.3	3 Elektroner injiceras från bas till emitter		
	8.4	Två pn-övergångar – en transistor		217
	8.5	Tillverkning       2         Sammanfattning       2         betssätt       2		
	8.6			
	8.7	Spänn	ningsförstärkning - gemensam-bas koppling	221
	8.8	$\operatorname{Geme}$	nsam-emitter koppling (npn transistor) $\ldots$ .	223
	8.9	Invert	eraren	226

## INNEHÅLL

9	Metall-halvledarövergången		
	9.1	Metall-halvledarövergången med pålagd spänning	233
10	Metall-Oxid-Halvledarövergången		
	10.1	MOS-övergången med pålagd spänning	238
		10.1.1 Ackumulation $\ldots$ $\ldots$ $\ldots$ $\ldots$ $\ldots$ $\ldots$	238
		10.1.2 Utarmning	239
		10.1.3 Inversion	240
	10.2	Charge-coupled device (CCD)	241
	10.3	MOSFET	245
	10.4	CMOS	249
$\mathbf{A}$	Vär	meledningsförmåga	251
в	3 Lådpotential i ett elektriskt fält		253
С	C Diffusionsekvationen		255
D	) Generation och rekombination – livstider		259
$\mathbf{E}$	Kon	stanter	263

INNEHÅLL

## Kapitel 1

## Inledning

Ibland läser man att fysikerna mot slutet av 1800-talet ansåg att fysikens gåtor i stort sett hade lösts. Mekanik, elektrodynamik och termodynamik – dessa tre områden som numera utgör den så kallade klassiska fysiken – bildade en stabil grund som kring mitten av seklet hade avrundats med lagen om energins bevarande. Visst kvarstod kanske en hel del smärre frågor, men var man en lovande ungdom skulle man nog helst söka utmaningarna på annat håll. Så skrev t ex den amerikanske fysikern A Michelson 1899:

The more important fundamental laws and facts of physics have all been discovered, and ... the possibility of their ever being supplanted in consequence of new discoveries is exceedingly small."

Om nu denna uppfattning verkligen var allmänt accepterad är kanske tveksamt, men på den tiden med sin oerhört snabba tekniska utveckling var det lätt att vara optimist. Landvinningarna inom t ex elektroteknik och kemi gick hand i hand med stora industriella satsningar och framtidsutsikterna verkade lovande. Det fanns en förhoppning att alla världens problem skulle kunna lösas lika rationellt som de hade lösts inom vetenskap och teknik.

Den första fyrtaktsmotorn byggs 1862, den första dieselmotorn 1897. Bensindrivna bilar börjar dyka upp efter 1886. Berlin får sin första tunnelbana och sin första kemtvätt. Julen 1895 öppnas Paris första biograf av bröderna Lumière. Edison erövrar marknaden med sin glödlampa och gatubelysning installeras på utvalda platser. Elektricitet börjar bli allmänt tillgänglig och Ferdinand Braun uppfinner det Braunska röret 1897, som med tiden utvecklas till (går-)dagens TV-rör. Vardagen förändras radikalt för vanligt folk under denna relativt långa period av fred i Europa kring sekelskiftet 1900 och



Mars 1888



September 1888



December 1888



Figur 1.1: Paris får sitt Eiffeltorn 1889.

man förväntar sig en fortsatt stark och framför allt kontinuerlig utveckling, både intellektuellt och materiellt, snarare än någon revolution.

Det politiska läget förefaller vara stabilt. I England regerar drottning Victoria, hon kröntes redan 1837 och sedan 1876 är hon även kejsarinna av Indien. Det Tyska Riket, nybildat 1871, blomstrar under kejsar Wilhelm II, Victorias barnbarn, och även USA har ekonomiskt växt ikapp de forna stormakterna England och Frankrike. Den snabbt ökande industriella produktionen leder emellertid snart till ökande konkurrens mellan Europas stater med en betydande kapprustning som följd. Så försöker t ex Tyskland att etablera sig som stormakt i världen genom att förvärva kolonier och planera byggandet av en mäktig flotta som skulle kunna mäta sig med Royal Navy. Samtidigt ökar oron bland arbetarna som bildar allt starkare arbetarrörelser. Fredsperioden som hade varat sedan 1871 slutar i första världskrigets katastrof och de följande sociala revolutionerna ritar om Europas karta, politiskt och ekonomiskt.

Inom vetenskapen, särskilt inom fysiken, står samtidigt dörren till en helt ny värld – framför allt atomernas inre struktur – redan på glänt under åren strax före sekelskiftet 1900. Ett antal avgörande upptäckter skulle de följande åren dramatiskt förändra vår bild av världen från grunden. Vi ska i denna inledning inte ens försöka skissera alla drag i denna omtumlande utveckling utan begränsa diskussionen till ett fåtal experiment och teoretiska aspekter som även utgör grunden för dagens beskrivning av fasta material.

Vi börjar denna historik med elektronen, den centrala figuren i elektronikens utveckling. När den identifierades togs det första steget för att förstå atomers struktur, men samtidigt också ett stort steg inom telekommunikationstekniken, dvs utvecklingen av vakuumeller radiorör. Dessa komponenter skulle så småningom byggas i halvledarmaterial, radikalt minskas i storlek, och leda till dagens miniatyriserade elektronik.



Figur 1.2: Elektriska urladdningar studeras i ett evakuerat glasrör med två elektroder, den negativa katoden K och den positiva anoden A.

## 1.1 Elektronen "upptäcks"1897

Ända sedan Faradays inledande experiment på 1830-talet studerar fysiker under senare hälften av 1800-talet hur elektricitet passerar genom gaser. Genom sådana experiment hoppas man få svar på frågan vad elektricitet egentligen är. Är det en fluid som flyter från ett ställe till ett annat, eller...?

Under normala omständigheter leder gaser ingen ström, som alla vet är luft vid atmosfärstryck en god isolator. Vi räknar ju inte med att ett batteri laddas ur där det ligger på bordet med oskyddade poler. För att en ström skall kunna flyta måste det finnas laddningar som kan transporteras. I gaser kan visserligen laddningar skapas av den joniserande strålning som alltid är närvarande, men den resulterande strömmen är ytterst liten och försumbar i sammanhanget. För att få en observerbar effekt hade man funnit att det krävdes förhållanden som åtminstone i den tidens laboratorier var rätt extrema, dvs höga elektriska fält och låga tryck. Då kunde det uppstå en elektrisk urladdning med mätbara strömmar och tydliga ljusfenomen. Blixtar är en relativt vanlig form av elektrisk urladdning men mindre lämpade för detaljerade experimentella studier. Man föredrar mer kontrollerbara förhållanden och använder sig av ett slutet glasrör med två elektroder, spänningskällans pluspol ansluts till anoden, minuspolen till katoden.

Sänker man trycket till storleksordningen 100 Pa, ca en tusendels atmosfärstryck, och lägger en spänning av storleksordningen 1000 V mellan elektroderna, kommer gasen i röret att börja lysa och strömmen ökar markant. Denna effekt utnyttjar man fortfarande i t ex lysrör och "neonrör". I detta läge måste det på något sätt skapas betydligt fler laddningar i gasen. Ett lägre tryck, p, innebär att antalet gaspartiklar per volymsenhet, N/V, minskar om vi håller temperaturen T konstant. Enligt ideala gaslagen gäller ju

$$pV = NkT$$
 eller  $\frac{N}{V} = \frac{p}{kT}$ .

Därmed ökar den fria medelväglängden i gasen, dvs en partikel kan röra sig en längre sträcka innan den kolliderar med en annan partikel. Tiden mellan två kollisioner blir längre eftersom partiklarnas medelhastighet förblir densamma vid konstant temperatur. Det elektriska fältet kan då under denna längre tid accelerera laddade partiklar till högre hastighet och därmed högre kinetisk energi. Och när partiklar med hög kinetisk energi krockar kan laddningar frigöras. Två kollisionsprocesser dominerar skeendet. På sin väg genom röret accelereras elektronerna så kraftigt av det elektriska fältet att de kan jonisera gasatomer när de krockar mot dem. Då skapas ytterligare elektroner och även positiva joner (figur 1.3). Det ökade antalet elektroner joniserar i sin tur ännu fler atomer, vi får en lavinartad ökning av antalet laddningsbärare, i det slutliga stationära tillståndet begränsas strömmen endast av motståndet R (se figur 1.2). I den andra viktiga kollisionsprocessen krockar de positiva jonerna med katoden, där de neutraliseras och även förmår slå ut sekundära elektroner som då börjar sin resa mot anoden (figur 1.4).

Ljuseffekterna i röret beror på att elektroner även om de inte joniserar en atom likväl kan excitera denna och när atomen återvänder



Figur 1.3: En elektron (svart) genererar 3 nya elektroner och tre positiva joner i kollisioner med neutrala partiklar (streckade).



Figur 1.4: En positiv jon neutraliseras vid katoden och slår samtidigt ut sekundära elektroner.

till sitt grundtillstånd gör den sig gärna av med energi genom emission av fotoner. Det utsända ljusets färg bestäms av vilken sorts gas som finns i röret.

Sänker man nu trycket i gasen ytterligare, till ungefär 0,1 Pa (ca  $10^{-6}$  atm), och håller spänningen tillräckligt hög för att upprätthålla urladdningen, försvinner ljuset och i stället observerar man blågrönt fluorescensljus från rörets glasväggar i närheten av anoden. Fortfarande slår joner ut sekundärelektroner från katoden som nu kan röra sig raka vägen genom röret, väsentligen utan att krocka med atomer. Fångas de ej upp av anoden träffar de glasväggarna där de exciterar fluorescensen, dvs atomer exciteras och emitterar ljus på vägen tillbaka till grundtillståndet.

Dessa strålar som utgår från katoden, *katodstrålarna*, hade observerats redan 1859 av J Plücker och senare av W Hittorf och andra. Dessa experiment hade endast blivit möjliga tack vare utvecklingen av väsentligt bättre (kvicksilver-) pumpar och högspänningskällor, och att man lyckats täta genomföringarna till elektroderna i glasröret. Man hade funnit att katodstrålen kunde avlänkas av en magnet i närheten av röret, och att strålen förde med sig negativ laddning som kunde fångas upp av en liten metallcylinder placerad i glasröret (JB Perrin). Försök att avlänka strålen med ett elektriskt fält hade dock inte gett några entydiga resultat (H Hertz). Vidare hade P Lenard konstaterat att strålen lätt kunde passera genom skärmar av olika material.

Men vad dessa strålar egentligen bestod av kunde man bara gissa sig till. England föredrog korpuskelteorin, dvs en ström av partiklar, Tyskland lutade mot "eteriska fenomen", dvs att strålen utbredde sig som en våg. Etern var ju den fluid som antogs fylla hela universum och som var det medium man trodde elektromagnetiska vågor utbredde sig i.

1897 kom den förbryllande upplösningen. JE Wiechert, W Kaufmann och JJ Thomson kommer fram till att strålen består av partiklar, alla tre finner liknande värden för förhållandet mellan laddningen och massan hos partiklarna. Antar man att laddningen är lika stor som den minsta laddning man hade uppmätt i elektrolys, elementarladdningen e, så följer att partikelns massa måste vara ca 2000 gånger mindre än väteatomens massa!

JJ Thomson, vars namn oftast förknippas med elektronens upptäckt, hade vid 28 år ålder utsetts till chef för Cavendish laboratoriet i Cambridge som efterträdare till Maxwell och Lord Rayleigh. Han studerar gasurladdningar, magnetiska fält. Ur avlänkningen kan han bestämma förhållandet e/m och påvisar därmed partikelkaraktären hos katodstrålarna. Resultatet är oberoende av vilket



Figur 1.5: Ett av de urladdningsrör Thomson använde. Katodstrålar utgår från katoden K, passerar en spalt i anoden A och en ytterligare kollimerande spalt i B. Därefter rör de sig mellan två aluminiumplattor D och E som kan anslutas till polerna av ett batteri för att avlänka strålen. Ett magnetfält kan läggas vinkelrätt mot det elektriska fältet i området mellan plattorna.

material katoden är gjord av eller vilken gas som finns i röret, samma partikel bör då finnas i all materia. Något år senare, 1899, lyckas hans student CTR Wilson med hjälp av en dimkammare bestämma e direkt och därmed är även m bestämd. (RA Millikan bestämde emed en promilles noggrannhet åren kring 1910.) Elektronen visar sig vara drygt 1800 gånger lättare än väteatomen. Atomen som sedan Demokritos dagar hade antagits vara odelbar visade sig bestå av betydligt mycket mindre delar. Elektronen blev den första elementarpartikeln man upptäckt, och den är fortfarande "elementär" eller odelbar. Namnet fick partikeln av den irländske fysikern GJ Stoney som därmed ursprungligen endast avsåg elementarladdningen e.

Hur gick Thomson nu tillväga med sitt experiment? Han använde ett urladdningsrör som visas schematiskt i figur 1.5. En spänning Upå tusentals volt läggs mellan katod K och anod A. Trycket i röret är så lågt att elektronerna som slås ut ur katoden nästan obehindrat kan röra sig rätlinjigt genom röret. Det elektriska fältet mellan anod och katod accelererar elektronerna, spalten B kollimerar strålen. 1895 hade H Lorentz tecknat kraften **F** som verkar på en laddning qsom befinner sig i ett elektriskt fält **E** och som rör sig med hastighet **v** i ett magnetiskt fält **B** 

Lorentz-kraften:  $\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B})$ 

Elektronernas rörelse bestäms då av Newtons ekvation

 $m\mathbf{a} = \mathbf{F} = -e(\mathbf{E} + \mathbf{v} \times \mathbf{B})$ 

som innehåller tre obekanta storheter, massan m, laddningen -e och hastigheten  $\mathbf{v}$ . Thomson bestämmer först  $\mathbf{v}$  genom att låta strålen passera korsade elektriska och magnetiska fält. som han väljer så att strålen passerar utan att avlänkas. Han lägger en spänning



Figur 1.6: En elektron passerar korsade elektriska och magnetiska fält. Väljs fälten så att avlänkningen försvinner, kan elektronens fart bestämmas.

mellan plattorna D (+) och E (-), det elektriska fältet **E** mellan plattorna pekar nedåt, kraften på elektronen uppåt i positiv y-riktning, se figur 1.6. Kraftens belopp är

 $F_{el} = e\mathscr{E},$ 

där vi med  $\mathscr{E}$  betecknar den elektriska fältstyrkans belopp för att undvika förväxling med energin E. Magnetfältet genereras av spolar utanför röret och är (ungefär) begränsat till området mellan plattorna. B pekar in i figuren, i negativ z-riktning, och kraften med belopp

$$F_{magn} = ev_x B$$

pekar nedåt i figuren. Om fälten är valda så att ingen avlänkning sker måste krafterna ta ut varandra, dvs det måste gälla att

$$e\mathscr{E} = ev_x B$$

och därmed blir

$$v_x = \frac{\mathscr{E}}{B}$$

Nu tar Thomson bort magnetfältet och studerar avlänkningen i enbart det elektriska fältet. Eftersom han redan vet elektronernas hastighet, ger avlänkningen i y-led honom, tillslut, ett uttryck för e/m som han bestämmer till

$$\frac{e}{m}=0.7\cdot 10^{11}~{\rm C/kg}$$

### 1.2 Elektronens barndom

När katodstrålar identifierades som en ström av elektroner tog man det första steget mot en djupare förståelse av materia som skulle bli ett huvudtema i fysik och kemi under 1900-talet. Men den omedelbara tillämpningen av katodstrålar inom telekommunikationsteknik skulle dramatiskt förändra världen inom några år efter sekelskiftet. Redan 1910 skulle man kunna höra Caruso sjunga opera på radio.

Det viktigaste steget mot praktiskt användbara radiorör togs när man upptäckte att elektroner mycket lättare kunde lämna katoden om denna var upphettad. Då behövde man inte längre accelerera jonerna för att slå ut elektronerna ur katoden. Denna effekt observerade A Edison under sina experiment där han försökte förbättra glödlampan. Han hade en gång introducerat en extra elektrod i form av ett litet metallbleck i glödlampan och upptäckte att det flöt en



Figur 1.7: En diod fungerar som en likriktare. Elektroner kan emitteras av den heta katoden men inte av den kalla anoden.

ström mellan glödtråd och bleck om blecket var positivt, men strömmen försvann om blecket var negativt. Systemet fungerade som en likriktare. Elektroner kunde lämna den heta (negativa) glödtråden och vandra till blecket (anoden), men inte lämna det kalla (negativa) blecket. 1884 publicerade han resultaten. 1901 framlade OW Richardson sin teori för denna termiska emission och fick Nobelpriset 1928.

För den termiska emissionen spelar dock inte bara temperaturen utan även katodens materialsammansättning en stor roll, olika material kräver olika utträdesarbeten som elektronen måste uträtta för att kunna lämna katoden. När Wehnelt belade platina-tråden i sin katod med bariumoxid kunde han därmed sänka spänningen över röret från ca 30 000 V som var brukligt i Röntgenrör till ungefär 100 V. JA Fleming utvecklade sedan den första dioden – se figur 1.7 – som kunde användas som likriktare för högfrekventa växelströmmar. Som i Edisons glödlampa strömmade elektronerna från glödkatoden till anoden men inte tvärtom. Denna diod ersatte då delvis kristalldetektorn som var den första "solid-state"-komponenten och som användes som detektor i radions barndom. Kristalldetektorn bestod av en halvledare med en instucken metallspets. En sådan metall-halvledarövergång kan vara likriktande som vi kommer att se senare.

Hur kan nu en likriktare vara till hjälp för att detektera radiosignaler? Vid överföring av radiosignaler krävs ett budskap man vill dela med sig av, t ex en lågfrekvent ljudsignal omvandlad till en spänning som varierar med tiden, som i figur 1.8 (a). Genom addition av en konstant spänning erhålls en enbart positiv signal (b). Sändning och mottagning av radiosignaler visar sig fungera bäst när antennen har en längd av storleksordningen en fjärdedels våglängd. Sänder man signal (b) direkt behöver mottagaren antenner som kanske är många kilometer långa. Man kan kringgå problemet genom att låta signalen modulera amplituden på en högfrekvent, och därmed kortvågig bärvåg (c) och sänder i stället signalen (d). Mottagaren återställer nu den ursprungliga signalen med hjälp av en likriktare, som endast släpper igenom ström i en riktning och ger en pulserande signal (e), och ett lågpass- filter som endast släpper igenom låga frekvenser.

Nu är den mottagna radiosignalen oftast för svag för att våra sinnen direkt ska kunna uppfatta den. Det behövs någon slags förstärkning innan vi kan låta signalen driva t ex en högtalare. R von Lieben och Lee de Forest undersökte därför möjligheten att på något sätt styra strömmen genom en diod för att därmed om möjligt kunna förstärka svaga signaler. 1907 introducerade de Forest sitt elektronrör Audion, en triod där han mellan glödkatoden och platinaanoden



Figur 1.8: Figuren visar signalen som skall sändas (a), och skiftad till att alltid vara positiv (b). Bärvågen (c) moduleras i amplitud (d), sänds och detekteras av mottagaren med en likriktare (e).



Figur 1.9: Audion-röret. Vid  $U_g = 0$  gallerspänning relativt katoden (a) tar sig de flesta elektronerna igenom och fram till anoden. En negativ gallerspänning (b) blockerar vägen mer eller mindre beroende på spänningens storlek, dvs styr anodströmmen.



$$I = \frac{\Delta g}{\Delta t} \quad \text{ström}$$

$$J = \frac{I}{A} \quad \text{strömtäthet}$$

$$\mathcal{E} = \frac{U}{L} \quad \text{elektriskt fält}$$

$$\rho = \frac{R \cdot A}{L} \quad \text{resistivitet}$$

$$\sigma = \frac{1}{\rho} \quad \text{konduktivitet}$$

Figur 1.10: Några elektriska samband.

placerat ett nickelgaller som tredje elektrod. Små ändringar i spänningen  $U_g$  mellan galler och katod,  $\Delta U_g$ , ledde till stora ändringar i anodströmmen, och därmed till en stor motsvarande ändring av spänningsfallet  $\Delta U_R$  över anodkretsens motstånd  $R_a$ . Förstärkningen ges av

förstärkning = 
$$\frac{\Delta U_R}{\Delta U_q}$$
.

En tillräckligt stor negativ spänning  $U_g$  hindrade elektronerna effektivt från att nå anoden. Gallret (eller styret) drog heller ingen ström eller effekt så länge det var negativt (relativt katoden). Genom återkoppling kunde även elektromagnetiska svängningar genereras. När slutligen I Langmuir i USA hade förbättrat evakueringspumpen kunde man börja tillverka reproducerbara radiorör i stor skala.

I detta avsnitt introducerades två centrala komponenter som möjliggjorde radiokommunikation, dioden och trioden. Dessa komponenter var några centimeter långa och ersattes 40 år senare kring mitten av 1900-talet av pn-övergångar och transistorer med dimensioner av storleksordningen  $10^{-6}$  m.

## 1.3 Elektronerna i en metall som en klassisk gas (P Drude, 1900)

När man nu visste att det finns elektroner i en metall, försöker P Drude 1900 tolka ledningsförmågan – både den elektriska och den termiska – som en transport av rörliga elektroner. Dessa fria elektroner betraktar han som en gas enligt klassisk gasteori. Ett av det fundamentala sambanden inom elektriska kretsars teori är

Ohms lag: 
$$U = R \cdot I$$

som empiriskt motiverades 1826 och hade visat sig gälla för metalliska ledare. Med hjälp av sambanden i figur 1.10 kan lagen skrivas som

$$J = \frac{\Delta q}{A \cdot \Delta t} = \sigma \mathscr{E}$$

en relation mellan strömtätheten och elektriskt fält. Konduktiviteten (eller ledningsförmågan)  $\sigma$  införs som en specifik egenskap hos ett material som är oberoende av provets dimensioner. Vid en viss spänning över provet begränsas strömmen av provets resistans R, dvs Ohms lag förutsätter att elektronrörelsen är utsatt för någon slags friktion. Drude tänker sig att elektronerna krockar mot atomerna i metallen. Vid kollisionerna tillintetgörs hastighetsbidraget på grund av det elektriska fältets acceleration, energi överförs till metallens atomer – som t ex en brödrost snabbt låter oss känna. Resultatet blir en liten medelhastighet hos elektronerna i fältets motriktning, strömmen blir konstant.

Om elektronerna leder ström, vilket ju tyder på en viss förmåga att röra sig, borde de även kunna transportera termisk energi. I heta områden ökar den termiska energi de för med sig när de slumpartat rör sig mot kallare områden. Varierar temperaturen längs med ett prov kommer ett flöde av termisk energi

$$J^Q = \frac{\Delta Q}{A \cdot \Delta t} = -\kappa \frac{dT}{dx}$$

där Q är termisk energi ("heat") och  $\kappa$  den termiska ledningsför-  $\kappa$  - termisk ledningsförmåga mågan. Det fanns ett empiriskt samband från 1853,

Wiedemann-Franz lag:  $\frac{\kappa}{\sigma T}\approx 2\times 10^{-8}~{\rm W}\Omega/{\rm K}^2$ 

som Drude lyckades förklara även om han inte kunde bestämma  $\kappa$  eller  $\sigma$  var för sig. Han kunde inte veta med hur många elektroner en atom i metallen bidrar till ledningsförmågan. Hans modell måste vänta på att få sitt svar tills man med kvantmekanikens hjälp kartlagt atomernas elektronstruktur. Han kan heller inte förklara varför vissa material är metaller, dvs ledande, medan andra är isolatorer. Koncentrationen av rörliga elektroner var ju okänd vid den tidpunkten.

Men som vi kommer att diskutera mera detaljerat nedan döljer Drudes modell än allvarligare svårigheter som tydligt pekar på den klassiska fysikens begränsningar. I hans härledning förekommer t ex elektronernas bidrag till den molära värmekapaciteten. Detta bidrag skriver han för  $N_A$  atomer med var sin elektron enligt den klassiska likafördelningslagen (se senare i kapitlet) som

$$C_V^{el} = \frac{3}{2} N_A k$$

där  $N_A$  är Avogadros tal och k Boltzmanns konstant.

Experimentellt visar det sig dock att den totala värmekapaciteten, dvs summan av atomernas och elektronernas bidrag, är

$$C_V \approx 3N_A k$$

oavsett om materialet är en metall eller isolator, vilket måste betyda att  $C_V^{el} \approx 0!$  Elektronerna i en metall förblir förvånansvärt

Värmekapacitet vid konstant volym:

$$C_V = \left. \left( \frac{\partial E}{\partial T} \right) \right|_{V = \text{konst}}$$

Värmekapaciteten är relaterad till en viss substansmängd. Om elektronkoncentrationen n anges i m<sup>-3</sup> får  $C_V$  enheten J/(m<sup>3</sup>·K). opåverkade vid en uppvärmning. De tar inte upp någon nämnvärd energi när temperaturen höjs!

I Thomsons experiment kunde vi beskriva elektronstrålen med klassisk mekanik, i en metall däremot där elektronerna i hög koncentration trängs bland relativt tätpackade atomer verkar en klassisk beskrivning inte längre duga.

## 1.4 Svartkroppsstrålning – M Planck (1900)

Elektronens upptäckt var ett stort steg mot förståelsen av materians uppbyggnad, men hade skett inom ramarna av den klassiska fysiken och orsakade heller ingen omedelbar revolution i fysikens begreppsvärld. En långt mera dramatisk och för många smärtsam brytpunkt i fysiken skulle knytas till året 1900. En liten konstant hdök upp för första gången och ledde till att alla invanda idéer måste ifrågasättas.

Mot slutet av 1800-talet hade man lagt ner mycket möda på att å ena sidan mycket noggrant uppmäta svartkroppsstrålningens spektralfördelning (se figur 1.11) och å andra sidan komma fram till en teoretisk beskrivning. 1894 började Max Planck i Berlin att studera problemet.

Alla kroppar emitterar elektromagnetisk strålning när de värms upp. Vi ser detta med blotta ögat när en eldgaffel lämnats liggande i brasan och börjar lysa rött (700°C).Vi ser också glödtrådars gulaktiga färg i våra lampor (3000°C) och solens nästan vita ljus vid middagstid (5700°C) (när atmosfärens absorption är som svagast). Men vi skulle bli synnerligen förvånade om en kokplatta glömd på maximal effekt började lysa blått. Men detta är just vad den klassiska teorin skulle förutsäga!

Nu är ju ögat ingen helt tillförlitlig detektor av elektromagnetisk strålning. Vi vet ju att vitt ljus är ett synintryck som framkallas av ljus sammansatt av vågor med alla våglängder i det synliga området. Detta visade Newton redan 1666, se figur 1.12. Här skiljer sig ögat från örat som lätt urskiljer flöjtens toner från cellons, dvs korta från långa våglängder. När vi lyssnar till orkestern hör vi inte bara en "ton" som när ögat bara "ser" en färg. Denna observerade "färg" beror på ljusets sammansättning, dvs hur intensiteten är fördelad över de olika spektralfärgerna. När man kring 1900 började använda glödlampor blev även industrin intresserad av ljuset som den upphettade glödtråden utsände.

Gustav Kirchhoff visade 1859 att spektralfördelningen hos den emitterade termiska strålningen blir helt oberoende av kroppens sam-





Svartkroppsstrålningens spektralfördelning vid olika temperaturer. Notera att ögat är känsligast ungefär vid de våglängder där solen lyser starkast.



The production of the spectrum by Newton in 1666

Figur 1.12: Newton delar upp vitt ljus i sina spektralkomponenter.

mansättning och beskaffenhet om kroppen är svart, dvs absorberar all infallande strålning. Spektralfördelningen beror då endast på temperaturen. Ett sätt att realisera en svart kropp är att betrakta ett hålrum med endast en liten öppning mot omvärlden, figur 1.13. All strålning som passerar öppningen absorberas förr eller senare i hålrummet, den strålning som emitteras genom öppningen utsänds då från en svart kropp och beror endast på hålrummets temperatur. Inne i hålrummet kan vi karakterisera strålningen genom att ange strålningens spektrala energifördelning antingen med avseende på våglängder eller frekvenser. Här föredrar vi att studera frekvensberoendet, dvs att låta vinkelfrekvensen vara den oberoende variabeln. I vakuum är ju frekvensen  $\omega = 2\pi f$  relaterad till våglängden  $\lambda$  genom

$$f = \frac{c}{\lambda} \quad \omega = 2\pi \frac{c}{\lambda}$$

Energitätheten  $u(\omega,T)$  definieras av att  $u(\omega,T) \cdot d\omega$  är energin per volymsenhet inom (vinkel)frekvensintervallet ( $\omega, \omega + d\omega$ ). Man kan visa att den genom en öppning med area A utstrålade effekten (energi/tidsenhet) ges av

$$I = A\frac{c}{4} \cdot \int u(\omega, T) d\omega.$$

Det gäller då för Planck att bestämma  $u(\omega,T)$ .

Planck representerar hålrummets väggar med oscillatorer (eller "resonatorer") som växelverkar med strålningen. Han visar först med elektrodynamik att strålningstätheten och det termiska medelvärdet av en oscillators energi  $\langle \varepsilon \rangle$  i termisk jämvikt är relaterade genom sambandet

$$u(\omega,T) = \frac{\omega^2}{\pi^2 c^3} . \langle \varepsilon \rangle.$$

Han är expert på termodynamik och fokuserar sedan på sambandet mellan entropi och energi hos oscillatorerna. Han lyckas först bestämma ett sådant samband (genom en slags djärv interpolation) som resulterar i en spektralfördelning i överensstämmelse med experimenten. Sedan försöker han att motivera sin empiriska formel teoretiskt. Men när han då under räkningens gång vill beräkna en oscillators medelenergi med hjälp av statistisk fysik tvingas han införa små energienheter som måste vara proportionella mot oscillatorns frekvens f,

$$\varepsilon = hf = \hbar\omega \quad \left(\hbar = \frac{h}{2\pi}\right),$$

om hans räkning ska ge den observerade spektralfördelningen. Hans resultat blir *Plancks spektrala energifördelning:* 

$$u(\omega,T) = \frac{\omega^2}{\pi^2 c^3} \cdot \frac{\hbar\omega}{e^{\hbar\omega/kT} - 1}.$$



Figur 1.13: En stråle passerar in i hålrummet, reflekteras och absorberas, men kommer aldrig ut.



Figur 1.14: Spektrala energitätheten  $u(\omega, T)$  plottad mot frekvensen  $\omega$  vid tre olika temperaturer.



Figur 1.15: Några vågor eller svängningsmoder som passar in i ett kubiskt hålrum, dvs har noder på ytorna.

Han kan inte låta proportionalitetskonstanten h gå mot noll vid räkningens slut som man brukade göra i klassisk statistisk mekanik. Då hade Plancks strålningslag blivit

$$u(\omega,T) = \frac{\omega^2}{\pi^2 c^3} \cdot kT$$
 eftersom  $e^x = 1 + x + \dots$ 

Överensstämmelsen med ytterst noggranna experimentella data hade gått förlorad. h måste förbli ändligt.

Planck hade utgått från oscillatorerna i hålrummets väggar som befann sig i termisk jämvikt med strålningen. Fokuserar man direkt på själva strålningen kan man visa att den första faktorn i den spektrala energifördelningen ger antalet elektromagnetiska svängningsmoder per volymsenhet inom intervallet  $(\omega, \omega + d\omega)$  – eller tillståndstätheten som vi kommer att återkomma till. Den andra faktorn ger medelenergin hos en svängningsmod (hos strålningen) med frekvens  $\omega$ , eller också antalet fotoner (kvanta)  $\hbar\omega$  som motsvarar denna medelenergi. Detta antal  $\langle n \rangle$  (i medel) är då

$$\langle n \rangle = \frac{1}{e^{\hbar \omega/kT} - 1}$$

Vi ser att för höga frekvenser  $\hbar\omega\gg kT$ går  $\langle n\rangle\rightarrow 0$ , dvs vi behöver inte befara att en kokplatta som glömts på hög effekt skulle stråla ut Röntgenstrålning. Däremot kommer plattan att emittera en hel del IR-strålning med frekvenser för vilka gäller  $\hbar\omega\ll kT$ , eller vid tillräckligt hög temperatur rentav rött ljus.

Hade vi utgått från klassisk statistisk fysik och använt oss av likafördelningslagen som i termisk jämvikt tilldelar varje svängningsmod en medelenergi kT, så skulle energitätheten blivit (som i gränsen  $h \rightarrow 0$  ovan)

$$u(\omega,T) = \frac{\omega^2}{\pi^2 c^3} \cdot kT$$

Med sin höga frekvenser hade blått ljus dominerat spektralfördelningen och ljuset blivit blåaktigt. UV ljus och Röntgenstrålning hade emitterats kraftigt – lyckligtvis i strid med all erfarenhet. Den totala strålningsenergin i hålrummet skulle blivit oändligt stor:

$$\int_0^\infty u(\omega,T)\cdot d\omega \to \infty,$$

man talade om en ultraviolett katastrof!

Hade detta h någon djupare innebörd? Det tog tio år för Planck att övertyga sig om att hans "energikvanta" var mer än rent formella objekt införda för att matematiskt kunna genomföra analysen. Men

snart ökade mängden experimentella data som bekräftade kvanteffekternas existens, och så småningom vann denna nya kvantfysik, kanske inte alltid genom att övertyga de misstänksamma utan ofta genom att motståndarna med tiden dog ut.

## 1.5 Fotoelektriska effekten

Under åren 1886 till 1891 utförde H Hertz i Karlsruhe en rad avgörande experiment för att generera och detektera de elektromagnetiska vågor som Maxwells teori hade förutsagt. Med en induktionsspole genererade han en hög spänning som lades mellan två tunna metallstavar som var separerade med ett litet gap som i figur 1.16. Den ena staven blev negativt laddad, den andra positivt. Vid tillräckligt hög spänning uppstod en bågurladdning mellan stavarna då den joniserade gasen mellan stavarna blev ledande. Laddningarna på stavarna (elektronerna) kunde nu flyta fram och tillbaka och oscillationen av laddningen alstrade elektromagnetiska vågor. Så småningom dämpades svängningen, spänningen över gapet sjönk och urladdningen avstannade.

Hertz detekterade strålningen med en ledare som var sluten till en ring så när som på ett mycket litet gap. Vågorna inducerade en spänning över detta gap som när den blivit tillräckligt stor genererade en gnista.

Hertz brukade skydda den sekundära (svagare) gnistan med ett hölje mot ströljus för att lättare kunna se den. 1887 observerade Hertz att det starka ljuset från den primära bågurladdningen (sändaren) underlättade tändningen av den sekundära gnistan (mottagaren) när han avlägsnade höljet.

Ljuset frigör elektroner ur metallen och dessa laddningar startar sedan urladdningen lavinartat redan vid ett längre avstånd mellan elektroderna, dvs vid ett lägre elektriskt fält. Han hade upptäckt den fotoelektriska effekten. Något år senare noterade även Hallwachs att en metall blir positivt laddad under belysning.

P Lenard bestämmer år 1900 förhållandet e/m för de emitterade partiklarna och får samma värde som Thomson hade bestämt för elektronerna i katodstrålarna. 1902 belyser han katoden med en bågurladdning mellan kolelektroder där han kan variera ljusintensiteten med upp till en faktor 1000. Utan pålagd spänning mellan anod och katod flyter det redan en ström eftersom de emitterade elektronerna har fått kinetisk energi och kan nå anoden A (se figur 1.17). Med en positiv potential på anoden hjälper det elektriska fältet elektronerna att nå A, strömmen ökar tills den mättas när alla



Figur 1.16: Elektromagnetiska vågor som utgår från sändaren detekteras av mottagaren genom att en gnista uppstår i ringens öppning.

I en *bågurladdning* vid högt tryck leder en stor ström till en så kraftig upphettning att elektroner emitteras utan att detta kräver jonbombardemang. Det uppstår en mycket ljusstark ljusbåge med temperaturer kring 6000 °C. Sådan bågurladdning mellan t ex två kolelektroder har använts som ljuskälla i biografers projektionsapparater och i Xenon-högtryckslampor som tänds av en mycket hög spänning men därefter brinner vid lägre spänning som en självständig urladdning.



Figur 1.17: Schematisk bild av Lenards urladdningsrör. Ljuskällan L belyser katoden K, de emitterade elektronerna kan passera anoden A genom en liten öppning och kommer då in i ett område (streckad cirkel) med ett magnetfält  $\mathbf{B}$  som pekar ut ur figuren. Elektronerna kan registreras av detektorerna C och D.

emitterade elektroner når anoden. Lägger Lenard nu på en negativ potential på anoden byter det elektriska fältet riktning och motverkar elektronernas rörelse mot anoden. Elektronernas potentiella energi ökar på vägen från katoden till anoden, allt färre elektroner har tillräcklig kinetisk energi för att "ta sig uppför backen" och strömmen sjunker. Vid en viss retarderande spänning  $-U_0$  når inte ens de energirikaste elektronerna längre fram till anoden, strömmen försvinner helt. Detta innebär att elektronerna högst har en kinetisk energi

$$E_{kin}^{max} = \frac{m}{2}v_{max}^2 = eU_0$$

när de emitteras från katoden. Ökar Lenard nu ljusets intensitet skulle man förvänta sig att en större energi överförs till elektronerna. Men experimentet visar att elektronernas maximala kinetiska energi förblir precis den samma, spänningen som hindrar dem att nå anoden är fortfarande  $-U_0$ . Att strömmen ökar med ökande intensitet måste då innebära att fler elektroner emitteras.

Elektronernas maximala kinetiska energi verkar öka med frekvensen, men experimenten är osäkra och svåra.

## 1.6 A. Einstein 1905

1905 publicerar Einstein en artikel där de kvanta som Planck införde i sina räkningar blir verkliga på ett ytterst påtagligt sätt. Det infallande ljuset antas består av odelbara energikvanta  $hf = \hbar\omega$ , som 1926 döps till fotoner av GN Lewis. Den enklast tänkbara processen är då att ett helt energikvantum överförs till en enda elektron. Om



Figur 1.18: Ökande ljusintensitet ger en ökande ström. Den spänning  $-U_0$  som hindrar även de energirikaste elektronerna från att nå anoden är däremot oberoende av ljusets intensitet.

frekvensen hos ljuset är tillräckligt hög kan elektronen få tillräckligt hög energi för att kunna lämna metallen, jfr figur 1.19. På väg ut ur metallen måste elektronen uträtta ett arbete, metallens utträdesarbete  $\phi$ , för att övervinna de attraktiva krafterna som försöker hålla den kvar. Einstein skriver följande samband som bestämmer elektronens maximala kinetiska energi när den befinner sig utanför metallen

$$E_{max}^{kin} = \hbar\omega - \phi$$

Enligt figur 1.19, som visar en "modern" schematisk bild av elektronerna i en metall, är alla elektroniska tillstånd fyllda upp till den så kallade Fermienergin  $E_F$ . (Denna storhet kommer vi att studera närmare senare.) Om nu fotonen exciterar elektronen med högst energi i metallen (vid  $E_F$ ) behöver elektronen uträtta ett arbete, dvs avge energin  $\phi$  för att ta sig ut ur metallen och resten behåller den som kinetisk energi. Exciteras en lägre liggande elektron (lägre i energi!) så blir elektronens kinetiska energi utanför metallen mindre. Ljusets intensitet motsvarar transporterad energi per ytoch tidsenhet, vilket motsvarar antalet fotoner per yt- och tidsenhet. Fotonens energi bestämmer elektronernas maximala kinetiska energi, men denna är oberoende av intensiteten. Antalet exciterade elektroner däremot bestäms av intensiteten.

Einsteins relation förklarar de experimentella observationerna, men förutsäger också att den maximala kinetiska energin

$$E_{max}^{kin} = eU_0 = \hbar\omega - \phi$$

ökar *linjärt* med ljusets frekvens. Inga detaljerade data finns dock att jämföra med vid denna tidpunkt. Först år 1916 lyckades Millikan



Figur 1.19: Elektronen som har högst energi inne i metallen har efter excitationen även högst kinetisk energi utanför.



Figur 1.20: Millikans data för litium. Den retarderande spänningen  $U_0$  vid vilken strömmen försvinner är plottad mot det infallande ljusets frekvens. Den räta linjens lutning bestämmer h/e.

i Chicago bestämma  $U_0 - \omega$  sambandet för natrium och litium och fann en perfekt rät linje ur vars lutning h/e han kunde bestämma Plancks konstant eftersom han själv 1910 noggrant hade bestämt elementarladdningen e. Millikan kommer fram till en

"astonishing situation that these [experimental] facts were correctly and exactly predicted nine years ago by a form of quantum theory which has now been pretty generally abandoned."

Man blir något förvånad!

## 1.7 Röntgenstrålning upptäcks (WC Röntgen 1895)

Flera forskare hade studerat katodstrålar under nästan 40 år under senare hälften av 1800-talet. Framsteg hade möjliggjorts av stadigt förbättrade evakueringspumpar och tillgång till allt högre spänningar. I en sådan lång utvecklingsprocess är det inte alltid lätt att identifiera en person som "upptäckare" eller att avgöra vems insats var mest värdefull.

När en upptäckt är rent tillfällig, när den som en blixt ur klar himmel lämnar upptäckaren lika förvånad som omvärlden, är det lätt att identifiera en händelse. Man kan ange personen, platsen och även tidpunkten. Så var fallet när Wilhelm Conrad Röntgen den 8 november 1895 upptäckte sina, som han kallade dem, *x*-strålar.



Figur 1.21: Ett Röntgenrör där en upphettad katod redan vid lägre spänningar emitterar elektroner.

Även Röntgen studerade katodstrålar och denna dag arbetade han ensam i sitt laboratorium. I elektronröret var gastrycket kring 0,1 Pa och han kunde generera spänningar upp till 100 000 V med hjälp av en induktionsspole. Hela apparaturen var omsluten av svart papp.

Av en slump observerar Röntgen fluorescens på en skärm belagd med barium platina-cyanid som han brukar använda som detektor och som råkar befinna sig i elektronrörets närhet. Den svarta pappen kring röret släpper inte igenom något ljus, ändå konstaterar Röntgen att strålningen kommer från röret. Han upptäcker snart att strålarna förmår tränga igenom inte bara papp utan de flesta material. Den första bilden han tar med strålarnas hjälp är av makans hand, både ringen och handens ben är tydligt synliga. Upptäckten blir en sensation. Medicinare är snabba att utnyttja strålningen för att undersöka kroppens inre. T. o. m. skoaffärer skaffar en Röntgenapparat för att kontrollera skons passform.

Röntgen fick så småningom tillträde till kejsar Wilhelm II för att demonstrera sitt experiment. Efter audiensen avlägsnade han sig som etiketten krävde gående baklänges. 1901 belönades han med fysikens första Nobelpris.

Man spekulerade länge om Röntgenstrålar snarare uppvisade partikeleller vågegenskaper. Man observerade snart att Röntgenstrålar kunde jonisera atomer. Om strålningen nu är en vågrörelse, hur är det möjligt att den energi som vågen överför till en elektron per tidsenhet inte ändras även om elektronens avstånd till strålningskällan ökas? Man förväntar sig ju att vågens energi skulle spridas ut över ett allt större område, dvs att vågens intensitet skulle avta med avståndet från källan. Man kunde tänka sig att den mera utspridda vågen behövde inverka på en mera avlägsen elektron en längre tid för att ge upphov till jonisation, men man fann aldrig någon fördröjning av excitationen. Mycket talade för att se strålningen som

#### Partikel-våg dualitet

It is like a struggle between a tiger and a shark, each is supreme in his own element, but helpless in that of the other. JJ Thomson, 1925



Figur 1.22: Hur är det möjligt att den snabbaste elektronen som emitteras av atomen under inverkan av Röntgenstrålningen har (väsentligen) samma fart som den snabbaste elektronen i Röntgenröret som gav upphov till strålningen? (A Dauvillier)

en ström av korpuskler eller energiknippen. William Bragg, exempelvis, lutade alltmer åt en dualistisk tolkning av experimenten där våg- och korpuskelaspekter kompletterade varandra. Röntgenstrålningen hade lyft fram en fråga som även borde gälla synligt ljus, vars vågegenskaper sedan länge var experimentellt bekräftade och allmänt accepterade. Kunde även synligt ljus ibland bete sig som partiklar? Nu var frågan inte längre korpuskler *eller* våg, utan hur kan något vara både korpuskel *och* våg, samtidigt!

För synligt ljus kände man väl till de interferensmönster som kan uppstå om en stråle träffar ett gitter. HA Rowland i USA hade vidareutvecklat tekniken att framställa gitter och klarade av att "rita" 400-800 linjer per millimeter. Eftersom Röntgenstrålar inte spreds av vanliga gitter förmodade man att strålningen hade betydligt kortare våglängd. 1912 föreslog Max Laue vid A Sommerfelds institut i München att man skulle kunna använda kristaller som gitter för Röntgenstrålar. Då kan våglängder av samma storleksordning som atomavstånden i kristaller bestämmas. W Friedrich och P Knipping, även de vid samma institution i München, fotograferade de första interferensmönstren. Röntgenstrålar verkade faktiskt vara vågor med liknande egenskaper som vanligt ljus. Även WH Bragg och hans son WL Bragg studerade interferensmönster men deras huvudsakliga intresse skiftade snart från att undersöka strålningens natur till att kartlägga uppbyggnad och struktur hos kristallina material. Vi gör därför en liten utvikning om kristaller.

On Mondays, Wednesdays and Fridays, we teach the wave theory and on Tuesdays, Thursdays and Saturdays the particle theory. W Bragg, 1921

### 1.7. RÖNTGENSTRÅLNING UPPTÄCKS (WC RÖNTGEN 1895)27

#### Kristaller

Kristaller kännetecknas av perfekt periodicitet – "longrange order". Atomernas lägen är regelbundna och denna ordning sträcker sig över många atomavstånd. Många olika material är kristallina. Ibland syns detta direkt när vi t ex hittar en bergkristall i naturen, periodiciteten sträcker sig genom hela systemet, vi har en så kallad *enkristall*. Ibland består materialet av många ytterst små, slumpartat orienterade sammanfogade kristalliter som i en metall. Kristallina materials egenskaper uppvisar stora variationer.

#### Metallerär

- goda reflektorer av synligt ljus ogenomskinliga,
- $\bullet\,$ goda elektriska ledare, och
- mekaniskt formbara.

#### Jonkristaller är

- transparenta,
- $\bullet\,$ elektriska isolatorer, och
- brukar ha bandgap omkring 8 eV. (Begreppet bandgap förklaras i kapitlet Bandstruktur!)

I en jonkristall som t ex NaCl är jonerna ordnade som i figur 1.24. Väsentligen sfäriska positiva Na<sup>+</sup> (svarta) och negativa Cl<sup>-</sup>-joner hålls samman av elektrostatiska (Coulomb) krafter.

Bland *kovalent bundna kristaller* återfinns viktiga halvledare, t.ex. kisel (Si), germanium (Ge) och många andra som galliumarsenid (GaAs), InP och andra kombinationer av atomslag ur grupp 3 och 5 i periodiska systemet.

Diamant är prototypen bland kovalent bundna kristaller. Kisel, germanium, GaAs och många andra halvledare kristalliserar i diamantstruktur. Vi behöver här inte närmare gå in på beskrivningen av kristallstrukturer eftersom vi framöver oftast använder oss av enkla endimensionella modeller.

I kisel och germanium är atomerna bundna av kovalenta krafter, strukturen är samma som i diamant. Varje atom är här bunden till 4 närmaste grannar. Dessa tre ämnen är mycket olika, vilket vi kommer att beröra mer senare



Figur 1.23: Schematisk kristallstruktur hos en metall. Notera periodiciteten



Figur 1.24: NaCl kristallstruktur.



Figur 1.25: Kristallstrukturen hos diamant, kisel och germanium.

#### Diamant är

- $\bullet$  transparent,
- hårt, och
- en elektrisk isolator (har stort bandgap)

#### *Kisel* är

- ogenomskinligt,
- $\bullet\,$ en halvledare

Braggs lag blev ett välkänt och behändigt samband för att relatera strålningens våglängd till avståndet mellan kristallens atomplan i analysen av Röntgenstrålningens interferensmönster.

Braggs lag (1913):  $2D\sin\theta = n\lambda$ 

För att förstå Braggs lag utgår vi från reflektionen i ett atomplan som i figur 1.26. Vi anger strålens riktning med hjälp av den så kallade glansvinkeln  $\theta$  (vinkeln mellan strålen och planet). Vägskillnaden hos de två vågor som uppstår pga spridning i två närliggande atomer blir noll, dvs vågorna interfererar konstruktivt, i en riktning som har samma glansvinkel som den inkommande strålen. Strålen reflekteras såsom i en spegel. Men Röntgenstrålningen tränger djupt in i kristallen och endast en bråkdel av strålningen sprids i varje enskilt plan. Vägskillnaden mellan strålar som sprids i närliggande plan (se figur 1.27) måste vara ett helt antal våglängder, dvs  $2D \sin \theta = n\lambda$  för konstruktiv interferens. Vi ser att detta kräver att våglängden  $\lambda$  måste vara mindre än 2D, där ett typiskt plan- eller atomavstånd ligger kring D = 0,2 nm.

Nu hade man ett redskap för att analysera Röntgenstrålningen om man kände kristallens struktur, eller så kunde man kartlägga strukturen om man kände strålningens våglängd. Med sin kristallspektrometer (se figur 1.28) lyckades far och son Bragg detektera de första skarpa linjerna i ett Röntgenspektrum och även bestämma deras våglängd med hjälp av rimliga antaganden om kristallens uppbyggnad. Det var inte helt enkelt att mäta strålningens spektralfördelning eftersom denna innehöll smala skarpa toppar man lätt kunde missa i ett svep över våglängder. Ett typiskt spektrum kunde se ut som i figur 1.29. Spektrumet består av en bred strukturlös bakgrund och ett antal skarpa toppar. Det visade sig snart att dessa toppar – deras våglängd och relativa intensitet – var karakteristiska för det valda anodmaterialet i Röntgenröret. Bakgrundstrålningen



Figur 1.26: Konstruktiv interferens vid reflektion.



Figur 1.27: Konstruktiv interferens kräver att vägskillnaden  $2D\sin\theta$  är ett helt antal våglängder (Braggs lag).

## 1.7. RÖNTGENSTRÅLNING UPPTÄCKS (WC RÖNTGEN 1895)29



Figur 1.28: Braggs kristallspektrometer. Röntgenstrålar vars våglängd uppfyller Braggs lag reflekteras in i en gasfylld kammare där man genom att mäta gasens jonisation kan bestämma strålningens intensitet.



Figur 1.29: Ett typiskt emissionsspektrum uppvisar kontinuerlig bromsstrålning och skarpa, för anodmaterialet karakteristiska toppar.



Figur 1.30: När elektroner skjuts in i metallen växelverkar de med materians laddningar, dvs accelereras. En accelererad laddning strålar ut elektromagnetiska vågor och tappar därmed energi, dvs bromsas.



Figur 1.31: En inkommande elektron med kinetisk energi  $E_{kin}$ avger (ungefär) hela sin energi till en foton  $\hbar\omega$ .

kunde tolkas som den strålning en elektron emitterar vid sin kraftiga inbromsning när den tränger in i anoden som i figur 1.30. Märkligt var dock att spektrumet hos denna så kallade **bromsstrålning** (bremsstrahlung) så abrupt försvann för våglängder kortare än en viss gränsvåglängd  $\lambda_{min}$ . W Duane och F Hunt vid Harvard undersökte 1915 hur  $\lambda_{min}$  var relaterad till spänningen över Röntgenröret. Sambandet de fann var

$$eU = \frac{hc}{\lambda_{min}}$$

När elektronen i Röntgenröret har passerat spänningsfallet U har den fått en kinetisk energi  $E_{kin} = eU$ . I denna omvända fotoelektriska effekt kan den nu som mest överföra hela sin energi till ett strålningskvantum

$$eU = \hbar\omega = hf = \frac{hc}{\lambda}$$

och relationen vi kommer ihåg från fotoeffekten

$$E_{kin} = \hbar\omega - \phi = hf - \phi$$

gäller. Jämfört med energin eU som låg på tiotusentals elektronvolt är utträdesarbetet försumbart. De karakteristiska linjerna kunde så småningom relateras till elektroniska övergångar djupt inne i atomen som figur 1.34 visar schematiskt.

### **1.8** Atomernas struktur

E Rutherford kom 1895 till Cavendish laboratoriet i Cambridge från Nya Zeeland för att studera under JJ Thomson. Där undersökte han den av AH Becquerel 1896 upptäckta radioaktiviteten och identifierade de tre strålningstyperna  $\alpha, \beta$  och  $\gamma$ . 1908 fick han Nobelpriset i kemi för sin upptäkt att radioaktivitet innebar en transformation av ett element till ett annat under emission av strålning.  $\alpha$ -strålningen identifierades som dubbelt positivt laddade heliumatomer (dvs Hekärnor). När han tillsammans med H Geiger och E Marsden lät strålar av  $\alpha$ -partiklar träffa en guldfolie kunde spridningsmönstret hos  $\alpha$ - partiklarna endast förklaras av att guldatomerna innehöll en ytterst liten positivt laddad, massiv "kärna" vars diameter var ca 100 000 gånger mindre än atomens. Detta var genombrottet på vägen att förstå atomers struktur. Märkligt var att Rutherford gjorde sin kanske mest betydande upptäckt efter att redan ha fått ett Nobelpris tidigare!

Men var fanns då de lätta elektronerna? Om elektronerna inte rörde sig skulle attraktionen från kärnan obevekligen dra dem inåt. Rörde de sig kring kärnan så skulle de hela tiden accelereras, deras hastighet skulle ju ständigt ändras. Och en accelererad laddning skulle ju enligt Maxwell alltid utstråla elektromagnetiska vågor. Därmed skulle elektronerna snabbt förlora sin kinetiska energi och falla in mot kärnan. Hur man en vred och vände på atomen kunde man inte komma på något stabilt arrangemang för elektronerna.

1911 kommer Niels Bohr till Cambridge. Det tänkta samarbetet med Thomson ser inte ut att bli lovande men där möter han Rutherford och flyttar till hans institution i Manchester och påbörjar arbetet med atomen. 1913 publicerar han sin första teori om atomens struktur. Hans modell omfattar bl a två postulat. i) Elektronerna rör sig på stationära banor utan att stråla, deras rörelsemängdsmoment är kvantiserat och kan endast anta heltalsmultipler av  $\hbar$ . ii) Övergångar mellan dessa stationära tillstånd ger upphov till strålning vars frekvens bestäms av relationen

$$hf = h\frac{c}{\lambda} = E_i - E_f$$

där  $E_i$  och  $E_f$  är energierna hos begynnelse- resp. sluttillstånd och hf fotonenergin. En viktig ledtråd hade Bohr funnit i en formel från 1884 som en schweizisk matematiklärare J Balmer rent empiriskt hade ställt upp för våglängderna hos väteatomens spektrallinjer.

$$\lambda = b \frac{m^2}{m^2 - n^2} \qquad (n = 2, m > n)$$

Fyra av dessa linjer hade noggrant uppmätts av Anders Ångström, m = 3, 4, 5 och 6. När Ångström senare även bestämde en femte spektrallinjes våglängd stämde även denna med Balmers formel. 1888 hade Janne Rydberg funnit en mera allmän formel som för väte skrevs

$$\frac{1}{\lambda} = R\left(\frac{1}{n^2} - \frac{1}{m^2}\right) \qquad (m > n)$$

och som tydligare relaterade vågtalet  $1/\lambda$  (eller fotonenergin) till en differens av två termer. Bohrs modell ledde till följande ekvation

$$\frac{1}{\lambda} = \frac{R}{hc} Z^2 \left( \frac{1}{n^2} - \frac{1}{m^2} \right) \qquad (m > n)$$

där

$$R = \frac{me^4}{32\pi^2\varepsilon_0^2\hbar^2}$$

är Rydbergs konstant och Z kärnans atomnummer eller antalet positiva kärnladdningar.



Figur 1.32: Elektroner med negativ energi är bundna i atomen, med positiv energi kan de lämna atomen. Antalet elektroner i de bundna tillstånden, de så kallade skalen, är begränsat. Observera den brutna energiskalan!

Bohrs modell mottogs med skepsis, man klamrade sig fast vid klassiska fysikens invändningar. Men dessa avfärdades bl a av J Jeans som framhöll att allt modellen krävde för sitt berättigande var framgång. Och framgången kom och med den Nobelpriset 1922. Men det skulle dröja till 1926 innan E Schrödinger löste sin ekvation för väteatomen och den gamla kvantteorin avlöstes av den nya.

Vi behöver här inte följa den vidare utvecklingen av atomers fysik men bör känna till några grundläggande fakta om elektronstrukturen i atomer. Elektronerna rör sig i den attraktiva potentialen från den positiva kärnan, figur 1.32. Vi kanske vet sedan tidigare att elektronerna sitter i skal, där antalet tillstånd i varje skal är begränsat. Skalen karakteriseras i första hand av ett huvudkvanttal

 $n = 1, 2, 3 \dots$ 

Inom varje skal med huvudkvanttal n finns ett eller flera underskal karakteriserade av ett bankvanttal l, som antar värdena  $0,1,\ldots,n-1$  och betecknas med bokstäverna s, p, d, f, g, ... Ett underskal betecknas  $nl^x$  där x är antalet elektroner.

I de atomer vi betraktar är underskalen ordnade i energi enligt 1s, 2s, 2p, 3s, 3p, 3d, 4s, 4p, ... (Skalen 3d och 4s har nästan samma

п	Underskal	Antal tillstånd
1	1s	2
2	2s	2
	2р	6
3	3s	2
	3р	6
	3d	10
4	4s	2
	4p	6



Figur 1.34: En elektron med hög energi (a) träffar en atom i metallen. Den slår ut en elektron ur ett inre skal, t ex en 1s-elektron. Det djupt liggande tomma tillståndet fylls därefter av en elektron ur ett högre liggande skal (b), energin utsänds som en foton.

energi och kan ibland byta plats.)

Elektroniska övergångar i Na-atomen (se figur 1.33) resulterar t ex i natriumlampors karakteristiska gula färg.

Viktigt att notera här är att elektronerna kan delas in i två grupper beroende på deras bindningsenergi. De yttersta, relativt svagt bundna elektronerna i atomer kallas valenselektroner. Dessa sköter t ex bindningen mellan atomerna och bestämmer även materialets elektriska och optiska egenskaper. De inre, hårt bundna elektronerna förblir i dessa sammanhang passiva och är endast undantagsvis av intresse när det gäller fasta ämnens egenskaper.

Bohrs formel gällde väte eller väteliknande joner, dvs joniserade atomer med endast en kvarvarande elektron. Det blev svårt att försöka tillämpa teorin på atomer med flera elektroner och deras linjespektra visade sig vara mera komplicerade och oregelbundna. Men krafterna på de innersta elektronerna borde domineras av Coulomb attraktionen till kärnan, bindningskrafterna mellan atomer skulle t ex inte kunna rubba de inre elektronerna nämnvärt. Efter diskussioner med Bohr genomförde H Moseley i Manchester 1913 sina systematiska studier av karakteristiska Röntgen-linjespektra.

När elektroner med hög energi träffar anoden i ett Röntgen- rör kan en elektron ur det innersta skalet (n = 1) i en atom slås ut. Luckan kan sedan fyllas genom att en elektron med högre energi i skal mtrillar ner och avger energi i form av en foton:

 $hf = E(m) - E(1) \quad (m > 1).$ 

Denna serie av linjer kallas K-serien. Skapas luckan i skalet med n = 2 leder övergångarna till L-serien. Moseley fann att frekvenserna hos



Figur 1.33: Atomära övergångar i natrium. Övergången från 3p till 3s ger natrium atomen den karakteristiska gula färgen i emission.



Figur 1.35: Övergångar till n = 1skalet bildar K-serien, övergångar till n = 2 skalet L-serien.

Röntgen-ljuset som emitterades följde sambandet

$$f \sim (Z - \sigma)^2$$

där Z är atomnumret och  $\sigma$  är en konstant mellan 1 och 2 för K-serien och ungefär 8 för L-serien. Enligt Bohrs modell låg det nära till hands att identifiera Z med kärnans laddning och  $Z - \sigma$ med den effektiva laddningen som gav upphov till Coulomb fältet som elektronerna rörde sig i. Detta fält härrörde ju inte enbart från kärnan utan även från andra elektroner nära kärnan – när det gällde t ex K-serien fanns ju den andra 1s-elektronen kvar, men antalet elektroner i enstaka skal kände man inte till vid denna tid. Moseley formulerade slutligen sin lag

$$f = \frac{R}{\hbar} (Z - \sigma)^2 \left(\frac{1}{n^2} - \frac{1}{m^2}\right)$$

Frekvensen uppvisar därmed en tydlig trend för alla atomer i det periodiska systemet. Moseley kunde då enkelt identifiera luckor i det periodiska systemet och även på några timmar analysera prover bestående av sällsynta jordartsmetaller som kemisterna hade brottats med i flera år. Samtidigt bekräftade hans resultat Bohrs modell även i mera komplicerade atomer bara man begränsade sig till situationer där den rimligtvis kunde förväntas gälla.

Moseleys experiment genomfördes inom loppet av några ytterst intensiva månader. Han stupade 1915 på Dardanellerna innan han kunnat tilldelas ett annars säkert Nobelpris.

## 1.9 Värmekapacitet

Vid 1800-talets början gick meningarna fortfarande isär om värmens natur. Var det en viktlös gas som kunde genomtränga alla kroppar eller skulle man se det som en osynlig rörelse hos materians molekyler? Inte heller begreppet "molekyl" hade man någon klarare uppfattning om. Allteftersom ingenjörerna borrade fler kanonrör och konstruerade fler ångmaskiner blev denna fråga ytterst aktuell. Kalorimetern (jfr figur 1.36) och termometern hade redan utvecklats under 1700-talet och man studerade nu systematiskt värmekapaciteten hos olika former av materia för att försöka avslöja värmens egenskaper.

Värmekapaciteten är definierad som förhållandet mellan tillförd värme  $\Delta Q$  och ändringen i temperatur  $\Delta T$ 

$$C = \frac{\Delta Q}{\Delta T}$$

Värmemängden är proportionell mot substansmängden och C anges då t ex med enhet  $J/(kg \cdot K)$  eller  $J/(mol \cdot K)$ . Enligt termodynamikens första lag gäller

$$\Delta E = \Delta Q + \Delta W$$

där  $\Delta E$  är ändringen av systemets inre energi och  $\Delta W$  arbetet som uträttas på systemet. Betraktar vi endast volymsarbete så är

$$\Delta W = -p\Delta V$$

dvs vid en expansion  $\Delta V > 0$  uträttar systemet arbete på omgivningen. För värmekapaciteten följer

$$C = \frac{\Delta Q}{\Delta T} = \frac{\Delta E}{\Delta T} + p \frac{\Delta V}{\Delta T}$$

Mätningen utförs normalt vid konstant tryck – vilket då bestämmer  $C_p$  – men i teorin föredrar man att hålla volymen konstant, dvs att studera  $C_V$ . Vi bortser här från den relativt lilla skillnaden mellan  $C_p$  och  $C_V$  i fasta material.

$$C_V = \left. \frac{\Delta Q}{\Delta T} \right|_V = \left. \frac{\Delta E}{\Delta T} \right|_V$$

1819 rapporterar PL Dulong och AT Petit att värmekapaciteten för en mol av ett (monoatomärt) fast ämne alltid har ungefär samma värde

### Dulong-Petits lag: $C = 3N_A k$ (3k per atom)

oavsett vilket fast material man studerar. $N_A$  är Avogadros tal som ger antalet atomer i en mol. Denna empiriska lag lyckas Boltzmann förklara 1876 med hjälp av likafördelningslagen (Equipartition theorem). Vi ska titta lite närmare på hur man kan gå tillväga.

I kristallina material är atomerna arrangerade i en perfekt periodisk struktur. Man kan tänka sig bindningarna som fjädrar mellan atomerna, som i figur 1.37. Vid uppvärmning ökar atomernas svängningar kring sina jämviktslägen. (Även ljud utbreder sig som vågor av atomära svängningar genom kristallen.) Vid temperaturer kring smältpunkten blir atomernas svängningar så våldsamma att de kan lämna sina jämviktspositioner för gott.

Låt oss nu betrakta en enkel modell för atomära vibrationer. Vi antar att alla atomer kan svänga i tre oberoende svängningsriktningar längs med de tre koordinataxlarna. Vi lägger origo i atomens jämviktsläge och kan skriva atomens energi som en summa av kinetisk och potentiell energi

$$\varepsilon = \frac{p_x^2}{2M} + \frac{p_y^2}{2M} + \frac{p_z^2}{2M} + V(x,y,z)$$



S



Figur 1.36: Lavoisier

introducerade iskalorimetern, figuren visar Bunsens utformning. Ett kärl a är omgivet av en sluten behållare b som innehåller vatten flytande på kvicksilver (svart) som genom ett kapillärrör leds till en skala S. Genom kylning fryses vattnet som omger kärl a inom det streckade området. En kropp med massa m, värmekapacitet Coch temperatur T placeras därefter i kärl a och kyls nu till 0 °C av den omgivande behållaren och avger därvid energin  $CmT = m_V \cdot L$  som smälter en ismängd med massan  $m_V$  och smältvärme L = 333 kJ/kg. Vid smältning minskar vattnets volym, volymändringen syns på skala S som är kalibrerad att ange  $m_V$ . Mätningen bestämmer då värmekapaciteten C som här har enheten  $kJ/(kg \cdot K)$ .

med till exempel rörelsemängdens komposant i x-led

$$p_x = Mv_x = M\frac{dx}{dt}$$

och där den harmoniska potentialen

$$V(x,y,z) = \frac{K}{2}(x^2 + y^2 + z^2)$$

har sitt minimum i atomens jämviktsläge.

Vi delar upp rörelsen i tre endimensionella svängningar, och får i x-led med kraften

$$F_x = -\frac{\partial V}{\partial x} = -Kx$$

rörelseekvationen

$$M\frac{\partial^2 x}{dt^2} = -Kx$$

Ansatsen för en harmonisk svängning

$$x(t) \sim e^{-i\omega t}$$

ger oss svängningens (vinkel-)frekvens

$$\omega = \sqrt{\frac{K}{M}}.$$

Energin hos den harmoniska oscillatorn kan nu tecknas

$$\varepsilon = \varepsilon_{kin} + \varepsilon_{pot} = \frac{p_x^2}{2M} + \frac{M}{2}\omega^2 x^2$$

För att bestämma energins termiska medelvärd<br/>e $\langle \varepsilon \rangle$ utnyttjar vi Boltzmannfaktorn

$$e^{-\varepsilon/kT} \sim P(\varepsilon)$$

som är proportionell mot sannolikheten  $P(\varepsilon)$  att ett system i termisk jämvikt med sin omgivning med temperaturen T har energin  $\varepsilon$ . Medelvärdet över alla tillstånd får vi inom klassisk fysik genom att integrera över systemets så kallade fasrum, dvs över alla tillgängliga lägeskoordinater och rörelsemängder

$$\langle \varepsilon 
angle = rac{\int \int \varepsilon e^{-\varepsilon/kT} dp_x dx}{\int \int e^{-\varepsilon/kT} dp_x dx}.$$

Nämnaren normerar Boltzmannfaktorn till att vara *lika med* sannolikheten

$$\frac{e^{-\varepsilon/kT}}{\int \int e^{-\varepsilon/kT} dp_x dx} = P(\varepsilon)$$



Figur 1.37: Bindningarna mellan atomerna kan representeras av fjädrar. Om modellen är vettig beror på vilka egenskaper hos kristallen man vill studera.
och det termiska medelvärdet får vi genom att multiplicera energin  $\varepsilon$  med sannolikheten  $P(\varepsilon)$  och summera (integrera) över alla tillstånd,

$$\langle \varepsilon \rangle = \int \int \varepsilon \cdot P(\varepsilon) dp_x dx.$$

Med

$$\beta = \frac{1}{kT}$$

kan vi skriva integralen på formen

$$\langle \varepsilon \rangle = -\frac{\partial}{\partial \beta} \ln \int \int e^{-\left(\frac{p_x^2}{2M} + \frac{M}{2}\omega^2 x^2\right)\beta} dp_x dx$$

Med substitutionen

$$\hat{p}_x = p_x \sqrt{\beta}$$
 och  $\hat{x} = x \sqrt{\beta}$ 

erhåller vi

$$\langle \varepsilon \rangle = \frac{\partial}{\partial \beta} \ln \frac{1}{\beta} \int \int e^{-\left(\frac{\hat{p}_x^2}{2M} + \frac{M}{2}\omega^2 \hat{x}^2\right)} d\hat{p}_x d\hat{x}$$

och slutligen

$$\langle \varepsilon \rangle = -\frac{\partial}{\partial \beta} \ln \frac{1}{\beta} = \frac{\partial}{\partial \beta} \ln \beta = kT$$

eftersom integralen inte längre beror av <br/>  $\beta$  och därmed

$$\frac{\partial}{\partial\beta}\ln\left(\int\int e^{-\left(\frac{\hat{p}_x^2}{2M}+\frac{M}{2}\omega^2\hat{x}^2\right)}d\hat{p}_xd\hat{x}\right) = 0.$$

Detta resultat kan generaliseras till

Likafördelningslagen eller ekvipartitionsteoremet, enligt vilken varje kvadratisk term (kvadratisk i koordinater eller rörelsemängdskomponenter) i uttrycket för energin bidrar till energins termiska medelvärde med kT/2.

En mol innehåller  $N_A$  atomer och varje atom motsvarar 3 endimensionella oscillatorer. Därmed blir kristallens energi (för en mol)

$$E(T) = 3N_A kT$$

 $\operatorname{och}$ 

$$C_V = 3N_A k,$$

Vad hade resultatet blivit om partikeln varit fri, dvs om V(x) = 0?

$$\operatorname{Svar}:kT/2$$

 ${\cal N}_A$  - Avogadros tal

som Dulong-Petits lag angav.

Även om denna empiriska lag stämde väl med observationerna hade man redan 1840 funnit undantag, framför allt bland de lättaste grundämnena som beryllium, bor och kol (diamant). Dessutom blev det allt tydligare att värmekapaciteterna uppvisar ett märkbart temperaturberoende.

Länge hade man varit beroende av kalla vintrar för att mäta värmekapaciteten vid lägre temperaturer, men mot slutet av seklet, 1895, lyckades Carl von Linde genom att utnyttja Joule-Thomsoneffekten att överföra luft till flytande fas, ett ofta använt kylmedel med kokpunkt kring 80 K. Nästan samtidigt kunde J Dewar bilda flytande väte med kokpunkt vid 20 K under tryck. Därmed hade så låga temperaturer blivit tillgängliga att fasta ämnen kunde undersökas från nästan 0 K till smältpunkten. (James Dewar hade förresten samarbetat med Alfred Nobel kring sprängämnet ballistit, men senare marknadsfört ett sprängämne i England under namnet Cordite med ett endast något modifierat recept. Den följande rättsprocessen förlorade Nobel då hans patentansökan var otydligt skriven.) Redan 1875 publicerade HF Weber data för bland annat diamant, som visas i figur 1.39 och som Einstein jämför sina resultat med i sitt arbete om fasta kroppars värmekapacitet 1907.

Einstein utgår ifrån att energin hos en oscillator är kvantiserad, dvs att dess energinivåer ges av

$$\varepsilon_n = \hbar \omega \left( n + \frac{1}{2} \right)$$

i motsats till klassisk mekanik där energin är en kontinuerlig storhet. Han använder här för mekaniska vågor samma kvantisering som för elektromagnetiska vågor i fotoelektriska effekten. Vi inkluderar här även nollpunktsenergin  $\hbar\omega/2$  för att överensstämma med dagens formalism, även om denna term introducerades av Planck först i hans andra version av teorin 1912. Einstein beräknar sedan det termiska medelvärdet av en oscillators energi genom att vikta tillståndens energi med Boltzmannfaktorn och summera över alla tillstånd n. Skillnaden till det klassiska fallet ovan är att energierna



Figur 1.38: Lindes apparatur för att omvandla luft till flytande fas.

nu är diskreta och integralen övergår i en summa.

$$\begin{aligned} \langle \varepsilon \rangle &= \frac{\sum_{n} \varepsilon_{n} \cdot e^{-\varepsilon_{n}/kT}}{\sum_{n} e^{-\varepsilon_{n}/kT}} = \frac{\sum_{n} \left(n + \frac{1}{2}\right) \hbar \omega \cdot e^{-\left(n + \frac{1}{2}\right) \hbar \omega/kT}}{\sum_{n} e^{-\left(n + \frac{1}{2}\right) \hbar \omega/kT}} \\ &= -\frac{\partial}{\partial \beta} \ln \left( \underbrace{\sum_{n} e^{-n\hbar \omega \beta} e^{-\frac{1}{2} \hbar \omega \beta}}_{\frac{1}{1 - e^{-\hbar \omega \beta}}} \right) \end{aligned}$$

där vi utnyttjat den geometriska serien

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + \dots$$

Vi får därmed

$$\langle \varepsilon \rangle = \frac{\partial}{\partial \beta} \left\{ \ln(1 - e^{-\hbar\omega\beta}) - \underbrace{\ln e^{-\frac{1}{2}\hbar\omega\beta}}_{-\frac{1}{2}\hbar\omega\beta} \right\} = \frac{\hbar\omega e^{-\hbar\omega\beta}}{1 - e^{-\hbar\omega\beta}} + \frac{1}{2}\hbar\omega$$

och slutligen

$$\langle \varepsilon \rangle = \frac{\hbar\omega}{e^{\hbar\omega\beta} - 1} + \frac{1}{2}\hbar\omega = \frac{\hbar\omega}{e^{\hbar\omega/kT} - 1} + \frac{1}{2}\hbar\omega$$

Detta resultat hade också Planck erhållit för sina oscillatorer i hålrummets väggar men utan att lika klart skilja på oscillatorns dynamik (energinivåerna) och den statistiska behandlingen (Boltzmannfaktorn). Så länge  $kT \ll \hbar \omega$  gäller kommer den första termen att vara mycket liten, oscillatorn förblir i sitt grundtillstånd (med nollpunktsenergi  $\hbar \omega/2$ ). Dess energi ändrar sig ytterst lite vid en uppvärmning, dess bidrag till värmekapaciteten blir då försumbart. Oscillatorn kommer "i gång" först när  $kT = \hbar \omega$ .

Nu kan ju likafördelningslagen uppenbarligen inte vara helt fel. Frågan är bara under vilka omständigheter den kan tänkas gälla. Eftersom avvikelserna från Dulong-Petits lag dök upp vid lägre temperaturer betraktar vi fallet där kT blir mycket större än ett energikvantum, dvs  $\hbar\omega \ll kT$ . Då går uttrycket ovan faktiskt över i det klassiska resultatet för en en-dimensionell oscillator, kT. Serieutvecklingen av exponentialfunktionen i nämnaren till andra ordningen ger

$$\begin{split} \langle \varepsilon \rangle &= \frac{\hbar\omega}{\frac{\hbar\omega}{kT} + \frac{1}{2} \left(\frac{\hbar\omega}{kT}\right)^2 + \dots} + \frac{1}{2} \hbar\omega \\ &= \frac{1}{\frac{1}{kT} \left(1 + \frac{1}{2} \frac{\hbar\omega}{kT}\right) + \dots} + \frac{1}{2} \hbar\omega \\ &= kT \left(1 - \frac{1}{2} \frac{\hbar\omega}{kT}\right) + \frac{1}{2} \hbar\omega = kT \end{split}$$

Kom ihåg:

$$\frac{1}{1+x} = 1 - x \pm \dots$$

När vi fryser kan det vara skönt att slå på IR-lampan så att denna sätter våra molekyler i behaglig svängning med väsentligen samma frekvens som IR-ljuset. Saknas lampa får väl en åkarbrasa duga.

Kom ihåg att

 $e^x \to 1+x$ när $x \to 0$ 

För att bestämma den totala termiska energin hos en kristall måste vi summera de termiska medelvärdena av energin för varje oscillator, och dessa energier

$$\left< \varepsilon \right> = \frac{\hbar \omega}{e^{\hbar \omega/kT-1}} + \frac{1}{2} \hbar \omega$$

beror ju på oscillatorernas frekvens. I vår enkla modell har alla atomer samma svängningsfrekvens. I verkligheten vet vi dock att det i en kristall förekommer svängningar inom ett brett frekvensintervall. Det finns ljudvågor med låga frekvenser upp till svängningar med frekvenser som motsvarar IR ljus, dvs frekvenser av storleksordningen  $10^{14}$  s<sup>-1</sup>. Summationen över tillstånden (frekvenser) är inte enkelt genomförbar – man kände inte heller närmare till svängningarnas frekvensspektrum vid denna tidpunkt – och Einstein nöjer sig med att låta alla 3N oscillatorer (N atomer som kan svänga i 3riktningar) ha samma frekvens  $\omega_E$  och får

$$E(T) = 3N \left\{ \frac{\hbar\omega_E}{e^{\hbar\omega_E/kT} - 1} + \frac{1}{2}\hbar\omega_E \right\}$$

och därmed

$$C_V = \left(\frac{\partial E}{\partial T}\right)_V = 3Nk \left(\frac{\hbar\omega_E}{kT}\right)^2 \frac{e^{\hbar\omega_E/kT}}{\left(e^{\hbar\omega_E/kT} - 1\right)^2}$$

I gränsen för höga temperaturer  $\hbar\omega_E\ll kT$  reproduceras Dulong-Petits lag

$$C_V = 3Nk$$

Likafördelningslagen gäller följaktligen bara för tillräckligt höga temperaturer där vi kan bortse från kvanteffekter.

När Einstein jämför sitt resultat med experiment (figur 1.39) är överensstämmelsen med tanke på de ingående approximationerna mycket god.

Varför uppvisar då diamant redan vid rumstemperatur kvanteffekter, dvs avviker från Dulong-Petits lag? Det måste ju innebära att  $\hbar\omega_E \ll kT$  inte gäller, dvs att svängningsfrekvensen är ovanligt hög. Då gäller, som vi såg ovan, inte heller likafördelningslagen.

Ovan hade vi funnit att

$$\omega = \sqrt{\frac{K}{M}}.$$

Kraftkonstanten K borde vara stor i diamant eftersom diamant är ett av de hårdaste material som finns, och den atomära massan är



Figur 1.39: Einstein jämförde sitt resultat för värmekapacitetens temperaturberoende med Webers data för diamant från 1875. Einstein- temperaturen definieras av sambandet  $kT_E = \hbar\omega_E$ . Hans anpassning ger ett värde  $T_E \approx 1300$  K. (1 cal = 4,186 J)

bland de lättaste. Därav den ovanligt höga atomära svängningsfrekvensen som i sin tur leder till en kraftigt reducerad värmekapacitet redan vid rumstemperatur. Så länge den termiska energin inte räcker till att excitera en typisk svängning i diamant, dvs så länge kTär mycket mindre än avståndet mellan de kvantmekaniska energinivåerna,  $\hbar\omega_E = kT_E$ , förblir svängningen väsentligen kvar i sitt grundtillstånd och energiändringen  $\Delta E$  och därmed värmetillförseln  $\Delta Q$  för en given temperaturökning  $\Delta T$  blir mycket liten.

Einstein har inte många data att jämföra med, men passningen till Webers data för diamant är god förutom vid de lägsta temperaturerna. 1913 inkluderade P Debye svängningarnas frekvensspektrum i beräkningen och fann att

 $C_V \sim T^3$  (Debyes  $T^3$ -lag)

när vi närmar oss 0 K. Överensstämmelsen med experiment blev nu ännu mer övertygande, även vid de lägsta temperaturerna.

Men Einstein hade redan kommit fram till det i detta läge viktigaste resultatet nämligen att  $C_V$  går mot noll när temperaturen går mot noll!

Varken Plancks eller Einsteins arbeten rönte något större omedelbart intresse. Analysen av svartkroppsstrålningen med hjälp av abstrakta oscillatorer som växelverkar med strålningen i hålrummet och energikvanta som introduceras mitt i en komplicerad beräkning



Figur 1.40: För en oscillator visas energins termiska medelvärde  $\langle \varepsilon \rangle$  som funktion av temperaturen. Vid låga T dröjer oscillatorn länge kvar i sitt grundtillstånd innan den vid högre temperaturer tar upp termisk energi och börjar svänga.

begränsade antalet intresserade läsare kraftigt. Vad gällde värmekapaciteterna så var Dulong-Petits lag väl etablerad och några undantag kunde man mycket väl leva med om man skyllde undantagen på komplicerade inre effekter i materian vid lägre temperaturer. Men intresset skulle strax växa när man började se en lösning till ett känt problem inom fysikalisk kemi, nämligen hur man skulle kunna bestämma olika ämnens entropi. Entropin behövdes när man analyserade kemisk jämvikt eller ville veta huruvida en process skulle kunna ske spontant eller ej. Entropin hade därmed även stor betydelse för den kemiska industrin, där man behövde veta under vilka omständigheter en process överhuvudtaget var möjlig.

1905 hade W Nernst formulerat en preliminär version av *termodynamikens tredje lag* som numera formuleras på följande sätt:

As the temperature tends to absolute zero, the entropy of a system tends to a universal constant that is independent of chemical or physical composition or of other parameters on which the entropy may depend. The constant can be taken to be zero.

eller i en annan formulering:

If the entropy of every element in its stable state at T = 0 K is taken as zero, every substance has a posi-

#### 1.9. VÄRMEKAPACITET

tive entropy which at T = 0 K may become zero, and does become zero for all perfect crystalline substances, including compounds.

Eftersom entropiändringen  $\Delta S$  vid konstant volym kan skrivas som

$$\Delta S = \frac{\Delta Q}{T} = \frac{C_V \Delta T}{T}$$

kan vi bestämma skillnaden mellan entropierna vid två olika temperaturer  $T_f$  och  $T_i$  genom integration

$$S(T_f) - S(T_i) = \int_{T_i}^{T_f} C_V \frac{dT}{T}.$$

Därmed följer att en mätning av värmekapaciteten mellan T = 0K och den aktuella temperaturen tillåter att bestämma entropin enligt

$$S(T) = \int_0^T \frac{C_V}{T} dT$$

om S(0K) = 0. Även om man i praktiken inte kan mäta ända ner till absoluta nollpunkten så kan man använda integralformeln genom att skarva på ett lämpligt teoretiskt uttryck för  $C_V$  som gäller vid låga temperaturer.

Men för att denna metod ska kunna fungera, måste integralen konvergera eftersom entropin ska vara ändlig. Detta förutsätter då att

$$C_V \to 0$$
 när  $T \to 0$  K.

Och detta var just vad Einstein förutspådde i sitt arbete!

Nernst, som var en inflytelserik och mäktig professor i fysikalisk kemi och nu ser att Einsteins resultat är just vad hans tredje lag och beräkningen av entropin behöver, reser 1910 från Berlin ända till Zürich för att besöka Einstein. Detta besök ökar Einsteins status inom vetenskapen betydligt. Nernst skriver efteråt (i engelsk översättning):

At this time, the quantum theory is essentially a computational rule, one may well say a rule with the most curious, indeed grotesque, properties. However, ... it has borne such rich fruits in the hands of Planck and Einstein that there is now a scientific obligation to take a stand in its regard and to subject it to experimental test. Studiet av värmekapaciteter och likafördelningslagen hörde till fysikens och fysikaliska kemins standardrepertoar. När den klassiska teorin ifrågasattes av Einstein, berörde detta många fler forskare än – det mera komplicerade – problemen med svartkroppsstrålningens spektralfördelning. Kvantteorin visade sig nu ha en omedelbar praktisk betydelse som gav de nya idéerna en mycket större genomslagskraft.

Det hade ju hänt förr att relativt okända vetenskapsmän hade kommit fram till oväntade och "galna" resultat, som i stället för att få ett entusiastiskt mottagande hade begravts i arkiven. Ett dylikt fall hade inspirerat Lord Rayleigh att skriva följande:

The history of this paper suggests that highly speculative investigations, especially by an unknown author, are best brought before the scientific world through some other channel than a scientific society, which naturally hesitates to admit into its printed records matter of uncertain value. Perhaps one may go further and say that a young author who believes himself capable of great things would usually do well to secure the favourable recognition of the scientific world by work whose scope is limited, and whose value is easily judged, before embarking on greater flights.

Men Einstein hade nu fått en mäktig bundsförvant. Nernst är god vän med den belgiske sodafabrikören Ernest Solvay och ber honom sponsra den första Solvaykonferensen i oktober 1911 för att samla alla fysiker av rang inom den nya kvantfysiken. Bland den illustra deltagarskaran syns även Maurice de Broglie som konferensens sekreterare (Maurice hade genomfört uppmärksammade Röntgenabsorptionsmätningar i atomer.) Han berättar för sin bror, prins Louis de Broglie, vad som avhandlats vid mötet. Brodern blir intresserad och vill gärna deltaga i nästa möte. Men detta är tyvärr omöjligt, deltagandet är "by invitation only". Prins Louis blir stött men fortsätter, kanske just därför med ännu större iver, att fundera på kvanta och partikel-våg dualiteten

1923 leder relativistiska beräkningar honom fram till slutsatsen att om energi är relaterad till en frekvens enligt Einstein, så bör rörelsemängden vara relaterad till en våglängd enligt

de Broglies relation:  $p = \frac{h}{\lambda}$ 

Partikel-våg dualiteten gäller inte bara elektromagnetiska vågor utan även partiklar! Den sista skansen mot den nya tiden hade fal-



Figur 1.41: Deltagarna i den första Solvay-konferensen 1911

lit. Elektronen, som 26 år tidigare identifierats som partikel av JJ Thomson, har även vågegenskaper!

Det avgörande experiment som påvisade elektroners vågkaraktär kom 1927. CJ Davisson och LH Germer studerade trioden, ett vakuumrör, där elektronströmmen från katod till anod kunde styras av en tredje elektrod – ett galler som beroende på den pålagda spänningen mer eller mindre blockerade elektronernas väg. De var speciellt intresserade av vad som händer när positiva joner eller elektroner träffar en elektrod. När de studerade elektroners reflektion vid metallytor fann de tydliga *interferensmönster*. De arbetade vid ett laboratorium i USA, som senare blev mycket känt som "Bell Labs". Trioden de studerade var föregångaren till transistorn som uppfanns drygt 20 år senare vid samma labb.

Villkoret för konstruktiv interferens av de vågor som sprids från ytan av den Ni en-kristall de studerade lyder

 $d\sin\varphi = n\lambda$ 

eftersom  $d\sin\varphi$  är vägskillnaden mellan strålarna som spridits i två närmast liggande atomer enligt figur 1.44.

Eftersom elektroner är laddade partiklar förväntar vi oss att deras inträngningsdjup i metaller är litet. Men några atomlager kan de tränga in och för vissa reflektionsvinklar kan det då finnas lämpligt orienterade atomplan som kan ge upphov till *Bragg-reflektion*. I



Figur 1.42: Trioden. Elektroner emitteras av den upphettade katoden och accelereras mot anoden av spänningen U. Spänningen  $U_s$  mellan styret och katoden "styr" elektronströmmen.



Figur 1.44: Vågor reflekterade i ytlagrets atomer interfererar konstruktivt om deras vägskillnad är ett helt antal våglängder.



Figur 1.43: CJ Davisson och LH Germer (1927) studerade reflektionen av en elektronstråle i en nickelyta. För vissa vinklar ökade reflektionen markant. Detta kunde förklaras genom att, som de Broglie föreslagit, tillskriva elektronerna vågkaraktär och tolka topparna som interferenstoppar.

dessa fall leder inte bara ytlagret utan även de inre atomplanen till konstruktiv interferens och reflexerna blir särskilt starka, se fig 1.46.

Låt  $\alpha$  vara vinkeln mellan ytan och ett atomplan i metallen enligt figur 1.45. Avståndet mellan planen ges av  $D = d \sin \alpha$ .

Ur Braggs lag  $2D\sin\theta = n\lambda$  följer med  $D = d\sin\alpha$ 

 $2d\sin\alpha\sin\theta = 2d\sin\alpha\cos\alpha = d\sin2\alpha = n\lambda$ 

(där vi använt att  $\theta = 90^{\circ} - \alpha$ ) och med  $\theta = 2\alpha$  får vi  $d\sin\theta = n\lambda$ , dvs samma villkor som för reflektion enbart från ytlagret. Starkare reflexer förväntar vi oss då för reflektionsvinkeln  $\theta = 2\alpha$ .



Figur 1.45: Om både ytlagret av atomer och inre atomplan enligt Braggs lag kan reflektera strålen förväntar vi oss särskilt starka reflexer.



Figur 1.46: Kristallstrukturer var väl kartlagda genom studiet av röntgen- interferens. Braggs lag (1913):  $2D \sin \theta = n\lambda$ 



Figur 1.47: Elektroners interferensmönster som uppstår när en elektronstråle passerar en polykristallin metallfolie. (GP Thomson 1927)

1927 skulle bl a Thomsons son, GP Thomson, experimentellt bekräfta vågkaraktären genom elektroninterferenser som uppstod när elektroner passerade tunna polykristallina metallfolier, se figur 1.47.

1927 blev de Broglie personligen inbjuden att delta i det årets Solvay-konferens. 1929 belönas han med Nobelpriset. Det låg 17 år mellan Einsteins ljuskvantum-hypotes och nobelpriset 1921. Under denna tid avfärdade de flesta hans teori. 1925 publicerade E Schrödinger sina grundläggande arbeten och den formella uppbyggnaden av kvantmekaniken satte igång. Nu hade fysikersamfundet nog börjat vänja sig vid "galenskaperna" och de Broglie fick sitt Nobelpris inom 5 år.



Figur 1.48: Deltagarna i den andra Solvay-konferensen 1927

#### 1.9. VÄRMEKAPACITET

### Tidtabell

(Året för Nobelpris inom parenteser)

Statistisk fysik Kvantmekanik	Kristallstruktur Gitterdynamik	Elektroner
	1819 PL Dulong, AT	
	Petit (DP): $C_V = 3Nk$	
		1833 M Faraday: gasurladdningar
		1853 Wiedemann-Franz lag
1860 Maxwell hastighetsfördeln-		1869 J Hittorf: katodstrålar
ıng. Likafördelningslagen		
	1876 L Boltzmann: Likaf- ördelning förklarar DP	
	1895 WC Röntgen (1901)	1894 P Zeeman (1902)
		1895 JB Perrin (1926): katodstrålar bär negativ laddning
		1897 JJ Thomson (1906):elektron: <i>e/m</i>
		1897 KF Braun (1909)
1898 Dewar: väte i vätskefas		1899 JJ Thomson, CTR Wilson (1927) dimkammare: e & m
1900 M Planck (1918) svartkroppsstrålning: h		1900 P Drude: elektroner i metaller som klassisk gas
1905 A Einstein: Brownsk rörelse		1905 A Einstein (1921): Fotoel. effekt
1905 W Nernst (1920) Termodynamikens tredje lag	1907 A Einstein $C_V \rightarrow 0$ när $T \rightarrow 0$	1906 "Audion" trioden, L De Forest
1908 J Perrin (1926): Brownsk rörelse – exp		
1911 First Solvay Council	1912 M v Laue (1914)	1910 RA Millikan (1923): e (exp)
1913 N Bohr: Väteatomen (1922)	1913 WH Bragg WL Bragg (1915)	1915 RA Millikan (1923): h (exp)
	1913 P Debye:	
	$C_{\rm V} \sim T^3$ när $T \to 0$	
1923 LV de Broglie (1929)		1917 RC Tolman, TD Stewart: Elek- troner i metaller
1925 W Pauli (1945): Uteslut- nings-princip		
1926 E Schrödinger (1933)		1927 CJ Davisson, LH Germer, GP Thomson (1937) Elektroninterfer- ens

KAPITEL 1. INLEDNING

### Kapitel 2

# Elektronen som partikel

Till en början betraktar vi elektronerna i en metall som en klassisk gas. Vi diskuterar ledningsförmågan och Ohms lag, där kollisioner med kristallens joner antas bromsa motverka elektronernas acceleration i det elektriska fältet. Elektronerna rör sig då i medel med drifthastigheten. Ett magnetfält vinkelrätt mot strömriktningen ger upphov till Halleffekten, med vars hjälp vi kan avgöra de rörliga laddningsbärarnas laddning samt deras koncentration. För vissa metaller ser det ut som att laddningsbärarna är positiva. Vi kan rädda situationen genom att tänka oss positiva "hål", men vad är då dessa? Diffusion diskuteras kortfattat, dyker senare upp i pn-övergången. Elektronernas bidrag till värmekapaciteten borde vara större än experiment visar, vilket tvingar oss att ifrågasätta den klassiska beskrivningen. Utgångspunkten i nästa kapitel blir de Broglies samband som relaterar rörelsemängden till våglängd, även "partiklar" beskrivs som vågor, och detta blir grunden till en kvantmekanisk beskrivning av ledningselektronerna i metaller och andra fasta material.

#### 2.1 Elektroner i metaller som en klassisk gas

Hur går strömtransport i en ledare till på mikroskopisk nivå? P. Drude försökte 1900 (då man visste att det fanns elektroner i metaller) beskriva ledningsförmågan i en metall som transport av rörliga elektroner. Inom Drudes modell kan valenselektronerna (som är lösare bundna än elektroner i de inre atomskalen) i en metall accelereras av ett elektriskt fält. De verkar inte vara bundna till en viss atom utan man kan tänka sig att de beter sig som fria, negativt laddade partiklar som rör sig bland metallens positiva joner som i figur 2.1. Kristallen i sin helhet är neutral. Drude betraktade då elektronerna i en metall som en (ideal) klassisk gas där alla elektrostatiska krafter mellan partiklarna försummas förutom kanske just i själva kollisionsögonblicket när en elektron krockar med en jon.

Innan vi går vidare ska vi antyda vad som avses med "klassisk gas". För elektroner (och andra fermioner) gäller inom kvantmekaniken att de måste lyda Pauliprincipen, som dikterar att varje fermion



Figur 2.1: Ledningselektronerna rör sig mellan jonerna i en metall.

måste besätta sitt eget unika tillstånd, definierat av en uppsättning kvanttal (se kapitel 1). I en kvantmekanisk fermiongas kan elektroner alltså inte besätta tillstånd där det redan finns en elektron. Detta krav släpper vi helt i en klassisk modell - inget hindrar att alla elektroner har precis samma energi, eller att denna energi är noll för alla elektroner ( $\bar{E} = 0$  om T = 0 i uttrycket för kinetisk energi nedan). Vi antar att elektronerna är termiskt fördelade utan hänsyn till vad andra elektroner har för energi och att de beskrivs av Maxwell-Boltzmanns energi- eller hastighetsfördelning. Vi tänker alltså på elektrongasen precis som t.ex. heliumgas.

I termisk jämvikt säger likafördelningslagen i statistiska mekaniken (ekvipartitionsteoremet, equipartition of energy, equipartition theorem) att varje energiterm som är kvadratisk i en koordinat eller rörelsemängd bidrar med kT/2 till energins termiska medelvärde. Energins termiska medelvärde blir (för fria elektroner som enbart har kinetisk energi)

$$\bar{E} = \left(\frac{p_x^2}{2m} + \frac{p_y^2}{2m} + \frac{p_z^2}{2m}\right) = \frac{1}{2}kT + \frac{1}{2}kT + \frac{1}{2}kT$$

och elektronen har inom denna modell endast kinetisk energi

$$\bar{E} = \frac{3}{2}kT = \frac{1}{2}mv_{th}^2 \tag{2.1}$$

där  $v_{th}$  är elektronernas fart. I ett avsnitt i slutet av kapitlet finns en härledning, eller snarare ett troliggörande, av att varje frihetsgrad bidrar med kT/2 till energin. I kapitel 1 finns en mer rigorös variant som kvalificerar som härledning, men som baseras på Boltzmannfaktorn och är lite tyngre tekniskt och konceptuellt.

Vi ska nu försöka komma fram till en beskrivning av elektronernas rörelse under inflytande av ett yttre elektriskt fält. Medelvärdet av elektronernas hastighet i termisk jämvikt är  $\bar{\mathbf{v}} = 0$  eftersom de rör sig slumpartat i alla riktningar som i figur 2.2. Medelvärdet av elektronernas termiska fart bestäms av temperaturen enligt (2.1) och vid T = 300 K är farten  $v_{th} \sim 10^5$  m/s.

Vårt första mål är att förstå vilken ström som uppstår när elektronerna (med laddning q = -e) utsätts för ett elektriskt fält **E**. Då är kraften **F** 

$$\mathbf{F} = q\mathbf{E} = -e\mathbf{E}$$

och rörelseekvationen för en elektron blir

$$m\mathbf{a} = m\frac{d\mathbf{v}}{dt} = \mathbf{F} = -e\mathbf{E}$$

Här är  $v_{th} = v_{rms} = \sqrt{(v^2)}$ .  $\bar{v}$  och  $v_{rms}$  skiljer sig åt. I figur 2.2 är  $\bar{v} = \sqrt{(\bar{v})^2} = 0$ , medan  $v_{rms}$  ger ett positivt värde som vi kan använda för att beräkna gasens medelenergi.



Figur 2.2: Hastigheterna  $\mathbf{v}$  hos några elektroner i gasen.

Om det elektriska fältet är konstant i tiden accelereras elektronerna (figur 2.3) enligt

$$\mathbf{v}(t) = \mathbf{v}(0) - \frac{e\mathbf{E}}{m}t$$

Slår vi på fältet vid t = 0, dvs är  $\mathbf{E} = 0$  för  $t \leq 0$ , då är  $\mathbf{\bar{v}}(0) = 0$ . Elektronernas medelhastighet (motriktad fältet) ökar linjärt med tiden, elektronerna blir allt snabbare.

Ohms lag, å andra sidan, säger att U = RI. Motståndet R och strömmen I beror på provets dimensioner. Vi föredrar att betrakta resistiviteten

$$\rho = \frac{RA}{L} \tag{2.2}$$

eller konduktiviteten (ledningsförmågan)  $\sigma = \frac{1}{\rho}$  som är materialparametrar. Ersätter vi sedan spänningen med det elektriska fältet enligt  $|\mathbf{E}| = \mathscr{E} = U/L$  och strömmen med strömtätheten J = I/Aantar Ohms lag vektorformen

$$\mathbf{J} = \sigma \mathbf{E}.\tag{2.3}$$

Vi förväntar oss att  $J \sim v$ , dvs att strömtätheten är proportionell mot medelhastigheten hos elektronerna, och eftersom  $\mathscr{E} \sim F$  innebär Ohms lag att  $v \sim F$  och detta stämmer ju inte med Newtons ekvation

$$m\frac{dv}{dt} = F.$$

Vår slutsats måste bli att om ett konstant yttre elektriskt fält  $\mathbf{E}$  skall resultera i en konstant strömtäthet  $\mathbf{J}$  (enligt Ohms lag), då måste det finnas någon spridningsmekanism som motverkar accelerationen så att vi når ett stationärt tillstånd där elektronerna rör sig med konstant medelhastighet.

#### 2.2 Driftström

Elektronernas framfart måste alltså begränsas. Vi tänker då på kollisioner eller spridningshändelser som en möjlighet. Dessa leder till en begränsning av strömmen och yttrar sig makroskopiskt som elektrisk resistans. Efter varje kollision antar vi att elektronen fullständigt har tappat minnet av tidigare acceleration på grund av fältet och måste börja om igen från noll. Vi kan också uttrycka det som att elektronens medelhastighet efter kollisionen är noll. Utan att för tillfället närmare gå in på kollisionsprocesserna i detalj utgår



Figur 2.3: Under inflytande av ett elektriskt fält accelereras elektronerna. Deras hastighet ändras.



vi nu från att krockar mot jonerna i kristallen kan förekomma. Vi betecknar sannolikheten för att elektronen sprids genom en kollision i tidsintervallet  $dt \mod dt/\tau$ . Parametern  $\tau$  kallar vi kollisionstid.

Låt N(t = 0) vara totala antalet elektroner och N(t) antalet elektroner som överlever tiden t utan att spridas. Sannolikheten att ingen spridning sker inom intervallet dt är  $1 - dt/\tau$  och därmed följer





Figur 2.4: Antalet elektroner som överlever tiden  $\tau$  utan att spridas.

Därmed är  $e^{-t/\tau}$  sannolikheten att en elektron överlever tiden tutan spridning (se figur 2.4). Sannolikheten att en sådan elektron nu sprids inom tidsintervallet dt är

$$e^{-t/\tau}\cdot \frac{dt}{\tau}$$

 $\int_0^\infty e^{-t/\tau} \cdot \frac{dt}{\tau} = \int_0^\infty e^{-x} \cdot dx = 1$ 

Om medelvärdet av tiden fram till nästa kollision är  $\tau$ , och medelvärdet av tiden sedan den senaste kollisionen är  $\tau$ , vad blir då medeltiden mellan två kollisioner?

Svar:  $\tau$ 

Viktar vi tiden t med denna sannolikhet så kan vi bestämma medeltiden fram till nästa spridning:

$$\bar{t} = \int_0^\infty t e^{-t/\tau} \frac{dt}{\tau} = \tau \int_0^\infty y e^{-y} dy = \tau$$

Medelvärdet av tiden sedan den senaste kollisionen är också  $\tau$ . Vi finner då att elektronerna vid t = 0 i medel har accelererats av den yttre kraften under en tid  $\tau$  och medelvärdet av deras hastighet vid t = 0 blir elektronernas drifthastighet

$$\mathbf{v}_d = -\frac{e\mathbf{E}}{m}\tau$$

Vi kan också härleda en rörelseekvation för elektronernas medelhastighet under inverkan av en (möjligtvis tidsberoende) kraft. Om ingen kollision sker i intervallet dt – med sannolikhet  $(1 - dt/\tau)$  – behåller elektronerna sin medelhastighet och accelereras ytterligare av kraften F(t). Om en kollision inträffar – med sannolikhet  $dt/\tau$  – förlorar elektronerna sin medelhastighet och startar från 0 igen. Vi skriver

$$m\bar{\mathbf{v}}(t+dt) = \left(1 - \frac{dt}{\tau}\right) \left(m\bar{\mathbf{v}}(t) + \mathbf{F}(t)dt + O(dt)^2\right) + \frac{dt}{\tau} \left(0 + \mathbf{F}(t)dt + O(dt)^2\right)$$

Vi utelämnar nu alla termer av ordningen  $(dt)^2$  eller högre, eftersom vi tänker oss att  $dt \to 0$ , och får

$$\frac{m\bar{\mathbf{v}}(t+dt) - m\bar{\mathbf{v}}(t)}{dt} = \mathbf{F} - \frac{m\bar{\mathbf{v}}(t)}{\tau}$$

eller

$$m\left(\frac{d}{dt} + \frac{1}{\tau}\right)\mathbf{\bar{v}} = \mathbf{F}$$

I ett stationärt tillstånd gäller $\frac{d\bar{\mathbf{v}}}{dt}=0$ och därmed för medelhastigheten

$$\bar{\mathbf{v}} = \frac{\tau \bar{\mathbf{F}}}{m} = -\frac{e\tau \bar{\mathbf{E}}}{m} = \mathbf{v}_d \tag{2.4}$$

Nu har vi hittat ett uttryck för elektronernas medelhastighet under påverkan av ett yttre elektriskt fält. Fältet accelererar de laddade elektronerna, men deras acceleration bromsas av spridning, t.ex. mot jonerna i kristallen, och resultatet blir en medelhastighet enligt ovan. Elektronerna rör sig alltså på grund av det elektriska fältet i medel med en konstant hastighet som Ohms lag påstår. Sträckan som elektroner i medel rör sig mellan två kollisioner kallar vi *fria medelväglängden*:

$$\ell = v_{th} \cdot \tau.$$

Vilken ström ger elektronernas drift upphov till? Betrakta en area A vinkelrät mot strömmen I. Med elektronernas medelhastighet  $v_d$  i motsatt riktning till strömmen kan vi teckna strömtätheten J som

$$J = \frac{I}{A} = \frac{\Delta Q}{\Delta t \cdot A} = \frac{-enAv_d\Delta t}{\Delta t \cdot A} = -nev_d$$

där n är elektronkoncentrationen och Q betecknar laddning. Med uttrycket för drifthastigheten följer

$$\mathbf{J} = -ne\mathbf{v}_d = \frac{ne^2\tau}{m}\mathbf{E} = \sigma\mathbf{E}$$

Genom att beteckna den elektriska konduktiviteten  $\sigma$ som

$$\sigma = \frac{ne^2\tau}{m} \tag{2.5}$$

har vi återfått Ohm's lag.

Man använder ofta begreppet mobilitet eller rörlighet

$$\mu_e = \left|\frac{v_d}{\mathscr{E}}\right| = \frac{e}{m}\tau\tag{2.6}$$



Figur 2.5: Alla elektroner inom volymen  $A \cdot v\Delta t$  passerar arean A inom tidsintervallet  $\Delta t$ . Därmed passerar laddningen  $\Delta Q = -en(A \cdot v\Delta t)$  arean A per tidsenhet  $\Delta t$ .

så att

$$\sigma = \mu_e n e \tag{2.7}$$

Mobiliteten anger absolutbeloppet av förhållandet mellan drifthastighet och det elektriska fältet. I mobiliteten sammanfattas två materialberoende parametrar, m och  $\tau$ . Vi kommer senare att se att elektronens massa kan anta olika värden i olika halvledare.

**Exempel:** Beräkna spridningstiden, mobiliteten, termiska hastigheten, fria medelväglängden och drifthastigheten för ledningselektroner i Cu vid 300 K då en yttre spänning ger upphov till ett elektriskt fält på 10 mV/m.

$$\begin{array}{rcl} \rho & = & 1.72 \cdot 10^{-8} \ \Omega \mathrm{m} \\ n & = & 8.45 \cdot 10^{28} \ \mathrm{m}^{-3} \end{array}$$

(hämtade från formelsamling) ger tillsammans med  $\sigma = n e^2 \tau / m$ att spridningstiden  $\tau$ är

$$\tau = 2.44 \cdot 10^{-14} \text{ s}$$

Mobiliteten blir

$$\mu_e = \frac{e\tau}{m} = 4,32 \cdot 10^{-3} \text{ m}^2/\text{Vs}.$$

Från

$$rac{m}{2} v_{th}^2 = rac{3}{2} kT$$
 fås  $v_{th} = 1,17\cdot 10^5 ~{
m m/s}$ 

samt fria medelväglängden som

$$\ell = v_{th} \cdot \tau = 2,85 \cdot 10^{-9} \text{ m}$$

För den givna fältstyrkan får vi drifthastigheten till  $v_d = \mu_e \mathscr{E} = 4.3 \cdot 10^{-5} \text{ m/s}$  och slutligen fås strömtätheten  $J \text{ som } J = -nev_d = 5.82 \cdot 10^5 \text{ A/m}^2$ .

En koppartråd med tvärsnittsare<br/>a $1~{\rm mm}^2$ skulle med detta elektriska fält leda en ström p<br/>å $I=0,58~{\rm A}.$ 

**Kommentarer:** Av alla storheter vi beräknade för koppar ovan är endast drifthastigheten och strömtätheten beroende av den elektriska fältstyrkan  $\mathscr{E}$ . Båda dessa storheten blir också noll där fältstyrkan går mot noll - utan yttre spänning eller elektriskt fält blir det ingen ström i en koppartråd.

Här antar vi att spridningstiden är oberoende av elektrisk fältstyrka  $\mathscr{E}$ . Detta är sant så länge  $v_d \ll v_{th}$ , då vi är nära termisk jämvikt.



Figur 2.6: Antag att elektronkoncentrationen varierar längs med ett prov. Av elektronerna som finns en fri medelväglängd  $\ell$  till höger eller vänster om x = 0 rör sig i medel hälften åt vänster och hälften åt höger.

### 2.3 Diffusion (i en dimension)

Koncentrationsgradienter leder till nettoflöden av partiklar. I t ex halvledarkomponenter kan dessa flöden av laddade partiklar ge ett stort bidrag till strömmarna. Vi diskuterar här diffusionen av elektroner i ett enkelt endimensionellt fall. Elektronernas fria medelväglängd är

$$\ell = v_{th} \cdot \tau \tag{2.8}$$

där den termiska hastigheten bestäms av

$$\frac{1}{2}mv_{th}^2 = \frac{1}{2}kT$$

(1/2 i en dimension!). Vi betraktar flödet av elektroner som passerar punkten x = 0 i figur 2.6. Provet har tvärsnittsarean A. Elektroner som kommer från vänster hade sin sista kollision vid  $x = -\ell$  där koncentrationen är  $n(-\ell)$ , hälften av dessa elektroner rör sig åt höger och flödet blir

$$F^{\rightarrow} = \frac{\frac{1}{2}n(-\ell) \cdot Av_{th}\Delta t}{A \cdot \Delta t} = \frac{1}{2}n(-\ell) \cdot v_{th}$$

där  $n(-\ell) \cdot Av_{th}\Delta t$  på motsvarande sätt som i figur 2.5 är det antal elektroner som passerar x = 0 under tiden  $\Delta t$ . På motsvarande sätt är flödet från höger

$$F^{\leftarrow} = \frac{1}{2}n(+\ell)v_{th}$$

och slutligen nettoflödet

$$F = F^{\to} - F^{\leftarrow} = \frac{1}{2} v_{th} [n(-\ell) - n(+\ell)] = -v_{th} \cdot \ell \frac{dn}{dx} \qquad (2.9)$$

I ekv. 2.9 använder vi MacLaurin-utveckling;

$$n(-\ell) \approx n(0) - l\frac{dn}{dx},$$

och analogt för  $n(+\ell)$ .

Vi har då antagit att n(x) ändras endast svagt över längden  $\ell$ .

Detta flöde av elektroner ger en strömtäthet pga av diffusion som är

$$J_e = -eF = eD_e \frac{dn}{dx} \tag{2.10}$$

Diffusionskonstanten  $D_e$  och mobiliteten  $\mu_e$  är relaterade genom *Einsteins relation* 

Vi inför diffusionskonstanten eller diffusiviteten  $D_e = v_{th} \cdot \ell$ 

$$D_e = \mu_e \frac{kT}{e}.\tag{2.11}$$

För att visa relationen utgår vi ifrån sambanden

$$\left\{ \begin{array}{l} \frac{1}{2}mv_{th}^2 = \frac{1}{2}kT\\ \ell = v_{th}\cdot\tau\\ \mu_e = e\frac{\tau}{m} \end{array} \right.$$

och finner

$$D_e = v_{th} \cdot \ell = v_{th}^2 \tau = \frac{kT}{m} \tau = \mu_e \frac{kT}{e}$$

I ett och samma prov kan drift och diffusion pågå samtidigt. Totalt blir då strömtätheten för negativa laddningsbärare i form av elektroner

$$J_e = ne\mu_e \mathscr{E} + eD_e \frac{dn}{dx}$$

Som vi senare ska se kan laddningsbärare vara positiva (så kallade hål); för dessa blir strömtätheten analogt

$$J_h = p e \mu_h \mathscr{E} - e D_h \frac{dp}{dx}$$

där vi betecknar hålkoncentrationen med p.

#### 2.4 Halleffekt

Med en konduktivitetsmätning kan vi inte avgöra om de rörliga laddningarna i metaller har positiv eller negativ laddning. Är de positiva rör de sig i samma riktning som strömmen, är de negativa rör de sig åt motsatt håll. Låter vi däremot en ström gå genom ett prov vi placerat i ett magnetfält enligt figur 2.7 kommer laddningarna på grund av Lorentzkraften  $\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B})$  att avlänkas åt samma håll (oavsett laddningarnas tecken). Eftersom kraften pga magnetfältet  $\mathbf{F}_B = q\mathbf{v} \times \mathbf{B}$  är vinkelrät mot strömriktningen kommer detta att leda till en ansamling av laddningsbärare på provets

Lägg märke till teckenskillnaden framför diffusionstermen i  $J_e$  och  $J_h!$  Varför är det så? Tänk på riktningen partiklar (negativa resp. positiva) rör sig i för de två fallen, och på vilket håll strömtransporten då sker.



Figur 2.7: a) Ett strömgenomflutet Hallprov placeras i ett magnetfält. b) Nettoladdningar uppstår vid ytorna när elektroner böjs av. c) Avböjningen av hål resulterar i nettoladdningar med ombytt tecken.

ena sida – och ett motsvarande underskott på andra sidan – och laddningens tecken kan bestämmas. Denna effekt benämns Halleffekt.

Laddningsbärare som rör sig i ett magnetfält böjs av. För negativa laddningsbärare är kraften

$$\mathbf{F}_B = -e(\mathbf{v} \times \mathbf{B}) \tag{2.12}$$

Med riktningar enligt figuren pekar kraften  $\mathbf{F}_B$  i positiv x-riktning och ges av  $F_B = ev_d B$ .

Elektronernas avböjning resulterar i nettoladdningar på provets ytor. Dessa leder i sin tur till ett elektriskt fält med styrkan  $\mathscr{E}_H =$  $|\mathbf{E}_{\mathbf{H}}|$  i positiv x-riktning och därmed ytterligare en kraft som verkar på elektronerna. När denna – relativt strömriktningen – transversella kraft blivit så stor att  $e\mathscr{E}_H = F_B$  (där  $F_B = ev_d B$ ), har systemet nått steady state. Nettokraften i x-riktningen på elektronerna är då noll och elektronerna rör sig rakt fram genom provet utan att avlänkas ytterligare. Vi har då

$$e\mathscr{E}_H = ev_d B$$
 eller  $\mathscr{E}_H = v_d B$  (2.13)

Med  $J = -env_d$  följer

$$\mathscr{E}_H = \frac{J}{ne}B = \frac{I}{Ane}B \tag{2.14}$$

där A är provets tvärsnittsarea som strömmen flyter genom. Genom att sätta en voltmeter över provet kan vi mäta upp Hallspänningen  $U_H$ , som ges av  $\mathscr{E}_H \cdot b$  där b är provets bredd i x-led. Vi får då

$$U_H = \mathscr{E}_H \cdot b = \frac{Ib}{Ane}B = \frac{I}{hne}B \tag{2.15}$$

där vi skrivit om  $A \operatorname{som} b \cdot h$ , där h är provets tjocklek i y-led. Genom att mäta Hallspänningen  $U_H$  som funktion av ett varierande magnetiskt fält B kan man ur riktningskoefficienten beräkna laddningsbärarkoncentrationen n. Har man väl kalibrerat ett Hall-prov, dvs tagit reda på laddningsbärarkoncentrationen, kan man sedan använda det för att enkelt mäta magnetfält genom att mäta spänningen över Hallelementet.

I många texter introducerar man Hallkonstanten  $R_H$  som

$$R_H = \frac{\mathscr{E}_{H,x}}{J_z B_y} = -\frac{1}{ne} \tag{2.16}$$

Definierad enligt ovan blir Hallkonstanten negativ för negativa laddningsbärare. För positiva laddningsbärare är kraften från magnetfältet densamma (samma riktning och storlek) men Hallfältet byter riktning och Hallkonstanten blir positiv.

För elektroner är  $R_H < 0$ , se vidstående tabell. För positiva laddningsbärare, "hål", vore  $R_H > 0$ . Experimentet lutar åt "elektroner som laddningsbärare", men bevisföringen är bristfällig då undantag ses förekomma i tabellen.

#### 2.5 Värmeledningsförmåga

Varierar temperaturen längs med ett prov kommer ett flöde av termisk energi

$$J^Q = \frac{Q}{A \cdot \Delta t}$$

från den heta änden av provet till den kalla änden att försöka utjämna temperaturen. Hur kan vi tänka oss att detta sker? En uppvärmning av kristallen sätter fart på atomära svängningar, svängningsvågor utbreder sig i materialet och transporterar termisk energi. Men även elektronerna förväntas röra sig snabbare vid högre temperatur, dvs öka sin kinetiska energi. Denna för de med sig när de slumpartat rör sig genom kristallen.

Ett empiriskt samband för metaller var känt på Drudes tid under namnet Wiedemann-Franz lag (1853) som säger att förhållandet

Element	$-1/R_H ne$
Na	1,2
Κ	$1,\!1$
$\mathbf{Rb}$	1,0
$\mathbf{Cu}$	1,5
$\operatorname{Au}$	1,5
$\mathbf{Be}$	-0,2
Al	-0.3

Hallkonstanten för några vanliga metaller. Om vår modell gäller borde  $-1/(R_H ne) = 1$ 

Q - termisk energi ("heat")

mellan värmeledningsförmågan  $\kappa$  och produkten av elektrisk ledningsförmåga och temperatur,  $\sigma \cdot T$ , är konstant,

$$\frac{\kappa}{\sigma T} = \text{konstant}$$

Det experimentella så kallade Lorenz-talet är

$$\frac{\kappa}{\sigma T} \sim 2 \cdot 10^{-8} \ \mathrm{W}\Omega/\mathrm{K}^2.$$

Med Drudes teori – där vi endast tar med elektronbidraget – får vi (se appendix i slutet av kompendiet)

$$\frac{\kappa}{\sigma T} = \frac{3}{2} (\frac{k}{e})^2 = 1.11 \cdot 10^{-8} \text{ W}\Omega/\text{K}^2 = \text{konstant}$$

Den teoretiska förklaring av lagen upplevdes som ett starkt argument för Drudes modell – även om det sägs att överensstämmelsen ursprungligen varit bättre tack vare ett räknefel.

#### 2.6 Värmekapacitet

Värmekapacitet beskriver hur mycket medelenergin hos något ämne ändras då dess temperatur ändras

$$C_V = \left(\frac{\partial \bar{E}}{\partial T}\right),$$

Vid en given temperatur T är en elektrons medelenergi

$$\bar{E} = \frac{1}{2}mv_{th}^2 = \frac{3}{2}kT$$

enligt ekvipartitionsteorement. Därmed ges elektronernas bidrag till värmekapaciteten (per volymsenhet) av

$$C_V^{el} = n \left(\frac{\partial \bar{E}}{\partial T}\right)_V = n \cdot \frac{3}{2}k \tag{2.17}$$

Atomerna i materialet har en annan medelenergi - de är inte fria (de är bundna till varandra) och har såväl kinetisk som potentiell energi. Vi antar i det följande att bindningarna mellan atomerna kan beskrivas som (linjära) fjädrar, och att atomerna utför harmoniska svängningsrörelser runt sitt jämviktsläge. Medelenergin för en sådan atom med massa M med avseende på en dimension blir då

$$\bar{E}_x = \frac{p_x^2}{2M} + \frac{M}{2}\omega^2 x^2 = \frac{1}{2}kT + \frac{1}{2}kT = kT$$

där vi har utnyttjat ekvipartitionsteoremet. Typiskt för harmoniska oscillatorer är att de kinetiska och potentiella (medel-)energierna är Vi betraktar värmekapaciteten vid konstant volym,  $C_V$ . För fasta ämnen, till skillnad från för gaser, skiljer  $C_V$  sig ytterst lite från  $C_p$  (värmekapaciteten vid konstant tryck).



Figur 2.8: Modell av atomära bindningar.

lika stora. I tre dimensioner blir atomernas bidrag till värmekapaciteten  $C_V = 3Nk$ , där N är atomkoncentrationen.

Den totala värmekapaciteten i en metall bör vara summan av bidraget från ledningselektronerna och bidraget från de bundna atomerna. Experiment visar dock klart att den totala värmekapaciteten i princip blir den från atomerna och att  $C_V^{el} \sim 0$ ! Vi har då fått ett mycket stort problem. Det är ju självklart att partiklar i en ideal gas rör sig snabbare när temperaturen stiger och att elektrongasens medelenergi alltså har ökat med temperaturen (likafördelningslagen är mycket bestämd på den punkten). Men elektronerna verkar inte bry sig, vilket vi får återkomma till.

Så, Drudes modell är framgångsrik (den förekommer i alla läroböcker) och ger oss t.ex. Ohms lag om vi antar spridning i någon form, men det finns ett antal allvarliga problem som kvarstår i den klassiska beskrivningen:

**Värmekapacitet:** klassisk teori ger att ledningselektronernas bidrag borde vara  $C_V^{el} = \frac{3}{2}kn$ , men experimentellt finner man att  $C_V^{el} \sim 0$  och därmed att värmekapaciteten för fasta ämnen ges nästan helt av bidraget från de atomära vibrationerna.

Hur och varför beror resistiviteten på temperatur? Experimentellt finner man för metaller vid "rimliga" temperaturer att  $\rho(T) \sim T$ . Kan vi förstå det på något sätt? Vi har sambanden

$$\sigma = rac{ne^2 au}{m} \qquad 
ho = rac{1}{\sigma} \sim rac{1}{ au} \qquad \ell = v_{th} \cdot au$$

Om vi antar att spridningen som ledningselektronerna upplever härör från kollisioner med atomerna borde fria medelväglängden vara av storleksordningen avståndet mellan atomer, alltså  $\ell \sim$  konstant, då följer  $\tau = \frac{\ell}{v_{th}} \sim \frac{1}{\sqrt{T}}$  och därmed att  $\rho \sim v_{th} \sim \sqrt{T}$ ! Resonemanget ger oss alltså inte det temperaturberoende experiment visar, och vi kan dra slutsatsen att det finns något i våra antaganden som inte stämmer eller som är ofullständigt.

Halleffekt: Mätningar av Halleffekten kan indikera positiva laddningsbärare - vad skulle dessa kunna vara? I smältor kan de positiva jonerna röra sig men i en fast metall?

### 2.7 Var kommer $E_{kin} = \frac{3}{2}kT$ ifrån?

Det är inte så svårt att troliggöra att den kinetiska energin för en partikel vars energi bestäms av omgivande temperatur i medel är  $\frac{kT}{2}$  per rumsdimension. Det som följer här (baserat på boken An Introduction to Thermal Physics av Daniel Schroeder) är dock ingen härledning från grundläggande principer utan förutsätter att ideala gaslagen är sann.

Betrakta en instängd molekyl som studsar mot en vägg i en låda. När molekylen studsar mot väggen byter den riktning, och om dess hastighet i x-led före kollisionen med väggen var  $v_x$  är den nu efter kollisionen  $-v_x$ . Ändringen i hastighet blir alltså  $\Delta v_x = (-v_x) - (v_x) = -2v_x$ .

När molekylen träffar väggen utövar den en kraft på väggen (väggen utövar en lika stor kraft på molekylen, det är den kraften som orsakar molekylens rörelsemängdsförändring). Trycket på väggen blir

$$p = \frac{F_{x,\text{på väggen}}}{A} = \frac{-F_{x,\text{på molekylen}}}{A} = -\frac{m}{A} \frac{\overline{\Delta v_x}}{\Delta t}$$

där  $\frac{\overline{\Delta v_x}}{\Delta t}$  är den genomsnittliga hastighetsändringen för molekylen om vi betraktar den under en lång tid  $\Delta t$ . Vidare har vi utnyttjat att kraft ges av massa multiplicerat med hastighetsändring per tidsenhet  $(\Delta v / \Delta t)$ .

Genom att låta  $\Delta t$  motsvara den tid den tar för molekylen att gå från vänstra väggen till högra (den vi betraktar) och tillbaka, dvs  $\Delta t = 2L/v_x$ , får vi med exakt en kollision under tidsintervallet  $\Delta t$ . Vi har redan räknat ut ändringen i hastighet vid en kollision och trycket på väggen blir nu

$$p = -\frac{m}{A} \frac{(-2v_x)}{2L/v_x} = \frac{mv_x^2}{AL} = \frac{mv_x^2}{V}$$

p

Om det finns ett stort antal N molekyler i gasen blir trycket

$$= N \frac{mv_x^2}{V} \quad \text{eller} \quad pV = Nmv_x^2$$

Om vi nu tror på ideala gaslagen pV = NkT kan vi identifiera att

$$mv_x^2 = kT$$
 eller  $\frac{mv_x^2}{2} = \frac{kT}{2}$ 

Kinetiska energin för en molekyl som rör sig inte bara i x-led utan även i y- och z-led blir då

$$E_{kin} = \frac{mv_x^2}{2} + \frac{mv_y^2}{2} + \frac{mv_z^2}{2} = \frac{3}{2}kT$$

Att visa att andra frihetsgrader som t.ex potentiell energi i en harmonisk oscillator också bidrar med  $\frac{kT}{2}$  är lite svårare - se diskussionen i kompendiet i avsnitt 1.9 (kapitel 1). Där beräknas väntevärdet av energin för en en-dimensionell harmonisk oscillator. Vid höga temperaturer går medelenergin mot gränsvärdet kT. Om oscillatorns totala energi kan skrivas som  $\langle E_{kin} \rangle + \langle E_{pot} \rangle$  kan vi dra slutsatsen att  $\langle E_{pot} \rangle = \frac{kT}{2}$  eftersom vi vet att  $\langle E_{kin} \rangle = \frac{kT}{2}$ .

## Kapitel 3

# Elektronen som våg

Vågornas entré inleds med elektroners interferensfenomen, och vi skiljer på stående och fortskridande vågor. Partiklar förväntas vara lokaliserbara, varför vi inför vågpaket och skiljer mellan fas- och grupphastighet. Lokalisering av en partikel innebär att vi inte kan specificera vågvektorn exakt, rörelsemängd och läge hänger samman i en osäkerhetsrelation.

#### 3.1 Davisson-Germer interferensexperiment

Det avgörande experiment som påvisade elektroners vågkaraktär kom 1927. CJ Davisson och LH Germer studerade trioden, ett vakuumrör, där elektronströmmen från katod till anod kunde styras av en tredje elektrod – ett galler som beroende på den pålagda spänningen mer eller mindre blockerade elektronernas väg. De var speciellt intresserade av vad som händer när positiva joner eller elektroner träffar en elektrod. När de studerade elektroners reflektion vid metallytor fann de tydliga *interferensmönster*. De arbetade vid ett laboratorium i USA, som senare blev mycket känt som "Bell Labs". Trioden de studerade var föregångaren till transistorn som uppfanns drygt 20 år senare vid samma labb.

Davisson och Germer kunde analysera sina data när de väl fått läsa de Broglies arbeten och överensstämmelsen mellan experiment och teori var mycket god. Elektronernas accelerationsspänning bestämde deras kinetiska energi och därmed kunde våglängden beräknas med hjälp av *de Broglies relation* 

$$p = \frac{h}{\lambda}$$
$$E_{kin} = \frac{m}{2}v^2 = \frac{p^2}{2m} = \frac{h^2}{2m\lambda^2}$$

1

Kristallstrukturen hade tidigare kartlagts med hjälp av röntgeninterferensexperiment.



Figur 3.1: Trioden. Elektroner emitteras av den upphettade katoden och accelereras mot anoden av spänningen U. Spänningen  $U_s$  mellan styret och katoden "styr" elektronströmmen.

Erwin with his psi can do calculations quite a few, but one thing has not been seen: Just what does psi really mean?



Figur 3.2: Sannolikheten att hitta elektronen i intervallet (a,b) ges av den gråa ytan under sannolikhetstäthetsfunktionen.



Figur 3.3: Vågfunktionen som beskriver en elektron direkt efter en mätning där den hade hittats nära  $x_0$ .

Samma år observerade även GP Thomson tydliga interferensmönster när elektronstrålar passerade tunna metallfolier.

#### 3.2 Schrödingerekvationen

1926 publicerade Erwin Schrödinger, en österrikisk teoretisk fysiker som var professor i Zürich vid denna tidpunkt, sina banbrytande arbeten under titeln "Kvantisering som ett egenvärdesproblem". Einstein hade uppmärksammat betydelsen av de Broglies avhandling och när han med sin vetenskapliga auktoritet stödde de Broglies idéer, inspirerades Schrödinger att utveckla sin vågmekaniska teori. Denna teori var besläktad med den klassiska vågläran och därmed mera lättillgänglig än Heisenbergs mer abstrakta samtida matrismekanik. Schrödingers teori kunde även enkelt appliceras på handfasta problem och därmed jämföras med experiment vilket ökade teorins genomslagskraft.

Den tidsberoende Schrödingerekvationen har formen

$$i\hbar\frac{\partial\Psi}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\Psi(\mathbf{r},t) + V(\mathbf{r})\Psi(\mathbf{r},t)$$

Partikeln tilldelas här de av de Broglie förutsagda vågegenskaperna genom vågfunktionen  $\Psi(\mathbf{r},t)$ . Kan vi bestämma vågfunktionen för en elektron i ett visst system så innehåller denna funktion all information vi har om partikeln. Den talar inte om för oss exakt var partikeln befinner sig, utan bestämmer endast sannolikheten att vid en mätning hitta partikeln inom intervallet (x, x + dx) genom

$$|\Psi(x,t)|^2 dx = \Psi^*(x,t)\Psi(x,t)dx$$

(för enkelhets skull betraktar vi ofta bara en dimension) eller så gäller för ett ändligt intervall (a,b) som i figur 3.2

$$\int_{a}^{b} |\Psi(x,t)|^2 dx$$

vilket ger sannolikheten att hitta partikeln mellan a och b.

Partikelns läge är inte längre entydigt bestämt som i klassisk mekanik. Innan vi mäter läget kan partikeln vara var som helst där vågfunktionen inte är noll. Ger en mätning resultatet att partikeln ligger i (en liten omgivning) av  $x_0$  då vet vi det efter mätningen och vågfunktionen måste återspegla detta resultat, dvs den måste efter mätningen ha en skarp topp kring  $x_0$  och försvinna för övrigt. Man säger att vågfunktionen efter mätningen har kollapsat, mätningen har alltså drastiskt förändrat vågfunktionen till att se ut ungefär som i figur 3.3. Eftersom en partikel säkert finns någonstans måste dess vågfunktion kunna normeras så att

$$\int_{-\infty}^{\infty} |\Psi(x,t)|^2 dx = 1$$

Gäller normeringen för en lösning till Schrödingerekvationen vid en viss tidpunkt så kan man visa att den även gäller i all framtid.

Beskrivs en partikel av vågfunktionen  $\Psi(x,t)$ , som vi normalt förutsätter vara normerad, så kan vi beräkna *väntevärdet* av dess läge enligt

$$\langle x\rangle = \int_{-\infty}^{\infty} \Psi^*(x,t) x \Psi(x,t) dx,$$

där vi har viktat läget x med sannolikhetstätheten

$$\Psi^*(x,t)\Psi(x,t) = |\Psi(x,t)|^2.$$

Detta väntevärde ger oss medelvärdet av resultaten från många mätningar på en *ensemble* – en stor mängd – av identiska system där partikeln i varje enskilt system beskrivs av samma vågfunktion  $\Psi(x,t)$ . Eftersom en mätning förändrar partikelns vågfunktion ger en andra, omedelbart efterföljande mätning på *samma* system samma resultat som den första mätningen.

Vi kallar x för *lägesoperatorn*, den talar om för oss vad vi ska göra med den (till höger) följande vågfunktionen som t ex i väntevärdet ovan. Lägesoperatorn innebär alltså att vi helt enkelt multiplicerar vågfunktionen med x. Även operatorn som motsvarar potentiella energin, V(x), är multiplikativ.

R"orelsem"angdsoperatorn däremot är inte multiplikativ utan definieras som

$$\mathbf{p} = -i\hbar \frac{\partial}{\partial x}$$

som ber oss derivera vågfunktionen (till höger) med avseende på x och sedan multiplicera med  $(-i)\hbar$ . Rörelsemängdens väntevärde blir

$$\langle \mathbf{p} \rangle = \int_{-\infty}^{\infty} \Psi^*(x,t) \left( -i\hbar \frac{\partial}{\partial x} \right) \Psi(x,t) dx$$

Partikelns kinetiska energi som enligt klassisk mekanik ges av

$$E_{kin} = \frac{m}{2}v^2 = \frac{p^2}{2m}$$

motsvaras nu av operatorn

3-dim:  $\int |\Psi(\mathbf{r},t)|^2 d^3 \mathbf{r} = 1$ 

3-dim:  $\langle \mathbf{r} \rangle = \int \Psi^*(\mathbf{r},t) \mathbf{r} \Psi(\mathbf{r},t) d^3 \mathbf{r}$ 

3-dim:  $\mathbf{p} = -i\hbar\nabla$ 

3-dim:  $\langle \mathbf{p} \rangle = \int \Psi^*(\mathbf{r},t) (-i\hbar \nabla) \Psi(\mathbf{r},t) d^3 \mathbf{r}$ 

3-dim: 
$$-\frac{\hbar^2}{2m}\nabla^2$$
 där,  
 $\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}.$ 

$$-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}.$$

där vi ersatt p med rörelsemängdsoperatorn  $\mathbf{p}$ .

Partikelns totala energi motsvaras då av operatorn

$$\mathbf{H} = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x)$$

som kallas Hamiltonoperator.

3-dim:  $\mathbf{H} = -\frac{\hbar^2}{2m} \nabla + V(\mathbf{r})$ 

Därmed kan vi skriva Schrödingerekvationen på formen

$$i\hbar\frac{\partial}{\partial t}\Psi(x,t) = \left(-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x)\right)\Psi(x,t) = \mathbf{H}\Psi(x,t)$$

eller i tre dimensioner

$$i\hbar\frac{\partial}{\partial t}\Psi(\mathbf{r},t) = \left(-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r})\right)\Psi(\mathbf{r},t) = \mathbf{H}\Psi(\mathbf{r},t)$$

#### 3.3 Stationära lösningar

För att hitta lösningar till Schrödingerekvationen startar vi med följande ansats som separerar variablerna  $\mathbf{r}$  och t:

$$\Psi(\mathbf{r},t) = f(t) \cdot \phi(\mathbf{r})$$

Efter insättning i Schrödingerekvationen erhåller vi

$$i\hbar\phi(\mathbf{r})\cdot\frac{df(t)}{dt} = f(t)\cdot\left(-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r})\right)\phi(\mathbf{r})$$

och efter division med  $\Psi(\mathbf{r},t)$ 

$$\frac{i\hbar\frac{df(t)}{dt}}{f(t)} = \frac{\left(-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r})\right)\phi(\mathbf{r})}{\phi(\mathbf{r})}$$

Här beror funktionen till vänster om likhetstecknet enbart av tiden t och funktionen till höger om likhetstecknet enbart av läget **r**. Likhet kan endast gälla om bägge funktioner är lika med en konstant, som vi här kallar E. Vi erhåller två ekvationer, dels

$$i\hbar \frac{df(t)}{dt} = Ef(t)$$
  $f(t) = e^{-i\frac{Et}{\hbar}}$ 

och dels den tidsoberoende Schrödingerekvationen

$$\mathbf{H}\phi(\mathbf{r}) = \left(-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r})\right)\phi(\mathbf{r}) = E\phi(\mathbf{r})$$

vars lösningar ger de stationära tillstånden hos partikeln.

Ur ekvationen

 $\mathbf{H}\phi(\mathbf{r}) = E\phi(\mathbf{r})$ 

kan vi bestämma både egenvärdena E och egenfunktionerna  $\phi(\mathbf{r})$  till hamiltonoperatorn  $\mathbf{H}$ , dvs lösningen av den tidsoberoende Schrödingerekvationen motsvarar lösningen av ett egenvärdesproblem.

Den fulla lösningen ges av produkten

$$\Psi(\mathbf{r},t) = \phi(\mathbf{r})e^{-i\frac{E}{\hbar}t} = \phi(\mathbf{r})e^{-i\omega t}$$

där vi identifierar (vinkel-)frekvensen  $\omega = \frac{E}{\hbar}$ .

De stationära tillstånden, som är oerhört viktiga inom kvantmekaniken, har den egenskapen att

 $|\Psi(\mathbf{r},t)|^2 = |\phi(\mathbf{r})|^2$ 

dvs sannolikhetstätheten är oberoende av tiden. Även alla väntevärden är tidsoberoende.

I ett stationärt tillstånd är hamilton<br/>operatorns väntevärde lika med egenvärdet ${\cal E}$ 

$$\begin{aligned} \langle \mathbf{H} \rangle &= \int \Psi^*(\mathbf{r}, t) \mathbf{H} \Psi(\mathbf{r}, t) d^3 \mathbf{r} = \int \phi^*(\mathbf{r}) \mathbf{H} \phi(\mathbf{r}) d^3 \mathbf{r} \\ &= E \int \phi^*(\mathbf{r}) \phi(\mathbf{r}) d^3 \mathbf{r} = E \end{aligned}$$

där vi återigen antagit att vågfunktionen är normerad.

Standardavvikelsen

$$\sigma = \sqrt{\langle \mathbf{H}^2 \rangle - \langle \mathbf{H} \rangle^2} = 0$$

som är ett mått på spridningen i de energivärden som mätningar av energin kan ge, försvinner för stationära tillstånd eftersom

$$\mathbf{H}^2\phi(\mathbf{r}) = \mathbf{H}E\phi(\mathbf{r}) = E\mathbf{H}\phi(\mathbf{r}) = E^2\phi(\mathbf{r})$$

Därmed har totala energin alltid ett visst välbestämt värde i ett stationärt tillstånd.

Har vi två lösningar till Schrödingerekvationen,  $\Psi_1(\mathbf{r},t)$  och  $\Psi_2(\mathbf{r},t)$ , så är en linjärkombination med (komplexa) koefficienter  $c_1$  och  $c_2$ 

$$c_1\Psi_1(\mathbf{r},t)+c_2\Psi_2(\mathbf{r},t)$$

också en lösning. Man kan visa att den allmänna lösningen till Schrödingerekvationen, som ju är linjär i  $\Psi(\mathbf{r},t)$ , kan skrivas som en linjärkombination (*superposition*) av stationära, separabla tillstånd med (komplexa) koefficienter  $c_n$ :

$$\Psi(\mathbf{r},t) = \sum_{n=1}^{\infty} c_n \phi_n(\mathbf{r}) e^{-i\frac{E_n}{\hbar}t}$$

#### 3.4 Fria elektroner

För en fri elektron är  $V(\mathbf{r}) = 0$  överallt och den tidsoberoende Schrödingerekvationen blir

$$\mathbf{H}\phi(\mathbf{r}) = -\frac{\hbar^2}{2m} \nabla^2 \phi(\mathbf{r}) = E\phi(\mathbf{r})$$

eller

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right)\phi = -\frac{2mE}{\hbar^2}\phi.$$

Vi separerar variablerna med ansatsen

$$\phi(\mathbf{r}) = X(x) \cdot Y(y) \cdot Z(z).$$

Efter insättning och division med ansatsen får vi

$$\left(\frac{1}{X}\frac{\partial^2 X}{\partial x^2} + \frac{1}{Y}\frac{\partial^2 Y}{\partial y^2} + \frac{1}{Z}\frac{\partial^2 Z}{\partial z^2}\right) = -\frac{2mE}{\hbar^2},$$

en ekvation vi också kan skriva på formen

$$\frac{1}{X}\frac{\partial^2 X}{\partial x^2} = -\frac{2mE}{\hbar^2} - \frac{1}{Y}\frac{\partial^2 Y}{\partial y^2} - \frac{1}{Z}\frac{\partial^2 Z}{\partial z^2}$$

Till vänster om likhetstecknet har vi en funktion av x, på högra sidan en funktion av y och z. Likheten kräver då att funktionen till vänster om likhetstecknet måste vara konstant, dvs

$$\frac{1}{X}\frac{\partial^2 X}{\partial x^2} = -\frac{2mE_x}{\hbar^2}$$

och med analoga argument följer

$$\frac{1}{Y}\frac{\partial^2 Y}{\partial y^2} = -\frac{2mE_y}{\hbar^2} \quad \text{samt} \quad \frac{1}{Z}\frac{\partial^2 Z}{\partial z^2} = -\frac{2mE_z}{\hbar^2}$$

Konstanterna  $E_i$  (i = x, y, z) uppfyller sambandet

$$E_x + E_y + E_z = E$$

Vi har då tre ekvationer av precis samma form:

$$-\frac{\hbar^2}{2m}\frac{\partial^2 X(x)}{\partial x^2} = E_x X(x).$$

Detta är den tidsoberoende Schrödinger-ekvationen för en fri elektron i *en dimension*. Vi tittar nu lite närmare på detta en-dimensionella fall. Lösningarna – eller egenfunktionerna till hamiltonoperatorn som motsvarar elektronens kinetiska energi – är

$$X(x) \sim \sin k_x x$$
 och  $X(x) \sim \cos k_x x$ 

#### 3.4. FRIA ELEKTRONER

och energins egenvärde blir

$$E_x = \frac{\hbar^2 k_x^2}{2m}.$$

De fulla lösningarna inkluderande tidsberoendet blir

$$\Psi_1(x,t) \sim \sin k_x x \cdot e^{-i\frac{E_x}{\hbar}t} = \sin k_x x \cdot e^{-i\omega_x t} \quad \text{och}$$
$$\Psi_2(x,t) \sim \cos k_x x \cdot e^{-i\frac{E_x}{\hbar}t} = \cos k_x x \cdot e^{-i\omega_x t}$$

Dessa lösningar beskriver, som vi nu ska se<br/>, $stående \ vågor.$  Plottar vi tex

$$\operatorname{Re}(\Psi_1) = \operatorname{Re}(\sin k_x x \cdot e^{-i\omega_x t}) = \sin k_x x \cdot \cos \omega_x t$$

ser vi i figur 3.4 att vågens noder, som bestäms av funktionens nollställen, dvs av sin  $k_x x = 0$ , ligger still, vi har ingen rörelse av vågen i x-led. Samma sak gäller för

$$\operatorname{Re}(\Psi_2) = \operatorname{Re}(\cos k_x x \cdot e^{-i\omega_x t}) = \cos k_x x \cdot \cos \omega_x t$$

För att kunna beskriva transportfenomen vill vi gärna beskriva elektroner som vågor som förflyttar sig och föredrar därför lösningar på formen

$$X(x) \sim e^{ik_x x}$$
 och  $X(x) \sim e^{-ik_x x}$   $(k_x \ge 0).$ 

Energins egenvärde är fortfarande

$$E_k = \frac{\hbar^2 k_x^2}{2m}.$$

och de fulla tidsberoende lösningarna är

$$\Psi_1(x,t) \sim e^{i\left(k_x x - \frac{E_x}{\hbar}t\right)} = e^{i\left(k_x x - \omega_x t\right)}$$

 $\operatorname{och}$ 

$$\Psi_2(x,t) \sim e^{i\left(-k_x x - \frac{E_x}{\hbar}t\right)} = e^{i\left(-k_x x - \omega_x t\right)}$$

Dessa lösningar beskriver nu *fortskridande vågor*. Vi plottar realdelen i figur 3.5

$$\operatorname{Re}(\Psi_1) = \operatorname{Re}(e^{i(k_x x - \omega_x t)}) = \cos(k_x x - \omega_x t)$$

och betraktar en punkt med konstant fas, t ex ett maximum,

$$k_x x - \omega_x t = \text{konst.}$$

Vi deriverar med avseende på tiden och får

$$k_x \frac{dx}{dt} - \omega_x = 0.$$



 $\operatorname{Re} \Psi_1(\mathbf{x}, \mathbf{t}) = \operatorname{sin} \mathbf{k} \mathbf{x} \cdot \cos \omega \mathbf{t}$ 



 $\operatorname{Re} \Psi_2(\mathbf{x}, \mathbf{t}) = \cos \mathbf{k} \mathbf{x} \cdot \cos \omega \mathbf{t}$ 

Figur 3.4: Vågens realdel vid t = 0 (heldragen linje) och vid en senare tid (streckad).



 $\operatorname{Re} \Psi_1(\mathbf{x}, \mathbf{t}) = \cos(\mathbf{k}\mathbf{x} - \omega \mathbf{t})$ 



Re  $\Psi_2(x, t) = \cos(-kx - \omega t)$ 

Figur 3.5: Vågens realdel vid t = 0 (heldragen linje) och vid en senare tid (streckad).

Den valda punkten rör sig alltså med den så kallade fashastigheten

$$v_p = \frac{dx}{dt} = \frac{\omega_x}{k_x}.$$

Vågen  $\Psi_1$  rör sig i positiv x-riktning medan den andra vågen i figur 3.5,  $\Psi_2$ , rör sig i negativ x-riktning.

Lösningen till den tidsoberoende Schrödingerekvationen i tre dimensioner får enligt vår ansats produktform

$$\Psi(\mathbf{r}) \sim e^{ik_x x} e^{ik_y y} e^{ik_z z} = e^{i(k_x x + k_y y + k_z z)} = e^{i\mathbf{k}\cdot\mathbf{r}}$$

där vi nu låter  $k_i$  anta både positiva och negativa värden.

Den fulla tidsberoende lösningen blir

$$\Psi(\mathbf{r},t) \sim e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)} \qquad \omega = \frac{E}{\hbar}$$

där  $\mathbf{k}$  är vågvektorn. I en dimension kallas k vågtalet.

Vi väljer ett koordinatsystem där x-axeln ligger i **k**-vektorns riktning,  $\mathbf{k} \cdot \mathbf{r} = k_x \cdot x = kx$  där  $k = |\mathbf{k}|$ , och kan skriva vågen som

$$e^{i(kx-\omega t)}$$

Denna våg rör sig enligt diskussionen ovan i positiv x-riktning.

Vågfunktionen

 $e^{ikx} = \cos kx + i\sin kx$ 

är periodisk med våglängden  $\lambda$ ,

$$e^{ikx} = e^{i(kx+\lambda)}.$$

Detta medför att  $e^{ik\lambda} = 1$  eller  $k\lambda = 2\pi$ . Därmed följer

$$k = \frac{2\pi}{\lambda}.$$

Vågfunktionen är, som vi såg, oberoende av y och z i vårt speciella koordinatsystem, dvs den är konstant i plan som är vinkelräta mot **k**-vektorn. En uppsättning plan med inbördes avstånd  $\lambda$  och vinkelräta mot **k**-vektorn ger vågens vågfronter. Våra vågor, vars vågfronter är plan, kallas *plana vågor*.

Vågfunktionen beskriver en elektron med en viss våglängd $\lambda$ och därmed med en väldefinierad rörelsemängd enligt de Broglies relation

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


#### 3.5. VÅGPAKET

eller i tre dimensioner  $\mathbf{p} = \hbar \mathbf{k}$ 

Rörelsemängdsoperatorn var definierad som

$$\mathbf{p}=-i\hbar\nabla$$

och ger för en plan våg  $\phi = e^{i\mathbf{k}\cdot\mathbf{r}}$ 

$$\mathbf{p}\phi = -i\hbar\nabla e^{i\mathbf{k}\cdot\mathbf{r}} = \hbar\mathbf{k}\phi$$

Egenvärdet till operatorn  $\mathbf{p}$  är  $\hbar \mathbf{k}$  i överensstämmelse med de Broglies relation. Den plana vågen är *samtidigt* egenfunktion till hamiltonoperatorn och till rörelsemängdsoperatorn, både energi och rörelsemängd har skarpa värden.

Men det kvarstår två allvarliga problem. För det första gavs vågens fashastighet av

$$v_p = \frac{\omega}{k}$$

där vi med  $\omega = \frac{E}{\hbar}$  och  $E = \frac{\hbar^2 k^2}{2m}$ får

$$\omega = \frac{\hbar k^2}{2m}$$
 och därmed  
 $v_p = \frac{\hbar k}{2m} = \frac{p}{2m}$ 

vilket just är hälften av en klassisk partikelns fart,  $\frac{p}{m}$ !

För det andra, var befinner sig elektronen? Vi har

$$|\Psi(\mathbf{r},t)|^2 = |\phi(\mathbf{r})|^2 = \text{konst.}$$

Sannolikhetstätheten är oberoende av läget, vi skulle kunna hitta elektronen precis var som helst. Och vad värre är, vi kan därmed inte normera vågfunktionen!

## 3.5 Vågpaket

För att kunna beskriva transportexperiment där elektronen startar på ett ställe och rör sig mot ett annat behöver vi kunna lokalisera elektronen, vågfunktionens utsträckning måste begränsas. Själva begreppet våg kräver visserligen alltid en viss utsträckning i rummet, något punktformat objekt är inte att tänka på, men utsträckningen kan göras liten. Detta påminner oss om en analog situation inom elläran. Vill vi t ex överföra information räcker det inte att tända en lampa, vi måste även blinka, dvs skicka t ex en serie vågpaket som i Morsekod. Vi tänker oss då korta pulser där våglängd

erna adraterna som är vågkonstant



Figur 3.8: Svävningar uppstår när två vågor med något olika våglängd superponeras.

och frekvens är åtminstone ungefär givna. Lyser ficklampan rött, behåller även kortare pulser sin röda färg.

För vår elektron kan vi åstadkomma denna lokalisering genom att beskriva elektronen med ett vågpaket som byggs upp av en superposition av plana vågor.

Vi kan inledningsvis närma oss begreppet *vågpaket* genom att betrakta ett ytterst enkelt fall. Vi adderar (superponerar) två plana vågor med nästan samma frekvens och vågtal:

$$\Psi_1(x,t) = ce^{i(kx-\omega t)}$$
$$\Psi_2(x,t) = ce^{i(\{k+\delta k\}x - \{\omega+\delta \omega\}t)}$$

och får

$$\Psi = \Psi_1 + \Psi_2 = ce^{i(kx - \omega t)} \underbrace{\left\{ 1 + e^{i(\delta k \cdot x - \delta \omega \cdot t)} \right\}}_B$$

Vi ser att faktorn  $B \mod 0 \le |B| \le 2$  modulerar den plana vågen så att vi får svävningar (se figur 3.8).

Sannolikhetstätheten varierar i rummet (men elektronen är fortfarande inte lokaliserad). För en punkt med konstant värde på Bgäller

$$\delta k \cdot x - \delta \omega \cdot t = konst.$$

Tidsderivatan ger

$$\delta k \cdot \frac{dx}{dt} - \delta \omega = 0$$



Figur 3.7: Elek ett vågpaket.

#### 3.5. VÅGPAKET

och därmed följer punktens fart

$$v_g = \frac{dx}{dt} = \frac{\delta\omega}{\delta k},$$

dvs modulationen rör sig med grupphastigheten

$$v_g = \frac{\delta \omega}{\delta k}$$
 som här blir $v_g = \frac{\hbar k}{m} = \frac{p}{m}$ 

vilket just är en klassisk partikels fart. Vi är på rätt väg!

Vi hade ovan att

$$\omega = \frac{E}{\hbar}$$
 samt  $E = \frac{\hbar^2 k^2}{2m}$ .

vilket tillåter oss att uttrycka grupphastigheten som

$$v_g = \frac{\delta\omega}{\delta k} = \frac{1}{\hbar} \frac{dE}{dk}$$

Men med dessa två vågor har vi ännu inte lyckats lokalisera elektronen. Vi behöver superponera många vågor för att de genom destruktiv interferens ska kunna leda till utsläckning utanför ett visst relativt begränsat område. Vi bildar då ett allmänt vågpaket genom att superponera plana vågor viktade med koefficienter a(k)

$$\Psi(x,t) = \int a(k)e^{i(kx-\omega t)}dk$$

där a(k) är en kontinuerlig funktion centrerad kring ett visst vågtal  $k_0$ . Vi vill ju att elektronen ändå ungefär ska ha en viss rörelsemängd,  $p = \hbar k_0$ . Till varje vågtal k i utvecklingen hör en frekvens  $\omega$  enligt

$$\omega = \frac{\hbar k^2}{2m}.$$

Vi väljer nu en ytterst enkel funktion a(k) som är konstant (a(k) = 1) inom ett litet intervall  $\Delta k$  kring  $k_0$ , se figur 3.9. Därmed blir integralen enkel att utföra,

$$\Psi(x,t) = \int_{k_0 - \frac{\Delta k}{2}}^{k_0 + \frac{\Delta k}{2}} a(k) e^{i(kx - \omega t)} dk.$$

Vi betraktar först vågpaketet vid tiden t = 0:

$$\Psi(x,0) = \int_{k_0 - \frac{\Delta k}{2}}^{k_0 + \frac{\Delta k}{2}} e^{ikx} dk$$



Figur 3.9: Funktionen a(k) anger koefficienterna i vågpaketets utveckling i plana vågor.

eller om vi inför variabeln  $s = k - k_0$ 

$$\Psi(x,0) = e^{ik_0x} \int_{-\frac{\Delta k}{2}}^{\frac{\Delta k}{2}} e^{isx} ds$$

$$\Psi(x,0) = e^{ik_0x} \frac{e^{i\frac{\Delta k}{2}x} - e^{-i\frac{\Delta k}{2}x}}{ix} = e^{ik_0x} \frac{2i\sin\left(\frac{\Delta k}{2}x\right)}{ix}$$

Kom ihåg  $e^{i\alpha} = \cos \alpha + i \sin \alpha$ ,  $\sin x = x - \frac{x^2}{6} \pm \dots$ och därmed  $\frac{\sin x}{x} \to 1$  när  $x \to 0$ .

$$\Psi(x,0) = e^{ik_0x} \cdot \Delta k \frac{\sin\left(\frac{1}{2}\Delta k \cdot x\right)}{\frac{1}{2}\Delta k \cdot x}$$

Denna vågfunktion låter sig nu normeras eftersom  $|\Psi(x,0)|^2$  avtar tillräckligt snabbt när  $x \to \pm \infty$ ,

$$\int_{-\infty}^{\infty} |\Psi(x,0)|^2 dx = \int_{-\infty}^{\infty} 4 \frac{\sin^2\left(\frac{1}{2}\Delta kx\right)}{x^2} dx = 2\pi \cdot \Delta k$$

där vi utnyttjat formeln

$$\int_0^\infty \frac{\sin^2 bx}{x^2} = \frac{\pi}{2} \cdot |b|$$

som vi hittar i tillförlitliga matematiska tabeller.

Vi skriver då slutligen den normerade vågfunktionen

$$\Psi(x,0) = e^{ik_0x} \cdot \sqrt{\frac{\Delta k}{2\pi}} \cdot \frac{\sin\left(\frac{1}{2}\Delta k \cdot x\right)}{\frac{1}{2}\Delta k \cdot x} = e^{ik_0x} \cdot A(x)$$

där vi introducerat envelopfunktionen A(x).

Detta normerade vågpaket är plottat i figur 3.10 a) som visar att elektronen nu är hyfsat lokaliserad. Envelop-funktionen A(x) som modulerar vågens amplitud är plottad i figur 3.10 b) och sannolikhetstätheten,

$$|\Psi(x,0)|^2 = A^2(x)$$

i figur 3.10 c).

Om vi definierar vågpaketets vidd, som är ett mått på "osäkerheten" i läget x, lite godtyckligt som avståndet mellan punkterna bestämda av (se figur 3.10)

$$\frac{\Delta k \cdot x}{2} = \pm \frac{\pi}{2}$$

följer

$$\Delta k \cdot \Delta x = 2\pi.$$



Figur 3.10: a) Vågfunktion, b) envelopfunktion och c<br/>) sannolikhetstätheten hos vårt normerade vågpaket.



Figur 3.12: a) Envelopfunktion A(x) och b) sannolikhetstätheten för ett vågpaket med dubbelt så stort vågtalsintervall  $\Delta k$ . I b) visas även det tidigare vågpaketet (streckat) som jämförelse.

En alternativ form av denna relation följer med hjälp av de Broglies samband $p=\hbar k$ 

$$\Delta p \cdot \Delta x = h$$

Ju mer vi försöker lokalisera elektronen, desto "osäkrare" blir dess rörelsemängd – och tvärtom. Vi har härmed kommit fram till en relation som är mycket lik

Heisenbergs osäkerhetsrelation: 
$$\Delta p \cdot \Delta x \geq \frac{\hbar}{2}$$



I Heisenbergs osäkerhetsrelation är osäkerheterna definierade av mätvärdenas standardavvikelser.



 $\Delta k$ 

2

k

a

1,0

0

0

#### 3.6 Vågpaketets rörelse

Ett allmänt vågpaket där en godtycklig funktion a(k) är begränsad till ett litet intervall  $\Delta k$  kring  $k_0$  är

$$\Psi(x,t) = \int_{k_0 - \frac{\Delta k}{2}}^{k_0 + \frac{\Delta k}{2}} a(k)e^{i(kx - \omega t)}dk.$$
$$\Psi(x,t) = e^{i(k_0x - \omega_0 t)} \int_{k_0 - \frac{\Delta k}{2}}^{k_0 + \frac{\Delta k}{2}} a(k)e^{i((k-k_0)x - (\omega - \omega_0)t)}dk$$

 $\operatorname{eller}$ 

$$\Psi(x,t) = e^{i(k_0 x - \omega_0 t)} \int_{k_0 - \frac{\Delta k}{2}}^{k_0 + \frac{\Delta k}{2}} a(k) e^{i(k - k_0) \left(x - \frac{\omega - \omega_0}{k - k_0} t\right)} dk.$$

Eftersom integralen antogs vara begränsad till ett litet område kring  $k_0$  utvecklar vi $\omega(k)$ i en Taylor-serie

$$\omega(k) = \omega(k_0) + \left. \frac{\partial \omega}{\partial k} \right|_{k=k_0} \cdot (k-k_o) + \dots$$

i en omgivning av  $k_0$  och försummar högre ordningens termer. Med

$$\frac{\omega - \omega_0}{k - k_0} \quad \rightarrow \quad \frac{\partial \omega}{\partial k} \bigg|_{k = k_0} = v_g$$

där  $\omega_0 = \omega(k_0)$ , får vi

$$\Psi(x,t) = e^{i(k_0 x - \omega_0 t)} \int_{k_0 - \frac{\Delta k}{2}}^{k_0 + \frac{\Delta k}{2}} a(k) e^{i(k - k_0)(x - v_g t)} dk$$
$$= e^{i(k_0 x - \omega_0 t)} \cdot A(x - v_g t)$$

där envelop-funktionen

$$A(x - v_g t) = \int_{k_0 - \frac{\Delta k}{2}}^{k_0 + \frac{\Delta k}{2}} a(k) e^{i(k - k_0)(x - v_g t)} dk$$

beror av läge och tid endast i kombinationen  $x - v_g t$ .

En generell vågpuls som rör sig åt höger med konstant hastighet v beskrivs av en funktion f(x - vt) (se fig 3.13).

Därmed rör sig även envelopfunktionen med hastighet  $v_g$ , dvs med grupphastigheten

$$v_g = \left. \frac{\partial \omega}{\partial k} \right|_{k_0}.$$



Figur 3.13: Vid t = 0 har punkten  $x_1$  funktionsvärdet  $f(x_1)$ . Samma funktionsvärde har vi i punkten  $x_2$  vid tiden t om  $x_1 = x_2 - vt$  eller  $x_2 = x_1 + vt$ .

Vi har hittills bortsett från en viktig effekt, nämligen att vågpaketets form inte förblir konstant i tiden. Eftersom vi har en osäkerhet i rörelsemängd,  $\Delta p$  (och därmed i hastighet) kommer osäkerheten i position,  $\Delta x$ , att öka med tiden. Om man inte vet exakt hur fort ett objekt rör sig kommer man ju med tiden att bli mer och mer osäker på var det befinner sig. Denna effekt försvann då vi utvecklade  $\omega(k)$  i en Taylorserie och endast inkluderade termer upp till första ordningen i k. Högre ordningens termer kan antas vara små då  $k \approx k_0$ , alltså för små  $\Delta k$ , men eftersom de är proportionella mot t kommer de trots allt att blir stora om vi väntar tillräckligt länge. Vi försöker oss här inte på någon matematisk analys av denna effekt, utan noterar bara att envelop-funktionen kommer att bli bredare och bredare med tiden, vilket uttrycker en ökad osäkerhet i elektronens position.

# 3.7 Dispersion

För fria elektroner gäller  $E = \frac{p^2}{2m} = \frac{\hbar^2 k^2}{2m}$ . Fashastigheten blir  $v_p = \frac{\omega}{k} = \frac{E}{\hbar k} = \frac{\hbar k}{2m} = \frac{p}{2m}$ och grupphastigheten  $v_g = \frac{\partial \omega}{\partial k} = \frac{1}{\hbar} \frac{\partial E}{\partial k} = \frac{\hbar k}{m} = \frac{p}{m}$ .

Elektronens grupphastighet överensstämmer med elektronens partikelhastighet i den klassiska teorin. Fas- och grupphastigheter är i allmänhet olika. Resultaten ovan gäller för fria elektroner - som vi ska se senare, kan E(k) se annorlunda ut vilket leder till andra fasoch grupphastigheter. Figur 3.14 visar ett fall där fashastigheten är större än grupphastigheten.

Fas- och grupphastighet beror av våglängden (eller frekvensen), elektronvågorna uppvisar *dispersion*.



Figur 3.14: Ett vågpaket rör sig med grupphastigheten. Kryss markerar vågpaketets tyngdpunkt". Fashastigheten markeras med pilar (följer en specifik vågtopp).

För elektromagnetiska vågor i vakuum är  $\omega = ck$  och därmed  $v_p = v_g = c$ . När elektromagnetiska vågor rör sig i ett material (glas, plast, osv) är brytningsindex frekvensberoende,  $n = n(\omega)$ , sambandet mellan frekvens och vågvektor blir

$$\omega = \frac{c}{n(\omega)}k$$

och därmed fashastigheten  $v_p = \frac{c}{n(\omega)}$ . Begreppet dispersion uppkom i optikens begynnelse när Newton upptäckte att vitt ljus sprids – "is dispersed" – efter passage genom ett prisma eller en välslipad diamant som i figur 3.15. Diamant uppvisar stark dispersion, brytningsindex varierar mellan 2,407(rött) och 2,465(violett) i det synliga området. Diamant har även ett ovanligt stort värde på brytningsindex, denna särställning delar diamanten med kisel (n = 3,5)och germanium (n = 4) som ligger strax under kol i grupp IV i det periodiska systemet. Ett stort brytningsindex ökar andelen ljus som totalreflekteras i botten av den slipade diamanten i figur 3.15 och bidrar till diamantens färgstarka glitter.



Figur 3.15: Newton uppmätte n = 2,44 för gult ljus i diamant.

# KAPITEL 3. ELEKTRONEN SOM VÅG

# Kapitel 4

# Frielektronmodellen

I vår första kvantmekaniska modell för en metall betraktar vi den som en oändligt djup potentialbrunn. Schrödingerekvationen ger elektronernas energier och vågfunktioner i brunnen. För en oändligt djup brunn är vågfunktionerna stående vågor. För att kunna beskriva elektroners rörelse genom metallen väljer vi att använda periodiska randvillkor vilket ger oss möjlighet att beskriva vågfunktionerna som fortskridande vågor. I ett stort system ligger tillstånden mycket tätt och vi beskriver dem med hjälp av tillståndstätheten. I metallens grundtillstånd fyller elektroner tillstånden nerifrån och upp till Fermienergin enligt Pauliprincipen. När temperaturen stiger exciteras i första hand elektroner nära Ferminivån. Fermi-Dirac funktionen håller reda på sannolikheten att ett tillstånd med viss energi är besatt vid en viss temperatur. Elektronernas reaktion på uppvärmning är ytterst svag eftersom Fermienergin i metaller är mycket större än den termiska energin kT. Först vid mycket höga temperaturer når elektrongasen gränsen där den kan beskrivas som en klassisk gas. Analysen av ledningsförmågan liknar den inledande partikel-diskussionen, dock ersätts den termiska hastigheten av Fermi-hastigheten.

Figur 4.1 visar en liten Na-"kristall". En Na-atom har elektronkonfigurationen  $1s^22s^2p^63s$ . De inre elektronerna  $(1s^22s^2p^6)$  är hårt bundna till kärnan. De förblir därmed atomära och påverkas endast



Figur 4.1: En liten Na-kristall med tre atomer. Atomära potentialer är streckade, deras superposition heldragen (tunn). Lådpotentialen inom frielektronmodellen är heldragen (tjock).

svagt av grannatomerna. Valenselektronen (3s) däremot förväntas bidra till ledningsförmågan, natrium är ju trots allt en metall, och vi antar att 3s-elektronen inte längre är bunden till en viss atom utan blir delokaliserad, endast bunden till kristallen i sin helhet.

I en första, enklast möjliga, modell ska vi nu studera valenselektronerna i en kvantmekanisk beskrivning. Vi antar att de är fria, oberoende partiklar så när som på att de befinner sig i en oändligt djup potentialbrunn. Att elektronerna är fria betyder att de rör sig i en konstant potential (V(x)=0 inuti brunnen) så att kraften som verkar på dem är

$$F = -\frac{\partial V}{\partial x} = 0 \tag{4.1}$$

Oberoende är de när vi bortser från deras inbördes Coulombrepulsion. Pauli-principen som har synpunkter på tillståndens besättning kommer däremot att vara väsentlig - den hindrar att mer än en elektron besätter ett visst tillstånd. Den enda funktion kristallens joner fyller i denna första modell är att stå för den positiva bakgrundsladdning som binder elektronerna till lådan (metallbiten) med elektriska krafter och se till att metallen i sin helhet är neutral.

Vi modellerar i det följande den ändliga, makroskopiska kristallen genom att beskriva den som en oändligt djup potentialbrunn.

#### 4.1 Elektron i en lådpotential

$$\mathbf{H}\phi(x) = \left(-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V(x)\right)\phi(x) = E\phi(x) \tag{4.2}$$

I vissa situationer är det väsentligt att lådpotentialens väggar är ändliga, t ex för termisk emission av elektroner ut ur metallen som i en glödkatod, men här låter vi väggarna vara oändliga så att V(x) =0 för  $0 \le x \le L$  och  $V(x) = \infty$  för övriga x.

Vi har då

$$-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\phi(x) = E\phi(x)$$

med den allmänna lösningen  $\phi(x) = A\sin(kx) + B\cos(kx)$ . Insättning i Schrödinger-ekvationen ger energin

$$E = \frac{\hbar^2 k^2}{2m}$$

med vågtalet k som är relaterat till våglängden  $\lambda$  genom

$$k = \frac{2\tau}{\lambda}$$

Hamiltonoperatorn **H** innehåller en del relaterad till kinetisk energi och en del relaterad till potentiell energi.

Inom spektroskopi används det så kallade spektroskopiska vågtalet;  $\nu = \frac{1}{\lambda}$ , som för elektromagnetiska vågor (ljus) ger energin genom  $E = hc\nu$ . Här använder vi vinkelvågtalet (angular eller circular wave number),  $k = \frac{2\pi}{\lambda}$ , som oftast kallas bara vågtal.



Figur 4.2: Lådpotentialens vågfunktioner  $\phi(x)$  och sannolikhetstätheter  $|\phi(x)|^2$ .

Randvillkoren för en oändlig potentialbrunn kräver att  $\phi(0) = \phi(L) = 0$ , vilket ger att B = 0 samt att  $\sin(kL) = 0$ . Det senare villkoret ger att vågtalen k måste uppfylla relationen  $k \cdot L = \pi \cdot p$ , och alltså att

$$k = \frac{\pi}{L} \cdot p$$
 för  $p = 1, 2, 3, \dots$ 

Endast våglängder som uppfyller  $\lambda = \frac{2L}{p}$  är tillåtna, vilket vi kan se i figur 4.2. T ex är våglängden hos tillståndet p = 4 i figur 4.2 given av  $\lambda = L/2$ .

Vi får alla tillåtna tillstånd om vi väljer k > 0. Det är inte meningsfullt att inkludera både negativa och positiva k-värden. Vågfunktionen  $\phi(x) \sim \sin(kx)$  skiljer sig bara på en fasfaktor  $(-1 = e^{i\pi})$  från  $\phi(x) \sim \sin(-kx)$ , då  $\sin(-kx) = -\sin(kx)$ . Energier, sannolikhetsfunktioner och väntevärden är de samma för båda vågfunktionerna och de är alltså i allt väsentligt samma vågfunktion.

De tidsoberoende lösningarna blir

$$\phi(x) = \sqrt{\frac{2}{L}}\sin(kx)$$

och den tidsberoende lösningen ges av

$$\phi(x,t) \sim \sin(kx)e^{-iEt/\hbar}$$

där vi ser att nodernas lägen inte ändrar sig med tiden (se även kapitel 3). Vågfunktionerna är således stående vågor.



Figur 4.3: Frielektronparabeln  $E = \frac{\hbar^2 k^2}{2m}$ . Tillåtna *k*-värden ges av  $\pi/L \cdot p$  om vi kräver att vågfunktionerna ska gå mot noll på randen.

#### 4.2 Lådpotential i tre dimensioner

I tre dimensioner lyder Schrödinger-ekvationen

$$-\frac{\hbar^2}{2m} \left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) \phi(x, y, z) = E\phi(x, y, z)$$
(4.3)

och vågfunktionerna antar då produktform som vi såg i kapitel 3.

$$\phi(x,y,z) = A \cdot \sin k_x x \cdot \sin k_y y \cdot \sin k_z z \tag{4.4}$$

där A är en normeringskonstant. Vi antar att vår "metall" har formen av en kub med volym  $V = L^3$ . Randvillkoren kräver då att vågfunktionen försvinner på kubens ytor, dvs t.ex. att sin  $k_z z = 0$ för z = L (analogt för de andra faktorerna i ekv. 4.4). Detta krav är uppfyllt om

$$k_i = \frac{\pi}{L} n_i \qquad (i = x, y, z)$$

Vi erhåller energierna

$$E = \frac{\hbar^2 k^2}{2m} = \frac{\hbar^2}{2m} (k_x^2 + k_y^2 + k_z^2) = \frac{\hbar^2}{2m} \frac{\pi^2}{L^2} (n_x^2 + n_y^2 + n_z^2)$$

Varje tillstånd i den tre-dimensionella kvantbrunnen definieras alltså av tre kvanttal  $(n_x, n_y, n_z)$ . Tillsammans med spinnkvanttalet bestämmer de entydigt ett elektroniskt tillstånd.

# 4.3 Planvågslösningar

Vi tittar igen på Schrödinger-ekvationen i tre dimensioner:

$$\mathbf{H}\phi(\mathbf{r}) = -\frac{\hbar^2}{2m} \nabla^2 \phi(\mathbf{r}) = -\frac{\hbar^2}{2m} (\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2})\phi(\mathbf{r}) = E\phi(\mathbf{r})$$

Lösningarna vi diskuterade i föregående avsnitt var stående vågor. För att beskriva rörelse hos elektronerna och därmed även laddningstransport vore det behändigt att i stället använda planvågslösningar, dvs gående vågor, på formen

$$\Psi(\mathbf{r},t) \sim e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)} \tag{4.5}$$

Vi kan som tidigare (se kap. 3) skriva lösningen till den tidsoberoende Schrödinger-ekvationen som

$$\phi(\mathbf{r}) \sim e^{i\mathbf{k}\cdot\mathbf{r}}$$

med energin

$$E = \frac{\hbar^2 k^2}{2m}$$

Elektronens (vinkel) frekvens är  $\omega = E/t$ 



Figur 4.4: Ett vågpaket  $\phi(x)$  med längd  $\Delta x$  har skapats genom att "klippa ut" en bit ur en oändligt utsträckt våg. Detta vågpaket kan byggas upp av planvågor med koefficienter a(k), vars topp ligger vid  $k = 2\pi/\lambda$  och som har en bredd  $\Delta k \sim 1/\Delta x$ . Är vågpaketet givet kan a(k) bestämmas med Fourieranalys.

men vågfunktionen  $\phi(\mathbf{r})$  är nu, till skillnad från de stående vågor som uppfyller fixa randvillkor, en egenfunktion till rörelsemängdsoperatorn med egenvärdet  $\hbar \mathbf{k}$ . Planvågslösningarna är egenfunktioner till både Hamiltonoperatorn och till rörelsemängdsoperatorn,

$$\mathbf{p}\phi = -i\hbar\nabla\phi = \hbar\mathbf{k}\phi$$

både (kinetisk) energi och rörelsemängd är då samtidigt entydigt bestämda i varje egentillstånd  $\phi$ . Vi kan skriva

$$E = \frac{\hbar^2 k^2}{2m} = \frac{p^2}{2m}$$

ett resultat som väl överensstämmer med den klassiska mekaniken.

Med dessa fortskridande vågor kan vi sedan bilda vågpaket (se figur 4.4 och 4.5) genom att addera vågor med lämpliga koefficienter, som finns mer utförligt diskuterat i kapitel 3.

Kom ihåg att en lokalisering i x, dvs i det verkliga rummet, kräver en viss osäkerhet i vågvektor vilket här leder till en relation mellan osäkerheten i läge och osäkerheten i rörelsemängd (Heisenbergs osäkerhetsrelation  $\Delta x \cdot \Delta p \geq \frac{\hbar}{2}$ ). Kom ihåg de Broglie:  $p = \frac{h}{\lambda} = \frac{h}{2\pi} \frac{2\pi}{\lambda} = \hbar k$ 

Rörelsemängden hos en fri elektron är  $\mathbf{p} = \hbar \mathbf{k}$ .

$$M/M \sim k_1$$

Figur 4.5: Elektronerna far fram som vågpaket.

#### 4.4 Periodiska randvillkor

Det som avgör om lösningarna till Schrödingerekvationen är på formen 4.4 (dvs stående vågor) eller på formen 4.5 (dvs planvågor eller gående/fortskridande vågor) är valet av randvillkor. Tar vi ett metallprov och kräver att vågfunktionen försvinner på ytan - som vi gjorde tidigare - tvingas vi använda stående vågor. Detta betyder att fortskridande vågor förr eller senare reflekteras av ytan. Så länge yteffekter inte intresserar oss vore det bekvämt att kunna göra sig av med ytor och därmed även reflektioner. För allt annat än mycket små prover kommer yteffekter inte att vara så viktiga för de egenskaper vi ska diskutera - de flesta elektroner är långt från ytan.

Vi inför nu lämpliga randvillkor som faktiskt tillåter fortskridande vågor. Antag att provet har formen av en kub med kantlängd L och volymen  $V = L^3$  - detta är vår metallbit som kan ha en makroskopisk storlek, t.ex. 1 mm<sup>3</sup>. Nu fyller vi hela rummet genom att periodiskt repetera kuben, som i figur 4.6. Då försvinner ytorna och därmed reflektionerna. Men hur kan vi nu ta hänsyn till att provet ändå är ändligt? Vår periodiska upprepning får ju inte leda till att vågfunktioner kan ha mycket längre våglängd, dvs breda ut sig över större volym. För stående vågor är tillståndens energier omvänt proportionella mot  $L^2$  och vi förväntar oss ett liknande resultat även här. Vi kräver då att provets kub och de repeterade kuberna är helt identiska även vad beträffar elektronernas vågfunktioner. Vi skriver dessa periodiska randvillkor som

$$\phi(x,y,z) = \phi(x+L,y,z) = \phi(x,y+L,z) = \phi(x,y,z+L) \quad (4.6)$$

I en dimension kan vi tänka oss dessa randvillkor som om vi hade knutit kristallen till en ring (se fig. 4.7) och krävt att vågfunktionen skall ha ett entydigt värde när vi gått ett varv runt ringen. För vågfunktionen

$$\phi(\mathbf{r}) \sim e^{i\mathbf{k}\cdot\mathbf{r}} = e^{i(k_x\cdot x + k_y\cdot y + k_z\cdot z)} = e^{ik_xx} \cdot e^{ik_yy} \cdot e^{ik_zz}$$

följer ur randvillkoren (eq. 4.6):

$$e^{ik_xL} = e^{ik_yL} = e^{ik_zL} = 1$$

och därmed

$$k_i = \frac{2\pi}{L} n_i, \qquad n_i = 0, \pm 1, \pm 2, \dots$$

Observera att vi nu – i motsats till fallet med stående vågor – behöver alla heltal  $n_i$ , positiva som negativa inklusive  $n_i = 0$ . Funktionerna  $e^{ikx}$  och  $e^{-ikx}$  är olika vågfunktioner och kan inte uttryckas i



Figur 4.6: Vi låter vår metall ha formen av en kub med kantlängden L, dvs ha volymen  $V = L^3$ .



Figur 4.7: Periodiska randvillkor innebär att vågfunktionen återfår samma värde (och samma derivata) efter ett varv i en tänkt cirkel vars omkrets motsvarar systemets (lådans) storlek.

varandra med hjälp av en multiplikativ fasfaktor, som i fallet med de stående vågorna. Vågvektorns komponenter  $\mathbf{k} = (k_x, k_y, k_z)$  antar därmed även negativa värden. Vågvektorns riktning bestämmer vågens utbredningsriktning.

Insättning av vågfunktionen i Schrödinger-ekvationen

$$-\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right)\phi(x,y,z)=E\phi(x,y,z)$$

ger som tidigare

$$E = \frac{\hbar^2 \mathbf{k}^2}{2m} = \frac{\hbar^2}{2m} (k_x^2 + k_y^2 + k_z^2)$$

I fallet med den oändligt djupa tre-dimensionella brunnen lät vi tre heltal  $(n_x, n_y, n_z)$  identifiera ett tillstånd. Framöver kommer vi att låta vågvektorns komponenter vara tillståndens "namn" – vi går över till *k-rummet* och identifierar varje vågfunktion eller tillstånd med hjälp av dess vågvektor  $(k_x, k_y, k_z)$ . Som vi kommer att se kan vi t ex beskriva kollisioner mellan elektronvågor och atomära vibrationsvågor som kollisioner mellan två "partiklar" med var sin "rörelsemängd"  $\hbar \mathbf{k}$ .

*k-vektorn*  $\mathbf{k} = (k_x, k_y, k_z)$  är ett mycket centralt begrepp i beskrivningen av fasta material. Den ingår i bevarandelagar och (därmed) i urvalsregler, samt är för fria elektroner en etikett som, tillsammans med spin-kvanttalet, entydligt definerar deras tillstånd och som även ger energi och rörelsemängd för tillståndet.



Figur 4.8: För fortskridande vågor bestämmer vågvektorns riktning (i en dimension dess tecken) vågens utbredningsriktning. Inom fri-elektronmodellen motsvarar ett visst k-värde en viss hastighet. En elektron i ett visst tillstånd har alltså en viss hastighet i rummet, men förflyttar sig inte i k - den ligger kvar i sitt tillstånd om den inte sprids.



Figur 4.9: För en oändligt djup tredimensionell potentialbrunn ger Schrödingerekvationen (med fixa randvillkor)  $E = \frac{\hbar^2}{2m} (\frac{\pi}{L})^2 (n_x^2 + n_y^2 + n_z^2) = \frac{\hbar^2}{2m} (\frac{\pi}{L})^2 n^2$ . Figuren visar totala antalet tillstånd dS(E) i små energiintervall dE vid olika energier E som funktion av  $n^2$  (dvs energin E). Totalt har ca 10<sup>5</sup> energivärden använts för att konstruera bilden. Den släta kurvan är en anpassning av en  $\sqrt{E}$ -kurva till medelvärdet av antalet tillstånd per energiintervall - dvs tillståndstätheten.

#### 4.5 Tillståndstäthet

För allt annat än mycket små prover kan vi anta att tillstånden ligger mycket tätt, då tillståndens energier beror på lådpotentialens storlek L som  $E \sim L^{-2}$ . Redan i ett litet prov kommer antalet elektroner som vi placerar i de tillåtna tillstånden att vara mycket stort. Varje tillstånd identifieras av  $\mathbf{k} = (k_x, k_y, k_z)$ , och en fullständig beskrivning av ett elektroniskt tillstånd kräver också att vi anger elektronens spinn. Det blir mycket opraktiskt för oss att hålla reda på **k**-vektorn och energin för alla enskilda elektroner.

I figur 4.9 visas antalet tillstånd inom ett litet energiintervall runt en viss energi som funktion av energin i en tredimensionell kubisk potentialbrunn.

Åven om figur 4.9 innehåller många tillstånd, är de långt ifrån tillräckligt många för att rymma alla valenselektroner ens i en mycket liten metallbit (ca  $10^{11}$  elektroner i ett prov som är  $1 \ \mu m^3$  om elektronkoncentrationen  $n = 5 \cdot 10^{28} m^{-3}$ ). Vi ser att vi inte har någon möjlighet att hålla reda på alla enstaka elektroniska tillstånd, istället inför vi begreppet tillståndstäthet (density of states).

Tillståndstätheten, Z(E), är definierad som antalet tillstånd per energiintervall;

$$Z(E) = \frac{dS(E)}{dE}$$

där S(E) är totala antalet tillstånd med en energi upp till E. Antalet tillstånd dS(E) med energier inom intervallet (E, E + dE) ges då av Z(E)dE. Tillståndstätheten Z(E) berättar alltså för oss hur många tillstånd det finns i ett litet intervall dE vid energin E.

Vi ska nu ta fram ett uttryck för tillståndstätheten och utgår då från planvågslösningarna till Schrödingerekvationen. En **k**-vektor  $(k_x,k_y,k_z)$  definierar ett tillstånd med energi  $E = \frac{\hbar^2}{2m}(k_x^2 + k_y^2 + k_z^2)$ . Varje punkt  $\mathbf{k} = (k_x,k_y,k_z)$  i **k**-rummet innebär två elektroniska tillstånd, då vi har två olika spin-tillstånd.

Om vi tänker oss att vi fyller vår tänkta kubiska potentiallåda med elektroner, så kommer den första elektronen att hamna i det allra lägsta energitillståndet i lådan, där  $(k_x,k_y,k_z) = (0,0,0)$ . Nästa elektron besätter tillståndet med  $(k_x,k_y,k_z) = (\pm 2\pi/L,0,0)$  (alternativt  $(k_x,k_y,k_z) = (0, \pm 2\pi/L,0)$  eller  $(0,0, \pm 2\pi/L)$ ). Vi inser att avståndet till origo är avgörande för energin och skriver

$$E = \frac{\hbar^2}{2m} (k_x^2 + k_y^2 + k_z^2) = \frac{\hbar^2 k^2}{2m},$$
(4.7)

där k är radien i en tänkt sfär på vars yta vi finner **k**-vektorer för de tillstånd som har en viss energi E.

Antalet tillstånd S(k) inom en sådan sfär, dvs antalet tillstånd med **k**-vektor kortare än k eller, ekvivalent, energi mindre än en viss energi E, ges av sfärens volym dividerat med den volym varje kpunkt upptar:

$$S(k) = 2\frac{4\pi}{3}k^3/(\text{volym per }\mathbf{k}\text{-punkt})$$
(4.8)

där faktorn 2 kommer från spin-degenerationen. Hur tätt ligger nu k-vektorerna i k-rummet, dvs vad är volymen per  $\mathbf{k}$ -punkt? I en dimension är avståndet mellan närliggande tillstånd  $\Delta k = 2\pi/L$ , det betyder att vi har en  $\mathbf{k}$ -vektor per kub  $(2\pi/L)^3$ . Vi kan också uttrycka detta som att  $\mathbf{k}$ -vektorernas täthet  $g(\mathbf{k})$  i  $\mathbf{k}$ -rummet är

$$g(\mathbf{k}) = \frac{L^3}{(2\pi)^3} = \frac{V}{(2\pi)^3}$$

och antalet tillstånd S(E) med en energi upp till E ges då av

$$S(E) = 2 \cdot g(\mathbf{k}) \cdot \frac{4\pi}{3} k^3 = 2 \cdot \frac{V}{(2\pi)^3} \cdot \frac{4\pi}{3} \left(\frac{2mE}{\hbar^2}\right)^{3/2}.$$



Figur 4.10: Alla tillstånd på ett visst avstånd  $k = \sqrt{k_x^2 + k_y^2 + k_z^2}$ från origo har samma energi.

**k**-rummet kallas också reciproka rummet eftersom avstånd här mäts i  $m^{-1}$ . Tillståndstätheten blir slutligen

$$Z(E) = \frac{dS(E)}{dE} = \frac{V}{2\pi^2} (\frac{2m}{\hbar^2})^{3/2} \sqrt{E}$$

eller

$$Z(E)dE = C\sqrt{E}dE \qquad C = \frac{4\pi L^3 (2m)^{3/2}}{h^3}$$
(4.9)

Ofta utelämnas faktor<br/>n $V=L^3$ - man talar då om tillståndstäthet per volymsenhet.

Vi har bestämt lådpotentialens elektrontillstånd och nu även tillståndstätheten. Det sista steget vi nu ska ta är att placera ut elektronerna i dessa tillstånd. Vår lådpotential representerar ju en bit metall med volym  $V = L^3$  och innehåller många elektroner, deras antal betecknar vi med N. När vi nu ska lägga elektronerna i lådans tillstånd måste vi komma ihåg Pauli-principen som utesluter möjligheten att fylla ett tillstånd med mer än en elektron. Vi börjar då med tillstånden med lägst energi, dvs med **k**-vektorer kring  $\mathbf{k} = 0$  för att sedan fylla tillstånden utåt i **k**-rummet tills elektronerna tar slut. Vi har då fyllt alla tillstånd inom den så kallade Fermisfären, se figur 4.11. Vi hittar alltså metallens grundtillstånd (vid T = 0) genom att fylla de lägsta tillstånden nerifrån och upp.

På detta sätt blir alla tillstånd fyllda upp till den så kallade Fermienergin  $E_F$  med de N elektronerna. Vi beräknar Fermienergin genom att sätta antalet elektroner lika med antalet fyllda tillstånd (se fig. 4.13),

$$N = \int_0^{E_F} Z(E) dE = \frac{2}{3} C E_F^{3/2} = Z(E_F) \cdot \frac{2}{3} E_F$$
(4.10)

och får sedan

$$E_F = \frac{h^2}{2m} \left(\frac{3}{8\pi} \frac{N}{V}\right)^{2/3} = \frac{\hbar^2 k_F^2}{2m}$$
(4.11)

där  $k_F$ , kallat *Fermivågtalet*, är radien av Fermisfären som innehåller alla **k**-vektorer som motsvarar besatta tillstånd vid T = 0K,

$$k_F = \left(3\pi^2 \frac{N}{V}\right)^{1/3}$$

Kvoten N/V, dvs antal elektroner per volymsenhet, kommer vi ofta att skriva som elektronkoncentrationen n. Tillståndstätheten vid Fermienergin ges av (från ekvation 4.10)

$$Z(E_F) = \frac{3}{2} \frac{N}{E_F}$$



Figur 4.11: Fermisfären har radien  $k_F$  och innehåller vågvektorerna **k** hos de besatta tillstånden.



Figur 4.12: Vid T = 0 K besätter elektronerna alla tillstånd upp till Fermienergin  $E_F$ .



Figur 4.13: Antalet elektroner är lika med den markerade arean som bestäms av tillståndstätheten och Fermienergin.



Figur 4.14: När en elektron i ett inre atomärt skal i en av metallens atomer har exciterats med så stor energi att den har lämnat atomen, kommer det tomma tillståndet snart att fyllas, t ex av en valenselektron i elektrongasen. I övergången skapas en foton vars energi talar om för oss vilken energi elektronen hade före övergången. Skillnaden mellan den högsta och lägsta emitterade fotonenergin bör då motsvara Fermienergin och intensitetsspektrumet borde likna tillståndstätheten – den streckade linjen.

 $E_F$  ligger över alla fyllda tillstånd och under alla tomma tillstånd. Men eftersom tillstånden i vår makroskopiska lådpotential (eller metallbit) ligger så oerhört tätt kan vi betrakta Fermienergin som den kinetiska energin hos elektronen i det högsta fyllda tillståndet i metallens grundtillstånd vid T = 0 K. Eftersom V(x) = 0 i lådan är elektronernas energi enbart kinetisk.

Hastigheten (eller snarare farten)  $v_F$  hos en elektron vid Fermienergin, Fermihastigheten, kan beräknas ur

$$\frac{m}{2}v_F^2 = E_F$$

Hur kan vi nu verifiera vår bild av elektrongasen och experimentellt t ex bestämma Fermienergin? Ett sätt är att beskjuta metallen med högenergetiska elektroner eller fotoner som slår ut en elektron ur ett inre skal i en av metallens atomer. Det tömda tillståndet markeras med en vit cirkel i figur 4.14. Förr eller senare kommer en av valenselektronerna att falla ner i det tomma tillståndet och göra sig av med övergångsenergin i form av en foton. Fotonenergin beror på vilken energi elektronen hade innan övergången. Skillnaden mellan största och minsta emitterade fotonenergi bör då motsvara Fermienergin. Emissionsspektrumet visar – i natrium – en förvånansvärt god överensstämmelse med teorin. Man kan också definiera Fermitemperaturen  $T_F$  från  $E_F = kT_F$  och Fermivåglängden  $\lambda_F$  från  $k_F = 2\pi/\lambda_F$ . **Exempel:** Beräkna Fermienergin och Fermi-hastigheten i Na och jämför dina värden med en elektrons medelenergi och medelhastighet vid rumstemperatur (T = 300 K) enligt klassisk statistisk fysik.

$$n = \frac{N}{V} = \frac{N_{atom}}{V} = 2,65 \cdot 10^{28} \text{ m}^{-3}$$

$$E_F = \frac{h^2}{2m} \left(\frac{3}{8\pi}n\right)^{2/3} = 3,24 \text{ eV}$$

Hastigheten hos en elektron med energi  $E_F$  är

$$rac{m}{2}v_F^2 = E_F$$
 vilket ger  $v_F = 1.07\cdot 10^6 \mathrm{~m/s}$ 

I en klassisk gas är termiska energins (dvs elektronens rörelse<br/>energi) medelvärde  $\frac{3}{2}kT$  omkring 1/40 eV<br/>  $\approx 25$  meV vid T = 300 K. Den termiska hastigheten blir då

$$rac{m}{2}v_{th}^2=rac{3}{2}kT$$
 vilket ger  $v_{th}=1,17\cdot 10^5~{
m m/s}$ 

**Kommentarer**: Exemplet visar att Fermienergin är ca 100 gånger större än den klassiska termiska energin vid rumstemperatur vilket är typiskt för metaller då  $E_F \sim 5-10$  eV. Notera att Fermienergin och Fermihastigheten inte är temperaturberoende. Även vid absoluta nollpunkten har elektronerna i elektrongasen alltså hög kinetisk energi och en hög hastighet inom frielektronmodellen.

#### 4.6 Fermi-Dirac fördelningen

Vi har beskrivit metallens grundtillstånd vid T = 0, men vad händer om vi ökar temperaturen? Genom att t ex värma katoden i ett vakuumrör kan man öka antalet elektroner som emitterades. Vi förväntar oss alltså att elektronernas fördelning över tillstånden ändrar sig vid en uppvärmning. Elektroner kommer nu att kunna besätta tillstånd med högre energi.

Vid ändliga temperaturer, T > 0 K, kan elektroner alltså exciteras termiskt. Eftersom den termiska energin per partikel – av storleksordningen kT – är så liten i förhållande till  $E_F$ , kommer framför allt elektroner strax under  $E_F$  att kunna exciteras till tomma tillstånd strax ovanför  $E_F$ . Elektroner med initialt lägre energi (långt mindre än  $E_F$ ) har inga tomma tillstånd i närheten och kan då inte tillgodogöra sig den lilla termiska energi som står till förfogande.

Vi utgår från att varje Na-atom bidrar med en ledningselektron (Na finns i grupp 1 i det periodiska systemet).





Figur 4.15: a) Fermi-Diracs fördelningsfunktion vid T = 0 K samt för två högre temperaturer med Ferminivån markerad. b) Produkten  $F(E) \cdot Z(E)$  som anger tätheten av besatta tillstånd. Den markerade arean är lika med antalet elektroner.

Den funktion som håller reda på elektronernas fördelning över tillstånden vid en godtycklig temperatur kallas Fermi-Diracs fördelningsfunktion F(E), och ges av uttrycket

$$F(E) = \frac{1}{e^{(E - E_F(T))/kT} + 1}.$$
(4.12)

Detta är fördelningsfunktionen för kvantpartiklar som lyder *Pauliprincipen*. Denna princip förbjuder oss att besätta elektroniska tillstånd med fler än en elektron.

F(E) ger sannolikheten att ett tillstånd med energi E är besatt vid temperaturen T. Fermi-Dirac fördelningen kommer att vara användbar i kommande kapitel - den talar om för oss i vilka av de tillåtna tillstånden vi kan hitta elektroner.  $E_F(T)$  i F(E) kallas Ferminivå eller kemisk potential.  $E_F(T)$  är temperaturberoende. Fermienergin som vi betraktade i tidigare avsnitt är gränsvärdet av Ferminivån vid T = 0. I uttrycket (4.12) för F(E) skrivs ibland  $E_F$  istället för  $E_F(T)$ , men det är Ferminivån som avses.

För metaller är Ferminivåns temperaturberoende mycket svagt inom det temperaturintervall där metaller överhuvudtaget existerar. Vi kan alltså för metaller (men inte för halvledare som vi kommer till i senare kapitel) approximativt kan betrakta Ferminivån som konstant och ge den samma värde som Fermienergin.

Vid T = 0 K är Fermi-Dirac fördelningen en stegfunktion som funktion av energi (tunn linje i fig. 4.15a). Vid ändliga temperaturer blir Kvantpartiklar som lyder Pauliprincipen kallas fermioner och har halvtaligt spinn, t.ex. elektroner, till skillnad från bosoner som har heltaligt spinn och vars fördelning på tillgängliga tillstånd inte begränsas av Pauliprincipen.

Sannolikheten att ett tillstånd med energi  $E = E_F(T)$  är besatt är 1/2. steget "mjukare". Energiintervallet inom vilket F(E) går från nära 1 till nära 0 är av storleksordningen kT.

Vi ser i figur 4.15 att nästan alla elektroner är låsta i sina tillstånd alla tillstånd under  $E_F$  är besatta vid T = 0 K, och detta är fortfarande till stor del sant även om vi värmer elektrongasen. Endast få elektroner, de nära  $E_F$ , kan ändra sin energi vid en temperaturökning  $\Delta T$ . Systemets energi är nästan oförändrad vid en uppvärmning, i en första approximation är energiändringen  $\Delta E \approx 0$ , och därmed har vi hittat en kvalitativ förklaring till varför elektronernas bidrag till värmekapaciteten,

$$C_V = \frac{\Delta E}{\Delta T},$$

är så mycket mindre än vad en klassisk beskrivning hade gett. Med hjälp av lite kvantmekanik har vi tagit det första stora steget mot en förståelse av elektronstrukturen i fasta material.

**Exempel:** Betrakta valenselektronerna i Na inom fri-elektronmodellen. Vad är sannolikheten för att ett tillstånd vid E = 0 (lägsta möjliga energin i fri-elektronparabeln) i Na är tomt vid 300 K?

Fermi-Dirac fördelningen ger sannolikheten att tillståndet är besatt

$$F(E=0) = \frac{1}{e^{(0-E_F(T))/kT} + 1}$$

Natrium är en metall och vi antar att Ferminivån  $E_F(T)$  kan approximeras med Fermienergin, som beräknats till 3,24 eV. F(E = 0) blir då så nära 1 att både Matlab och miniräknare visar 1. Vi kan dock ta oss runt denna begränsning hos Matlab/miniräknare och direkt beräkna sannolikheten att tillståndet är obesatt:

$$1 - F(0) = 1 - \frac{1}{e^{(0 - E_F(T))/kT} + 1} = \frac{e^{(0 - E_F(T))/kT}}{e^{(0 - E_F(T))/kT} + 1}$$

Exponentialtermen är här mycket liten, och vi<br/> kan då stryka den i nämnaren så att 1 - F(0) blir

$$1 - F(0) = e^{(0 - E_F(T))/kT} \sim 10^{-55}$$

Sannolikheten är alltså utomordentlig låg att vi ska hitta ett tomt lägsta tillstånd (vid E=0) i natrium vid rumstemperatur. Den ökar visserligen snabbt då vi ökar T, men vid 1000 K (långt över natriums smältpunkt) är den fortfarande så liten som  $10^{-17}$ .

Om vi betraktar ett tillstånd närmare  $E_F$ , t.ex. vid E = 3 eV, är sannolikheten  $10^{-4}$  att ett sådant tillstånd är tomt vid 300 K. Elektroner även vid denna energi är alltså till god approximation opåverkade av en temperturhöjning från 0 till 300 K.

**Kommentar:** Uppvärmning förmår alltså bara påverka elektroners fördelning på tillstånd mycket nära  $E_F$  i metaller. Ofta använder man approximationen att elektroner med energier inom avstånd kT (eller ibland något mer, t ex 3kT) från  $E_F$  är de som kan ändra tillstånd vid en temperatur T.

Vi skall här inte försöka härleda Fermi-Dirac fördelningen, men det följande resonemanget visar i korta drag hur man i princip skulle kunna gå tillväga och hjälper oss att förstå vad som ligger bakom fördelningsfunktionen.

Låt oss betrakta ett system (se fig. 4.16) med åtta elektroner där de möjliga en-elektrontillstånden ligger på konstant avstånd i energi,  $\Delta$ , från varandra. I systemets grundtillstånd besätter elektronerna de 8 lägsta en-elektrontillstånden. Vi bortser här för enkelhets skull från elektronernas spinn, men beaktar Pauli-principen.

Grundtillståndets konfiguration visas längst till vänster i figur 4.16. Vi sätter systemets totala energi i grundtillståndet till 0. Exciterade konfigurationer skapas genom att en eller flera elektroner besätter högre liggande en-partikeltillstånd och de exciterade konfigurationerna får då högre energi än grundtillståndet. Exciterade konfigurationer är i figuren ordnade efter ökande systemenergi (åt höger i figuren).

Vi vill nu bestämma sannolikheten att ett visst en-elektrontillstånd är besatt vid en given temperatur. Vi vet att *Boltzmann-faktorn* 

$$P(E) \sim e^{-E/kT}$$

är proportionell mot sannolikheten att systemet har energin E vid temperaturen T. Därmed kan vi beräkna sannolikheten för varje konfiguration. Summan av alla sådana sannolikheter ska bli 1 - om vi betraktar alla möjliga konfigurationer kommer vi garanterat att finna att systemet är i någon av dem. Genom att normera med summan kan vi beräkna varje konfigurations sannolikhet. Sannolikheten att systemet befinner sig i en konfiguration med energi  $E_i$  ges således av

$$P(E_i) = \frac{e^{-E_i/kT}}{\sum_{n} e^{-E_n/kT}}$$
(4.13)

där vi i nämnaren summerar över systemets alla möjliga konfigurationer n med energier  $E_n$ . Med *en*-elektron tillstånd (eller orbital) menar vi ett kvantmekaniskt tillstånd för en enstaka elektron. Systemets konfiguration är en beskrivning av samtliga elektroners tillstånd vilket ger hela systemets tillstånd och energi.

Pauliprincipen kräver att vi betraktar hela systemet med sina åtta elektroner. En enskild elektrons möjligheter att besätta olika en-elektrontillstånd begränsas starkt av vilka tillstånd de övriga sju elektronerna besätter.



Figur 4.16: Elektronkonfigurationer ordnade efter total energi med systemets grundtillstånd längst till vänster.

Besättningsgraden för ett enstaka en-elektrontillstånd får vi genom att summera sannolikheterna för alla konfigurationer där detta enelektrontillstånd är besatt. Som exempel tittar vi på det 11:e enelektrontillståndet i figur 4.16 och bestämmer sannolikheten att detta en-elektrontillstånd är besatt för att belysa förfarandet.

Vi ser att en-elektrontillståndet nummer 11 är besatt en gång i en konfiguration med energi  $3\Delta$ , en gång i en konfiguration med energi  $4\Delta$ , två gånger i konfigurationer med energi  $5\Delta$ , och tre gånger i konfigurationer med energi  $6\Delta$ .

Sannolikheten att det 11e en-elektrontillståndet är besatt ges av

$$P(11) = \frac{e^{-\frac{3\Delta}{kT}} + e^{-\frac{4\Delta}{kT}} + 2 \cdot e^{-\frac{5\Delta}{kT}} + 3 \cdot e^{-\frac{6\Delta}{kT}}}{1 + e^{-\frac{\Delta}{kT}} + 2e^{-\frac{2\Delta}{kT}} + 3e^{-\frac{3\Delta}{kT}} + 5e^{-\frac{4\Delta}{kT}} + 7e^{-\frac{5\Delta}{kT}} + 11e^{-\frac{6\Delta}{kT}}}$$

enligt ekv. 4.13. Vi summerar alltså sannolikheterna för var och en av de konfigurationer som innebär att en-elektrontillstånd 11 är besatt - i nämnaren finns summan över alla möjliga tillstånd så att vi får normerade sannolikheter.

I uträkningen här har vi tagit med konfigurationer med energi upp till 6 $\Delta$ . Konfigurationer med högre energi blir alltmer osannolika och bidrar därmed allt mindre. Notera att t ex sannolikheten för var och en av konfigurationerna med energi 6 $\Delta$  är  $e^{-6\Delta/kT} \approx 2.5 \times 10^{-3}$  gånger mindre än för grundtillståndskonfigurationen.

Resultatet för de lägsta 14 orbitalerna visas i figur 4.17, där man har valt att sätta  $kT = \Delta$ . För denna specifika temperatur blir P(11) =0,0456. Konfigurationer med energier upp till  $6\Delta$  inkluderades i den numeriska beräkningen. Våra resultat verkar skilja sig något från

98



Figur 4.17: Sannolikheten för de 14 lägsta en-elektrontillstånden att vara besatta vid temperaturen  $T = \Delta/k$ . Som jämförelse är Fermi-Dirac fördelningen plottad som cirklar för  $E_F = 8.5\Delta$  vid samma temperatur.

Fermi-Diracs fördelningsfunktion, men i gränsen  $\Delta \to 0$  hade vi nått överensstämmelse.

### 4.7 Klassisk gräns hos en elektrongas

Vi har sett hur viktigt det är att beskriva elektronerna i en metall med hjälp av kvantmekanik. Å andra sidan vet vi också att klassisk mekanik inte är fel i sig, men att användningsområdet kan vara begränsat. Klassisk mekanik duger t ex fint för att beskriva en fotbolls rörelse under en match, vågegenskaperna hos partikeln är här inte ens observerbara. Inget interferensmönster uppstår när bollen passerar spalterna mellan försvarsspelarna.

När kan en elektrongas beskrivas klassiskt? Så länge Pauli-principen är viktig kan klassisk fysik inte duga. Men om sannolikheten att två elektroner skulle råka vilja besätta samma tillstånd är försumbart liten, då lär vi också kunna glömma Pauli-principen. Detta förutsätter i sin tur att sannolikheten att ett tillstånd är besatt av en elektron är mycket liten, dvs

$$F(E) = \frac{1}{e^{(E - E_F(T))/kT} + 1} \ll 1$$

för alla E som är tillåtna för elektroner. Detta kräver att

$$e^{(E-E_F(T))/kT} \gg 1.$$

Under dessa förutsättningar kan F(E) skrivas

$$F(E) = e^{-(E - E_F(T))/kT}$$
(4.14)

F(E) är en avtagande funktion som antar sitt största värde för systemets lägsta energi. Inom frielektronmodellen har vi valt energiskala så att lägsta tillståndet för valenselektronerna energi är E = 0. Genom att kräva att  $F(E) \ll 1$  för E = 0, blir detta sant även för alla högre energier. Vi når alltså den klassiska gränsen när

$$e^{E_F(T)/kT} \ll 1$$
 eller  $e^{-E_F(T)/kT} \gg 1$ .

Eftersom T > 0 måste Ferminivån  $E_F(T)$  vara negativ och ligger alltså under understa tillåtna tillståndet! Observera att vårt tidigare antagande att  $E_F(T) \approx E_F(T = 0)$ , dvs att Ferminivån hade endast ett svagt temperaturberoende, bara gällde för "normala" temperaturer!

Slusatsen så här långt är att om vi kräver att det ska vara låg sannolikhet för besättning för alla tillstånd i vårt system, så måste Ferminivån ligga under den lägsta energin i systemet.

Nu vill vi applicera vår diskussion på frielektrongasen och hitta ett kriterium som avgör vid vilken temperatur vi kan tänka på systemet som klassiskt. Detta kriterium får gärna involvera Fermienergin, dvs Ferminivån vid T = 0, som vi lätt kan beräkna. I den följande diskussionen betecknar vi Ferminivå med  $E_F(T)$  och Fermienergin, dvs gränsvärdet av Ferminivån då  $T \rightarrow 0$ , med  $E_F(T = 0)$ .

Vi sätter som tidigare (ekv. 4.10) antalet elektroner, N, lika med antalet besatta tillstånd. Vi fann ovan (ekv. 4.14) att F(E) kunde förenklas i situationer då  $F(E) \ll 1$  för alla E - vi vill hitta ett villkor som innebär att denna approximation gäller och använder nu den förenklade formen:

$$N = \int_{0}^{\infty} Z(E)F(E)dE = Ce^{E_F(T)/kT} \int_{0}^{\infty} \sqrt{E}e^{-E/kT}dE$$

 $Z(E)dE = C\sqrt{E}dE$ där

 $C = \frac{4\pi L^3 (2m)^{3/2}}{h^3}$ 

 $F(E) = e^{-(E - E_F(T))/kT}$  $\int_{-\infty}^{\infty} \sqrt{y} e^{-y} dy = \frac{\sqrt{\pi}}{2}$ 

$$\frac{N}{V} = e^{E_F(T)/kT} \frac{2}{h^3} (2\pi m kT)^{3/2}$$

Om  $e^{E_F(T)/kT} \ll 1$ gäller, dvs om vi<br/> kan bortse från Pauliprincipen, följer då att

$$n = \frac{N}{V} \ll \frac{2}{h^3} (2\pi m kT)^{3/2}.$$
(4.15)

Vi ser att temperaturen T måste vara hög, eller elektronkoncentrationen n låg, för att en klassisk beskrivning ska fungera - endast då



Figur 4.18: a) Kvantregim:  $kT \ll E_F(T=0)$ . Pauliprincipen styr hur elektronerna fördelas på tillstånden. b) Klassisk gräns. Temperaturen har ökat relativt i figur a), så att  $kT \gg E_F(T=0)$  Elektroner fördelas termiskt på tillstånden, och besättningsgraden är låg för alla tillstånd.

blir sannolikheten för besättning låg för varje tillstånd i systemet och endast då behöver vi inte beakta Pauli-principen.

Vi kan ur detta resultat härleda ett kriterium som Fermienergin

$$E_F(T=0) = \frac{h^2}{2m} \left(\frac{3}{8\pi} \frac{N}{V}\right)^{2/3}$$

måste uppfylla för att vi ska befinna oss i denna klassiska regim:

$$E_F(T=0) \ll \left(\frac{3}{4\pi}\right)^{2/3} \pi kT$$
 eller  $E_F(T=0) \ll kT$ .

Detta är villkoret för att kunna beskriva elektrongasen som en klassisk gas. Vi kan alltså genom att jämföra Fermienergin med den aktuella temperaturen avgöra om en kvantmekanisk beskrivning mha Pauliprincipen är nödvändig eller ej. Fermienergier i metaller ligger kring 5 eV och de temperaturer som är nödvändiga för att systemet ska kunna beskrivas klassiskt ligger då över 58 000 K! Metaller kräver alltså i regel en kvantmekanisk beskrivning.

Vi kan nu jämföra kvantregimen och den klassiska gränsen. Notera att vi med klassisk resp. kvantmekanisk här avser vad som styr hur elektroner fördelas på tillstånd - temperaturen eller Pauliprincipen. Kvantregim:  $kT \ll E_F(T=0)$ . Elektronerna trängs i de lägsta tillstånden, Pauli-principen är väsentlig och begränsar elektronernas möjligheter att besätta tillstånd. Vi måste beskriva elektronerna som en kvantmekanisk gas.

Vi kan beräkna medelenergin för elektronerna i gasen vid $T=0~{\rm K}$ som

$$E_{av} = \frac{\int_{0}^{\infty} EZ(E)F(E)dE}{\int_{0}^{\infty} Z(E)F(E)dE} = \frac{\int_{0}^{E_{F}} EZ(E)dE}{\int_{0}^{E_{F}} Z(E)dE} = \frac{3}{5}E_{F}$$
(4.16)

där vi nu måste använda fulla uttrycket för F(E). Elektronernas (kinetiska) energi bestäms alltså i gränsen  $kT \ll E_F(T=0)$  av vilket tillstånd inom fri-elektronmodellen de besätter. Temperatur påverkar fördelningen av elektroner mycket lite då, som vi sett tidigare, endast elektroner ganska nära  $E_F$  påverkas av en temperaturförändring.

Klassisk gräns:  $kT \gg E_F(T=0)$ . Är detta villkor uppfyllt vet vi att  $F(E) \ll 1$  för alla energier i systemet. Tillstånden i figurerna är desamma, men besättningen på dem skiljer sig markant. Betraktar vi ett specifikt tillstånd i den klassiska regimen är sannolikheten att det är besatt liten oavsett vilket tillstånd vi råkar välja. De besatta tillstånden ligger alltså mycket glest, elektronerna är inte i vägen för varandra när det gäller att besätta tillstånden.

Pauliprincipen har då ingen betydelse och elektronerna kan beskrivas som en klassisk gas där elektronerna är termiskt fördelade på de olika tillstånden. Deras medelenergi ges av  $\frac{3}{2}kT$  enligt ekvipartitionsteoremet (kapitel 2). Som bekräftelse på detta kan vi beräkna medelenergin på samma sätt som i ekv 4.16:

$$E_{av} = \frac{\int_{0}^{\infty} EZ(E)F(E)dE}{\int_{0}^{\infty} Z(E)F(E)dE} \approx \frac{\int_{0}^{\infty} E\sqrt{E}e^{-E/kT}dE}{\int_{0}^{\infty} \sqrt{E}e^{-E/kT}dE}$$
$$= \frac{\left[-(kT)e^{-E/kT}E^{3/2}\right]_{0}^{\infty} + \frac{3}{2}kT\int_{0}^{\infty} e^{-E/kT}E^{1/2}dE}{\int_{0}^{\infty} e^{-E/kT}E^{1/2}dE} = \frac{3}{2}kT$$

Approximation:  $F(E) \approx e^{-E/kT}$ i klassiska regimen.

partialintegration

Den klassiska gränsen krävde att

$$e^{E_F(T)/kT} = \frac{h^3}{2(2\pi m kT)^{3/2}} n \ll 1.$$

Vi ser att om partikeltätheten n är stor, som i en metall, krävs i princip mycket höga temperaturer för att nå den klassiska gränsen. Allmänt ser vi att gränsen även nås för låg partikeltäthet n och/eller för stor massa m. Elektrontätheten kan t ex i en halvledare vara mycket liten.

**Exempel:** Hur låg måste elektrontätheten vara för att elektrongasen skall kunna beskrivas klassiskt vid rumstemperatur?

$$T = 300 \text{ K:} n \ll \frac{2(2\pi m k T)^{3/2}}{h^3} = 2.5 \times 10^{25} \text{m}^{-3}$$

Elektronkoncentrationen i metaller är, som vi har sett, flera storleksordningar större än n ovan, och villkoret för klassisk gräns är därmed inte uppfyllt. Valenselektroner i metaller måste beskrivas i en kvantmekanisk modell. I halvledare kommer vi däremot att kunna ha elektronkoncentrationer som är låga nog för att en klassisk beskrivning ska vara rimlig, som vi kommer att se i kapitel 6. I den klassiska gränsen ser vi ingen skillnad mellan bosoner och fermioner, besättningsgraden av alla en-partikeltillstånd måste vara mycket liten. Då kan vi för elektronerna bortse från Pauliprincipen, två elektroner kommer aldrig att konkurrera om samma en-elektrontillstånd.

#### 4.8 Värmekapacitet

Vi såg i kapitel 2 att elektronernas bidrag till värmekapaciteten hos fasta ämnen inte kan beskrivas av klassisk fysik. Vi ska nu titta lite mer kvantitativt på elektronernas bidrag till värmekapaciteten med hjälp av frielektronmodellen, men fortfarande nöjer vi oss med approximativa resultat. I en elektrongas med N elektronernas ges deras totala kinetiska energi vid T = 0 av:

$$E(T=0) = N \cdot E_{av} = \frac{3}{5}NE_F$$

där medelenergin per elektron  $E_{av}$  beräknades i ekv. 4.16. Endast elektroner inom ett energi-intervall av storleksordningen kT under  $E_F$  kommer att kunna exciteras termiskt - se fig 4.19. Under excitationsprocessen är deras energiökning i medel av storleksordningen kT. Vi uppskattar då energin vid temperaturen T till

$$E(T) \sim E(T=0) + kT \cdot Z(E_F) \cdot kT$$

Därmed kan vi även uppskatta elektrongasens värmekapacitet:

$$C_V^{el} = \left(\frac{\partial E}{\partial T}\right)\Big|_V \sim 2Z(E_F) \cdot k^2 T = N \cdot 3\frac{k^2 T}{E_F}$$

där vi utnyttjat  $Z(E_F) = \frac{3}{2} \frac{N}{E_F}$  som visats tidigare.

Värmekapaciteten per elektron blir

$$c_V \sim 3 \frac{k^2 T}{E_F}$$

vilket överensstämmer väl med ett mer noggrannt beräknat resultat:

$$c_V = \frac{\pi^2}{2} \frac{k^2 T}{E_F}$$

vars härledning är betydligt svårare.

Vi har då

$$c_V = \frac{\pi^2}{2} \frac{k^2 T}{E_F} = \frac{\pi^2}{2} \cdot k \cdot \frac{kT}{E_F} \ll \frac{3}{2}k$$

eftersom  $kT/E_F \ll 1$  vid vanliga temperaturer ( $kT \approx \frac{1}{40}$  eV vid T = 300 K och typiskt  $E_F \sim 5$  eV!). En ledningselektron i en metall ger följaktligen ett bidrag till värmekapaciteten som är storleksordningen 100 gånger mindre än vad en klassisk uppskattning ger vid handen.



 $\sim kT$ 

F(E)

1

Figur 4.19: Endast elektroner inom ett intervall  $\sim kT$  nära  $E_F$ påverkas nämnvärt av en uppvärmning. Dessa elektroners energiökning uppskattar vi till i medel kT.

Klassiskt resultat enligt ekvipartitionsprincipen:  $c_v = \frac{3}{2}k$  per elektron, för Nelektroner blir  $C_V = \frac{3}{2}Nk$  och för en elektronkoncentration n blir  $C_V$  per volymsenhet  $\frac{3}{2}nk$ .



Figur 4.20: Elektrongasens totala (kinetiska) energi som först vid mycket höga temperaturer närmar sig det klassiska värdet.

Figur 4.20 visar den totala energin hos elektrongasen, E, som funktion av temperaturen. Först vid mycket höga temperaturer,  $kT \sim E_F$ , närmar sig E(T) det klassiska värdet

$$E_{klass}(T) = \frac{3}{2}NkT.$$

Observera att  $kT \sim E_F$  innebär att  $T \sim 58000$  K i metaller.

Vi har tidigare diskuterat i kapitel 1 och 2 att atomära svängningar bidrar till värmekapaciteten. När  $n_{atom}L^3$  atomer svänger i 3 dimensioner motsvarar det att vi har  $3n_{atom}L^3$  en-dimensionella oscillatorer. Så länge temperaturen är hög nog, kan vi bortse från kvantiseringseffekter och behandla oscillatorerna klassiskt. Enligt likafördelningslagen bidrar var och en av oscillatorerna i medel med kT till termiska energin,  $\frac{kT}{2}$  är potentiell energi,  $\frac{kT}{2}$  är kinetisk. Energin (per volymsenhet) hos atomernas svängningar blir

$$E_{atom} = 3n_{atom} \cdot kT$$

där  $n_{atom}$ är atomkoncentrationen. Atomernas bidrag till värmekapaciteten blir då

$$C_V^{atom} = \left(\frac{\partial E_{atom}}{\partial T}\right)\Big|_V = 3n_{atom}k.$$

Denna relation kom man fram till empiriskt långt innan Boltzmann hade utvecklat sina teorier och sambandet är känt som *Dulong-Petits lag* (1819). Som synes domineras värmekapaciteten av det atomära bidraget och valenselektronernas bidrag är i metaller litet.





Figur 4.21: I termisk jämvikt är elektronernas medelhastighet noll. Vid ett pålagt elektriskt fält accelereras elektronerna och rör sig i medel med drifthastigheten. Notera att i denna bild, från kapitel 2, då beskrev elektronerna som termiskt fördelade i hastighet. Inom frielektronmodellen styrs deras fördelning av temperatur eller av Pauliprincipen, se fig 4.22.

### 4.9 Konduktivitet - en andra gång

I den klassiska bilden av elektrongasen har elektroner slumpmässig hastighet och riktning och det termiska medelvärdet av elektronernas kinetiska energi ges av

$$E = \frac{3}{2}kT$$

och vi beskrev gasen i termisk jämvikt och i ett elektriskt fält som i figur 4.21.

Inom *frielektronmodellen* kan elektroner inte ha vilka hastigheter och riktningar som helst - de tillåtna tillstånden, dvs olika  $\mathbf{k}$ , får vi genom att lösa Schrödingerekvationen och vilka av dessa tillstånd som är besatta styrs av Pauliprincipen. Nu gäller

$$\mathbf{p} = m\mathbf{v} = \hbar\mathbf{k}$$

och rörelseekvationen är

$$m\frac{d\mathbf{v}}{dt} = \hbar \frac{d\mathbf{k}}{dt} = \mathbf{F} = -e\mathbf{E}.$$

Kraften leder till acceleration men alltså också till en ändring av vågvektorn enligt

$$\mathbf{k}(t) - \mathbf{k}(0) = -\frac{e\mathbf{E}t}{\hbar}$$

Som i kapitel 2 behöver vi anta någon spridningsmekanism som bromsar elektronernas acceleration, även om vi inom frielektronmodellen egentligen inte kan ha någon friktion. Elektron-elektronkollisioner kan visserligen förekomma men i en sådan process bevaras totala rörelsemängden och laddningstransporten påverkas inte.

Den följande diskussionen är mycket lik den tidigare analysen i kapitel 2 där vi betraktade elektronen som en partikel.

Kom ihåg  $\mathbf{v}(t) = \mathbf{v}(0) - \frac{e\mathbf{E}t}{m}.$ 

Men med tanke på kristallens joner, atomära svängningar eller defekter – som ju finns i en verklig metall – verkar det rimligt att återigen införa en kollisionstid  $\tau$  och vi får

$$\delta \mathbf{k} = -\frac{e\mathbf{E}\tau}{\hbar} = \frac{m}{\hbar}\mathbf{v}_d.$$

Alla elektroner ändrar sin vågvektor med  $\delta \mathbf{k}$ , dvs ändrar sin rörelsemängd med  $\hbar \delta \mathbf{k}$ . *Fermisfären* förskjuts i medel med  $\delta \mathbf{k}$  i kraftens riktning, i analogi med den klassiska bilden där alla hastigheter i medel ändrades med drifthastigheten  $\mathbf{v}_d$ . Med

$$m\mathbf{v}_d = \hbar\delta\mathbf{k} = -e\mathbf{E}\tau$$

får vi samma strömtäthet som tidigare,

$$\mathbf{J} = -ne\mathbf{v}_d = \frac{ne^2\tau}{m}\mathbf{E} = \sigma\mathbf{E}.$$

Med Fermi-hastigheten som vi beräknar ur Fermienergin

$$E_F = \frac{m}{2}v_F^2$$

kan vi nu bestämma elektronernas fria medelväglängd

$$\ell = v_F \cdot \tau.$$

Notera att fria medelväglängden i den kvantmekaniska beskrivningen beräknas med Fermihastigheten  $v_F$ ! Fermisfärens förskjutning är mycket liten i förhållande till Fermisfärens radie, allting händer nära Fermisfärens yta. En elektron sprids från ett tillstånd nära  $E_F$ till ett annat tillstånd nära  $E_F$ , se figur 4.23. En elektron längre in i Fermisfären behöver mycket energi för att hitta ett tillgängligt tomt tillstånd utanför Fermisfären och har alltså mycket liten möjlighet att spridas. Likaså kan en elektron vid Fermisfärens yta inte spridas till tillstånd längre in i Fermisfären, då dessa tillstånd med mycket stor sannolikhet redan är upptagna.

För koppar kan vi ur resistivitetsdata bestämma fria medelväglängderna

300 K: 
$$\tau \sim 10^{-14} \text{ s} \rightarrow \ell = v_F \cdot \tau \sim 10^{-8} \text{ m} > a$$
  
70 K:  $\tau \sim 10^{-13} \text{ s} \rightarrow \ell = v_F \cdot \tau \sim 10^{-7} \text{ m} \gg a$ 

Vi ser att fria medelväglängden kan vara bra mycket större än typiska atomavstånd a i kristallen, särskilt vid låga temperaturer. Det verkar alltså som att närvaron av jonerna, dvs den statiska periodiska potentialen, i sig inte leder till spridning.

Om nu den statiska periodiska kristallen inte sprider elektronerna måste vi identifiera andra spridningsmekanismer. Det visar sig



Figur 4.22: Fermisfären förskjuts ur sitt jämviktsläge med  $\delta \mathbf{k}$ , elektronerna får en drifthastighet  $\mathbf{v}_d$ .



Figur 4.23: Spridningsprocesser involverar framför allt elektroner nära Fermienergin.

ph: phonon eller fonon vibrationskvanta i kristallen. imp: impurity dvs orenhet eller defekt i kristallen.

107

att både atomernas termiska svängningar kring sina jämviktslägen  $(\tau_{ph})$  och defekter i kristallstrukturen  $(\tau_{imp})$  bidrar, det är alltså avvikelser från den periodiska potentialen som är viktiga här,

$$\frac{1}{\tau} = \frac{1}{\tau_{ph}} + \frac{1}{\tau_{imp}}$$

Experimentellt finner man

$$\rho = \rho(T) + \rho_0 \quad \text{(Matthiessens regel)}.$$

Eftersom energin hos atomära svängningar ökar med temperaturen,  $E_{atom} \sim T$ (se nedan) kan vi anta att

$$au_{ph} \sim \frac{1}{T}$$

och därmed blir  $\rho(T) \sim T!$  Mer om spridning i kapitel 6.

Slutsatsen, att en periodisk struktur inte leder till elektronspridning men att en störning av periodiciteten gör det, bekräftas när vi jämför resistivitetsdata hos legeringar och de ingående rena metallerna. Så hittar vi t ex resistiviteterna för

platina (Pt)  $\rho = 9.81 \cdot 10^{-8} \Omega m$  och för

rodium (Rh)  $\rho = 4.3 \cdot 10^{-8} \ \Omega m \text{ vid } T = 300 \text{ K}.$ 

En legering däremot som består av platina och 10% rodium har resistiviteten  $\rho = 18.7 \cdot 10^{-8} \Omega m!$ 

Varför adderas *inversa* kollisionstider för att bestämma den sammanlagda kollisionstiden?
# 4.10 Slutsatser

Inom frielektronmodellen kan vi förstå varför valenselektronernas bidrag till värmekapaciteten näst intill försvinner vid alla temperaturer under smältpunkten, dvs så länge metaller är kristaller.

Analysen av ledningsförmågan i metaller antyder att fria medelväglängden är mycket större än atomavståndet. Vi drar slutsatsen att den statiska strukturen inte sprider elektronerna.

Framför allt har vi med frielektronmodellen och lite kvantmekanik tagit ett första viktigt steg mot att förstå elektronstrukturen i fasta material. Det kvarstår dock fortfarande vissa stora problem.

*Metaller - Isolatorer*: Olika material uppvisar förvånansvärt stora skillnader i ledningsförmåga.

Al  $(1s^22s^2p^63s^2p)$  med tre valenselektroner är en metall, Si  $(1s^22s^2p^63s^2p^2)$  med ytterligare en valenselektron är en halvledare (isolator). Borde inte Si leda ännu bättre?

Ej heller kan vi ännu förklara

**Halleffekten:** I vissa fall är  $R_H > 0$  vilket tyder på att laddningsbärarna är positivt laddade?!

# 4.11 Tillståndstäthet igen - nu med fixa randvillkor!

Tidigare härledde vi tillståndstätheten för tillstånd som uppfyller periodiska randvillkor. Vi hoppas förstås att vårt val av randvillkor inte har alltför allvarlig inverkan på de resultat vi får - tillståndstätheten är ju vad den är i ett verkligt prov av t.ex. natrium, oavsett vad vi hittar på med avseende på randvillkor.

Vi gör om proceduren, men använder nu vågfunktioner som uppfyller fixa randvillkor, dvs tillstånden och egenenergierna i en oändligt djup tredimensionell kvantbrunn.

$$E = \frac{\hbar^2 \pi^2}{2mL^2} (n_x^2 + n_y^2 + n_z^2) = \frac{n^2}{K}$$

Tillståndstätheten, Z(E), är definierad som antalet tillstånd per energiintervall. Tre godtyckliga heltal  $(n_x, n_y, n_z)$ , dvs en punkt i figur 4.24 identifierar ett tillstånd med energi  $E = n^2/K$ .

Antalet tillstånd inom en sfär med radie n ges av sfärens volym

$$S(n) = \frac{4\pi}{3}n^3.$$

Eftersom  $E = n^2/K$  kan vi också skriva

$$S(E) = \frac{4\pi}{3} K^{3/2} E^{3/2}.$$

Detta är antalet tillstånd inom den heldragna sfären i figur 4.24, dvs antalet tillstånd med energier mellan 0 och E.

Tillståndstätheten ges sedan, som tidigare, av

$$Z(E) = \frac{dS(E)}{dE}$$

där dS(E)är antalet tillstånd inom det sfäriska skalet, mellan den heldragna och den streckade sfären i figur 4.24. Detta skal motsvarar energiintervallet dE.

Nu måste vi komma ihåg att för fixa randvillkor är bara positiva vågtal och positiva kvanttal relevanta. Vi får alla distinkta tillstånd om vi endast betraktar positiva n. Vi tar hänsyn till detta genom att bara räkna tillstånden i en kvadrant, dvs i 1/8 av sfären. Slutligen inkluderar vi även elektronens spinn genom en faktor 2. Varje vågvektor motsvarar ju två tillstånd, två elektroner kan ha samma vågvektor bara de har olika spinn. Vi får då

$$S(E) = \frac{\pi}{3} K^{3/2} E^{3/2}$$



Figur 4.24: Varje punkt  $(n_x, n_y, n_z)$  motsvarar ett tillstånd med energi  $E = n^2/K$ , eller två tillstånd om vi inkluderar spinn.

och tillståndstätheten blir

$$Z(E) = \frac{dS(E)}{dE} = \frac{\pi}{2} K^{3/2} E^{1/2} = \frac{L^3}{2\pi^2} (\frac{2m}{\hbar^2})^{3/2} \sqrt{E}$$

vilket är exakt samma resultat som vi erhöll genom att betrakta de energitillstånd vi fick då vi löste Schrödingerekvationen för periodiska randvillkor.

# 4.12 Utträdesarbete

Elektroner i en metall kan frigöras om vi belyser metallen med (vanligtvis ultraviolett) ljus (fotoelektriska effekten). Den emitterade elektronens maximala kinetiska energi ges av

 $E_{kin}^{max} = \hbar\omega - e\phi$ 

där  $e\phi$  är utträdesarbetet hos metallen.

Utträdesarbetet är den minsta energin som krävs för att frigöra en elektron från metallen. Vakuum-nivån är energin hos en elektron i vila som inte känner av metallens närvaro. En elektrons potentiella energi på grund av makroskopiska elektriska potentialer brukar man dock inkludera i vakuum-nivån.

# 4.13 Övergång mellan två metaller med olika utträdesarbeten

Ferminivån - eller kemiska potentialen - är en storhet som karaktäriserar ett system i termisk jämvikt. Två system är i termisk jämvikt med varandra om Ferminivåerna (och temperaturerna och trycken) är lika.

När två metaller med olika Ferminivåer sätts i kontakt med varandra, kommer elektroner att gå från metallen med högre  $E_F$  till den



Utträdesarbete: eng. work function



med lägre  $E_F$ , från vänster till höger i figuren. Detta resulterar i en positiv nettoladdning nära ytan av den vänstra metallen, och en motsvarande negativ laddning nära ytan av den högra metallen.

Dessa laddningar leder till ett elektriskt fält som pekar åt höger, dvs en kraft som verkar på elektronerna riktad åt vänster. I termisk jämvikt förhindrar denna kraft ett ytterligare elektronflöde genom gränsytan. Dessa ytladdningar leder - som i en vanlig plattkondensator - till en potentialdifferens mellan de två metallerna som förskjuter Ferminivåerna så att  $E_{F1}$  blir lika med  $E_{F2}$  samt förskjuter även vakuumnivån. Potential-differensen kallas kontaktpotential.

# 4.14 Termisk elektronemission



Om elektroner har tillräcklig kinetisk energi på grund av sin termiska rörelse kan de lämna metallen. Men de måste förstås också röra sig i rätt riktning, dvs mot ytan.

En elektron kan ta sig över ytbarriären om

$$\frac{p_x^0}{2m} > \frac{p_{x0}^0}{2m} = E_F + e\phi,$$

dvs om dess rörelsemängd i x-riktningen är större än  $p_{x0}$  som bestäms av metallens Fermienergi  $E_F$  och utträdesarbete  $e\phi$ . Detaljerna av analysen återfinns i ett appendix, här nämner vi bara slutresultatet, nämligen att elektronernas flöde ut ur metallen ger upphov till en strömtäthet

$$J = -\frac{4\pi emk^2}{h^3}T^2e^{-e\phi/kT}$$

Denna strömtäthet är starkt temperaturberoende. Vill vi t ex i ett vakuumrör uppnå en viss elektronström mellan katod och anod anger sambandet hur mycket vi behöver värma katoden, dvs vilken temperatur som krävs.

# 114 KAPITEL 4. FRIELEKTRONMODELLEN

# Kapitel 5

# Bandstruktur

Fri-elektronmodellen förmår inte skilja på isolatorer och ledare. Vi behöver modifiera vår modell och ta hänsyn till kristallens periodiska struktur: vi introducerar nästan-fri-elektronmodellen. Vi studerar således effekten av en periodisk potential på elektrontillstånden. Vi finner att planvågor för vissa vågvektorer kan reflekteras av potentialen. Det kan då bildas stående vågor vars sannolikhetstäthet varierar i rummet vilket leder till olika potentiell energi hos tillstånden. Bandgap uppstår. De flesta vågor anpassar sig till den nya situationen och beskriver elektroner som med konstant grupphastighet rör sig genom kristallen. Vi får band med tillåtna tillstånd. Rörelseekvationen anger hur elektronernas vågvektor ändrar sig under inverkan av en yttre kraft. Bandstrukturen ger oss nu möjlighet att skilja på metaller och isolatorer. Kan ett yttre fält åstadkomma en förändring i elektronernas fördelning på tillstånden, så att fler rör sig åt ett håll än åt ett annat, är kristallen en metall. Åstadkommer ett fält ingen förändring har vi en isolator. Detta betyder att metaller har delvis fyllda band, isolatorers band är helt fyllda eller tomma.

# 5.1 Frielektronmodellen (FEM) - en påminnelse

FEM är utgångspunkten för den följande diskussionen av elektroner i en kristall. Den är då vår nollte approximation till "verkligheten" vilket vi markerar med ett superskript "0".

$$\mathbf{H}^{0}\phi^{0} = -\frac{\hbar^{2}}{2m}\frac{d^{2}\phi^{0}}{dx^{2}} = E^{0}\phi^{0}$$

Vi väljer att använda periodiska randvillkor och lösningarna blir

$$\phi^0(x) = \frac{1}{\sqrt{L}} e^{ikx} \qquad \int_L |\phi^0(x)|^2 dx = 1 \qquad E^0 = \frac{\hbar^2 k^2}{2m}$$

som är normerade över kristallens längd L. Periodiska randvillkor ger oss även de tillåtna vågvektorerna.

$$k = \frac{2\pi}{L}n$$



Figur 5.1: Frielektronparabeln



Figur 5.2: En liten Na-metall med tre atomer. Atomära potentialer visas som streckade tunna linjer. Den fulla potentialen erhålls genom superposition av atomära potentialer och visas som en tunn heldragen linje. FEM motsvarar den tjocka streckade potentialbrunnen. En periodisk *modellpotential* visas som en tjock heldragen kurva.

Vågvektorn k (tillsammans med spinnet) specificerar entydigt ett elektroniskt tillstånd. Dessa lösningar är samtidigt egenfunktioner till rörelsemängdsoperatorn,

$$\mathbf{p}\phi^0 = -i\hbar \frac{d\phi^0}{dx} = \hbar k\phi^0, \quad \hbar k = mv = p$$

I varje tillstånd har både energin och rörelsemängden ett entydigt värde och läget måste då vara obestämt.

### 5.2 Introduktion till nästan-frielektronmodellen

När vi nu vill försöka förbättra vår modell, dvs gå längre än frielektronmodellen för att t ex förstå skillnaden mellan metaller och isolatorer, ligger det nära till hands att ta hänsyn till att elektronerna i en kristall faktiskt känner av jonerna i kristallen, vilket vi helt har ignorerat inom FEM. De rör sig då i en potential och i en kristall är denna potential *periodisk*, se figur 5.2.

Inom frielektronmodellen skrev vi lösningarna på formen  $\phi^0 \sim e^{ikx}$ . För dessa funktioner är sannolikhetstätheten  $|\phi^0(x)|^2$  konstant, dvs sannolikheten att hitta elektronen i ett visst intervall  $\Delta x$  är för alla tillstånd oberoende av x.

I närvaro av en periodisk potential förväntar vi oss att sannolikhetstätheten modifieras så att den på något sätt återspeglar kristallens periodiska struktur, dv<br/>s $|\phi^0(x)|^2 \neq \text{konst.}$ Bara en enstaka planvåg kan då ej längre vara en lösning till Schrödingerek<br/>vationen.

För att diskutera bindningen i kristaller brukar man vanligtvis, särskilt inom kemin, utgå från lokala interatomära bindningskrafter.

För en periodisk V(x) kommer elektroner i vissa tillstånd att föredra "dalarna" i potentialen och undviker potentialtopparna, medan elektroner i andra tillstånd föredrar att vara "på toppen" som vi kommer att se nedan. Man talar då t ex om kovalent bindning eller jonbindning. I den kovalenta bindningen tänker man sig att grannatomer delar på par av valenselektroner och därmed uppnår ädelgaskonfiguration som i diamant eller kisel. I jonbindning, å andra sidan, förutsätter man att en eller flera elektroner har gått över från en atom till en annan, som t ex i NaCl, och att jonerna sedan binds ihop av elektrostatiska Coulombkrafter. I bägge fallen beror bindningen väsentligen på den lokala fördelningen av valenselektronerna i rummet, dvs valenselektronernas sannolikhetstäthet uppvisar kraftiga variationer som ligger helt utanför fri-elektronmodellen.

Innan vi behandlar den formella lösningen till Schrödingerekvationen som inkluderar en periodisk potential, försöker vi med lite intuition komma fram till vad vi borde förvänta oss. Låt oss utgå från följande situation. En elektron befinner sig i ett tillstånd vars vågfunktion är en plan våg – som i frielektronmodellen. Men nu låter vi elektronen röra sig i en periodisk potential. En liknande situation möter vi när en ljusstråle träffar ett optiskt gitter. Om ljusets våglängd är av samma storleksordning som gitterkonstanten, dvs avståndet mellan spalterna, kan strålen brytas av gittret, vi observerar *interferens*. Låt oss nu titta lite närmare på dessa optiska interferenser innan vi återgår till elektronen.

#### 5.2.1 Röntgeninterferenser

När röntgenstrålningen hade upptäckts och man ville påvisa strålningens vågkaraktär – och i så fall även bestämma dess våglängd – behövde man ett gitter med gitterkonstant  $a \sim 1$  Å = 0,1 nm. M von Laue föreslog då 1912 att man borde kunna använda kristaller som gitter. Hade man väl bestämt våglängden med hjälp av en känd kristallstruktur, kunde strålningen sedan användas för att studera andra, ännu ej kartlagda kristallstrukturer, en metod som far och son Bragg framgångsrikt utvecklade. Denna röntgeninterferens används fortfarande för strukturbestämningar.

I en enkel bild tänker vi oss att atomerna i en kristall ligger i plan som i figur 5.3. Vi antar att varje plan – som en halvtransparent spegel – reflekterar en liten del av strålen, den större delen släpps igenom. För att de reflekterade strålarna från första och andra planet skall interferera konstruktivt, dvs efter reflektionen vara i fas med varandra, måste deras vägskillnad vara ett helt antal våglängder. Detta formuleras som

Braggs lag:  $2a\sin\theta = n\lambda$ 

Notera att detta krav är mycket selektivt, för en godtycklig infallsvinkel – och givna a och  $\lambda$  – blir det i allmänhet destruktiv



I härledningen av Braggs lag och när vi betraktar optiska gitter antar vi att spridningen är *elastisk*, dvs att strålningens våglängd inte ändras. Innan partiklarna (fotoner, elektroner, ...) träffar kristallen har de en viss energi och vi studerar de partiklar som efter spridningen lämnar kristallen med samma energi.



Figur 5.3: Vägskillnaden  $2a\sin\theta$ mellan de två reflekterade strålarna måste vara lika med ett helt antal våglängder för konstruktiv interferens.

interferens mellan de reflekterade strålarna. Man brukar då t ex se till att kunna variera  $\theta$ , antingen genom att rotera kristallen eller att ha provet i pulverform där kornen är slumpartat orienterade så att interferensvillkoret är uppfyllt för vissa kristallina korn. Vid normalt infall,  $\theta = 90^{\circ}$ , ger Braggs lag  $2a = n\lambda$ ,

$$\lambda = \frac{2a}{n}$$
 eller  $k = \frac{2\pi}{\lambda} = \frac{\pi}{a}n.$  (5.1)

Röntgenstrålningen tränger in djupt i kristallen, många reflekterade vågor interfererar, diffraktionsmönstret blir mycket skarpt. Men även partiklar, som ju uppvisar vågegenskaper, kan Bragg-reflekteras, exempelvis elektroner och neutroner.

Vi noterar sammanfattningsvis att en våg i ett periodiskt medium kan spridas, dvs det uppstår interferensmönster, och att denna spridning styrs av strikta krav eller urvalsregler.

#### 5.2.2 Tillbaka till elektronen

Om nu en elektron reflekteras betyder det att den ändrar sin vågvektor, t ex  $\mathbf{k} \to \mathbf{k}'$  som i figur 5.3. Lägger vi en elektron i ett tillstånd som beskrivs av en viss planvåg och låter den röra sig i en periodisk potential så kan vi senare även hitta den i andra planvågstillstånd. En enstaka planvåg kan därmed inte vara en stationär lösning.

Vi förenklar problemet till att vara en-dimensionellt. Vi tänker oss då att planvågor rör sig i en en-dimensionell kristall, dvs ett pärlband av atomer med avstånd *a* från varandra. De planvågor som kan undergå stark Braggspridning är de vars våglängd, och därmed vågtal, uppfyller Braggvillkoret för normalt infall (ekv. 5.1). Vi har då t.ex. planvågorna

 $e^{i(\pi/a)x}$  och  $e^{-i(\pi/a)x}$ 

som Braggreflekteras av potentialen från atomerna. En elektron som initialt är i första (andra) vågfunktionen kan återfinnas i andra (första) vågfunktionen. Vi vet inte längre i vilket tillstånd den är och vi får anta att elektronen befinner sig i en superposition av de två olika tillstånden.

Initialt hade vi två tillstånd (de två planvågorna) och när vi nu tänker oss att potentialen *mixar* eller *kopplar ihop* de två planvågorna så måste vi fortfarande ha lika många tillstånd - vi tar inte bort eller lägger till tillstånd, vi bara blandar en uppsättning (två stycken) tillstånd till en ny uppsättning. Vi kan bygga två nya tillstånd av de två planvågorna:

$$\phi \sim e^{i(\pi/a)x} \pm e^{-i(\pi/a)x} \sim \begin{cases} \cos(\frac{\pi}{a}x) \\ \sin(\frac{\pi}{a}x) \end{cases}$$

Elektroner används främst för ytstudier - de kan inte tränga in så djupt i kristallen eftersom de växelverkar kraftigt med kristallens laddningar. Dessa två nya vågfunktioner är stående vågor, till skillnad från de gående planvågor vi startade med.

För att ta reda på var i kristallen elektroner som besätter de två olika tillstånden finns, får vi hitta sannolikhetstätheterna  $|\phi|^2$ . Sannolikhetstätheterna tillsammans med potentialen visas i figur 5.4. Vi ser att en elektron i det ena tillståndet (fig A) har hög sannolikhet att befinna sig över topparna i potentialen, medan det andra tillståndet (fig B) innebär stor sannolikhet att elektronen befinner sig vid potentialens dalgångar. Vi kan då ana att dessa två tillstånd får olika potentiell energi. I frielektronparabeln bör det då öppnas "gap" vid de vågtal som ges av Braggvillkoret (ekv. 5.1).

Vi kan resonera oss fram till samma slutsats på ett alternativt sätt, utan att hänvisa till Braggreflektion.

Utan någon periodisk potential är alla lösningar av formen

 $\phi \sim Ae^{ikx} + Be^{-ikx}$ , eller  $\phi \sim C\sin(kx) + D\cos(kx)$ 



Figur 5.4: Sannolikhetstätheten för de nya tillstånden samt potentialen. I figur A och B markerar de tunna streckade linjerna  $\phi$  medan de heldragna markerar  $|\phi|^2$ .

likvärdiga lösningar till Schrödingerekvationen med samma energi. Vi väljer vanligtvis exponentiallösningarna med (A,B) = (1,0) eller (0,1) som ju beskriver elektroner som befinner sig överallt med samma sannolikhet. Vågfunktionerna ovan representerar dock vågfunktioner som i allmänhet är periodiskt varierande. Utan potential verkar det rimligt att energin är oförändrad även om sannolikhetstätheten varierar periodiskt i rummet.

Om vi nu lägger till en periodisk potential så kan vi fortfarande anta att vågfunktionernas energi är opåverkad för vågfunktioner med sådana k-värden att vågfunktionernas periodicitet inte överensstämmer med perioden hos potentialen, eftersom elektronerna för vissa xlokaliseras vid en potentialgrop och för andra x vid en potentialtopp och att det totalt inte blir något bidrag till den potentiella energin. Enbart vid de k-värden där periodiciteten i sinus- och cosinusfunktionerna överensstämmer med den hos potentialen får vi överallt en situation som i figur 5.4, där cosinusfunktionen motsvarar lokalisering kring potentialgropar medan sinusfunktionen motsvarar lokalisering kring potentialtoppar. Detta är precis då  $k = n\pi/a$ .

I de följande avsnitten ska vi fördjupa denna kvalitativa diskussion.

## 5.3 Nästan-frielektronmodellen

#### 5.3.1 Antaganden och ansats

Vi nöjer oss här med att betrakta problemet i en dimension, vid behov kan vi alltid extrapolera till tre dimensioner. Vi ska nu försöka få en uppfattning om *vilka tillstånd och energier* som finns i vårt system då vi har en periodisk potential. Senare ska vi diskutera vilka av dessa tillstånd som är besatta av elektroner.

Kristallstrukturen introducerar en periodisk potential

$$V(x) = V(x+a)$$

där a är avståndet mellan atomerna. Schrödingerekvationen i en dimension lyder:

$$\left\{-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V(x)\right\}\phi(x) = E\phi(x)$$

För att studera vilken effekt en periodisk potential har på elektronerna i en kristall, betraktar vi nu en enkel modellpotential i en dimension:

$$V(x) = -V_0 \cos\left(\frac{2\pi}{a}x\right) \qquad (V_0 > 0)$$

Vi tänker oss att potentialgroparna representerar de attraktiva positiva jonerna som ligger på avståndet a ifrån varandra. Med  $N_a$ atomer blir kristallens längd

$$L = N_a a$$

och vi förutsätter periodiska randvillkor när vi löser Schrödingerekvationen.

Denna modellpotential är "snäll" och uppvisar t ex inte Coulombpotentialens singulära beteende nära kärnorna. Men vår kvalitativa analys kommer ändå att låta oss förstå de väsentligaste aspekterna hos elektronstrukturen i fasta material.

Vi förutsätter i det följande att potentialen är "svag", dvs att den motsvarar en störning som endast svagt förändrar FEM-resultaten. Därför utgår vi ifrån elektroniska tillstånd i form av plana vågor och studerar hur dessa modifieras av den periodiska potentialen, dvs vi betraktar *nästan-frielektronmodellen*.

Vi tänker oss följande situation. En elektron beskrivs av vågfunktionen

$$\phi_1^0(x) = \frac{1}{\sqrt{L}} e^{ik_1x}$$





För modellpotentialen gäller  $\int_0^a V(x) dx = 0.$ 



Figur 5.5: Tillstånden  $\phi_1^0(x)$  och

Hamiltonoperatorn med periodisk

väldefinierad. Detta innebär att

en elektron kan spridas från ett fri-elektrontillstånd till ett annat

även om tillståndens fri-elektron-

inkluderar en periodisk potential, endast är parametrar utan omedelbar fysikalisk betydelse –

 $\phi_2^0(x)$  är inga egentillstånd till

potential, deras energi är ej

energier – som nu, när vi

är olika.

som är en lösning till frielektronfallet men som inte är en egenfunktion i närvaro av en periodisk potential. Vi räknar nu med att denna elektron kommer att kunna Bragg-spridas av potentialen. Låt den efter spridningen t ex beskrivas av vågfunktionen

$$\phi_2^0(x) = \frac{1}{\sqrt{L}} e^{ik_2x}$$

Därefter kan den mycket väl spridas tillbaka till  $\phi_1^0(x)$  (eller något annat tillstånd). För enkelhets skull begränsar vi analysen här till dessa två vågfunktioner.

Vågfunktionerna  $\phi_1^0(x)$  och  $\phi_2^0(x)$  beskriver två möjliga icke-stationära tillstånd för vår elektron, tillstånden kopplas av potentialen, dvs elektronen kan spridas mellan tillstånden. En stationär lösning bör omfatta bägge möjligheter, dvs på något sätt innehålla bägge våg-funktioner.

Vi tar hänsyn till kopplingen genom att göra ansatsen

$$\phi(x) = c_1 \phi_1^0(x) + c_2 \phi_2^0(x) \tag{5.2}$$

där  $\phi(x)$  är en approximativ lösning till den tidsobero<br/>ende Schrödingerekvationen

$$\mathbf{H}\phi = E\phi. \tag{5.3}$$

Hamiltonoperatorn innehåller nu, förutom den kinetiska energin, även den periodiska potentialen V(x),

$$\mathbf{H} = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V(x) = \mathbf{H}^{\mathbf{0}} + V(x).$$

Operatorn  $\mathbf{H}^{\mathbf{0}}$ , som inom FEM ger elektronens totala energi;

$$\mathbf{H}^{\mathbf{0}}\phi_{1(2)}^{0} = E_{1(2)}^{0}\phi_{1(2)}^{0},$$

ger nu endast elektronens kinetiska energi. Då vi har  $V(x) \neq 0$  finns ju möjligheten att elektronen också har potentiell energi, till skillnad från vad den kan ha inom fri-elektronmodellen. Totalenergin kan nu alltså skilja sig från den kinetiska energin.

#### 5.3.2 Lösning av SE med periodisk potential

En lösning till vårt problem är funnen när vi har bestämt koefficienterna  $c_1$  och  $c_2$  i ekv. 5.2 samt de tillhörande energierna E i Schrödingerekvationen 5.3.

För att ta fram ett ekvationssystem för koefficienterna  $c_1$  och  $c_2$  sätter vi in ansatsen i Schrödingerekvationen;

$$\mathbf{H}(c_1\phi_1^0 + c_2\phi_2^0) = E(c_1\phi_1^0 + c_2\phi_2^0).$$

Sedan multiplicerar vi från vänster med  $\phi_1^{0^*}$  (komplexkonjugat) och integrerar över kristallens längd L.

$$\int_0^L \phi_1^{0^*} \mathbf{H}(c_1 \phi_1^0 + c_2 \phi_2^0) dx = \int_0^L \phi_1^{0^*} E(c_1 \phi_1^0 + c_2 \phi_2^0) dx$$

Ekvationens vänstra sida blir

$$\int_{0}^{L} \phi_{1}^{0^{*}} \mathbf{H}(c_{1}\phi_{1}^{0} + c_{2}\phi_{2}^{0}) dx = c_{1} \underbrace{\int_{0}^{L} \phi_{1}^{0^{*}} \mathbf{H}\phi_{1}^{0} dx}_{H_{11}} + c_{2} \underbrace{\int_{0}^{L} \phi_{1}^{0^{*}} \mathbf{H}\phi_{2}^{0} dx}_{H_{12}}_{H_{12}}$$
$$= H_{11}c_{1} + H_{12}c_{2}$$

och ekvationens högra sida blir

$$\int_0^L \phi_1^{0^*} E(c_1 \phi_1^0 + c_2 \phi_2^0) dx = Ec_1.$$

Här utnyttjade vi att vågfunktionerna är normerade

$$\int_{0}^{L}\phi_{1}^{0^{*}}\phi_{1}^{0}dx=1$$

och att två vågfunktioner är ortogonala

$$\int_0^L \phi_1^{0^*} \phi_2^0 dx = \frac{1}{L} \int_0^L e^{-i(k_1 - k_2)x} dx = \begin{cases} 1 \text{ om } k_1 = k_2 \\ 0 \text{ om } k_1 \neq k_2 \end{cases}$$

Därmed har vi hittat den första ekvation som vi behöver för att bestämma  $c_1$  och  $c_2$ .

$$H_{11}c_1 + H_{12}c_2 = Ec_1$$

En andra ekvation för att bestämma koefficienterna får vi när vi upprepar stegen ovan, dvs multiplicerar Schrödingerekvationen med  $\phi_2^{0^*}$  och integrerar över kristallens längd. Detta ger

$$H_{21}c_1 + H_{22}c_2 = Ec_2$$

ten  $H_{ij}$  som är integraler.

Matriselementet  $H_{11}$  beräknar vi enligt

Vi sammanfattar dessa två ekvationer på matrisform

$$\begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = E \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$$
(5.4)

Innan vi kan lösa detta egenvärdesproblem, dv<br/>s bestämma egenvärdenaEoch egenvektorerna<br/>  ${\bf c},$ måste vi först beräkna matriselemen-

$$\mathbf{c} = \left(\begin{array}{c} c_1 \\ c_2 \end{array}\right)$$

Här utnyttjar vi att

$$\int_0^a V(x)dx = \int_0^L V(x)dx = 0$$

för den potential vi har valt.

Vi förutsätter att vågfunktionerna uppfyller periodiska randvillkor, dvs

$$k_i = \frac{2\pi}{L} n_i.$$

Integralen över kristallens längd L blir med  $n_1 \neq n_2$ 

$$\int_{0}^{L} e^{-i(k_{1}-k_{2})x} dx =$$
$$= \int_{0}^{L} e^{-i\frac{2\pi}{L}(n_{1}-n_{2})x} dx =$$
$$= \frac{e^{-i\frac{2\pi}{L}(n_{1}-n_{2})L} - 1}{-i\frac{2\pi}{L}(n_{1}-n_{2})} = 0$$

$$H_{11} = \int_0^L \phi_1^{0^*} \mathbf{H} \phi_1^0 dx = \frac{1}{L} \int_0^L e^{-ik_1 x} \mathbf{H} e^{ik_1 x} dx$$
$$= \frac{1}{L} \int_0^L e^{-ik_1 x} \left\{ \frac{\hbar^2 k_1^2}{2m} + V(x) \right\} e^{ik_1 x} dx = \frac{\hbar^2 k_1^2}{2m} = E_1^0$$

och på samma sätt får vi

$$\begin{aligned} H_{22} &= \int_0^L \phi_2^{0^*} \mathbf{H} \phi_2^0 dx = \frac{1}{L} \int_0^L e^{-ik_2 x} \mathbf{H} e^{ik_2 x} dx \\ &= \frac{\hbar^2 k_2^2}{2m} = E_2^0 \end{aligned}$$

Diagonalelementen  $H_{11}$  och  $H_{22}$  är inget annat än tillståndens frielektronenergier. Vi kan betrakta dessa energier som elektronernas ostörda energier som "störs" av en periodisk potential. Vi kallar vårt tillvägagångssätt för störningsräkning eller störningsteori (*perturbation theory*).

I integralen

$$H_{12} = \int_0^L \phi_1^{0^*} \mathbf{H} \phi_2^0 dx = \frac{1}{L} \int_0^L e^{-ik_1 x} V(x) e^{ik_2 x} dx$$

skriver vi lämpligen om potentialen på formen

$$V(x) = -V_0 \cos\left(\frac{2\pi}{a}x\right) = \frac{-V_0}{2} \left(e^{i\frac{2\pi}{a}x} + e^{-i\frac{2\pi}{a}x}\right)$$

och får

$$H_{12} = -\frac{V_0}{2} \frac{1}{L} \int_0^L e^{-ik_1 x} \left( e^{i\frac{2\pi}{a}x} + e^{-i\frac{2\pi}{a}x} \right) e^{ik_2 x} dx =$$
  
=  $-\frac{V_0}{2} \frac{1}{L} \left( \int_0^L e^{-i(k_1 - \frac{2\pi}{a} - k_2)x} dx + \int_0^L e^{-i(k_1 + \frac{2\pi}{a} - k_2)x} dx \right) =$   
=  $\begin{cases} -\frac{V_0}{2} \text{ om } k_1 = k_2 \pm \frac{2\pi}{a} \\ 0 \text{ annars} \end{cases}$ 

Slutligen får vi också

$$H_{21} = \frac{1}{L} \int_0^L e^{-ik_2 x} V(x) e^{ik_1 x} dx =$$
  
=  $\left(\frac{1}{L} \int_0^L e^{-ik_1 x} V(x) e^{ik_2 x} dx\right)^* = H_{12}^* = H_{12}.$ 

Resultatet ovan, att  $H_{21} = H_{12} \neq 0$  endast om  $k_1$  och  $k_2$  skiljer sig med  $\frac{2\pi}{a}$ , betyder att endast under denna förutsättning kan potentialen koppla de två tillstånden. För en planvåg med ett givet  $k_1$  finns

Eftersom

$$\frac{d^2}{dx^2}e^{ikx} = -k^2e^{ikx}$$

försvinner termerna som motsvarar kinetiska energin i $H_{12}$  på grund av vågfunktionernas ortogonalitet.

$$\cos\alpha = \frac{1}{2}(e^{i\alpha} + e^{-i\alpha})$$

Integralen kan lösas på samma sätt som ortogonalitetsintegralen ovan.



Figur 5.6: Två tillstånd vars vågtal skiljer sig med  $\frac{2\pi}{a}$  kan kopplas av potentialen. Den streckade linjen är frielektronparablen.

Vi ser att om potentialen försvinner, dvs  $V_0 = 0$ , ges energiegenvärdena E av  $E_1^0$  och  $E_2^0$ , vi är tillbaka i frielektronfallet.

En trivial lösning vore

$$\mathbf{c} = \left(\begin{array}{c} 0\\ 0 \end{array}\right)$$

det alltså endast två, eller som vi ska se senare något fler, planvågor med vågtal  $k_2$  som vår planvåg kan mixas med till nya tillstånd. Potentialen är mycket selektiv i valet av tillstånd den kopplar!

Denna urvalsregel innebär en enorm förenkling av vårt problem. Antag att vi försöker oss på en allmän utveckling av den sökta vågfunktionen i fri-elektronvågfunktioner och börjar vår utveckling med vågvektorn  $k_1$ , då behöver vi inte ta med andra vågvektorer  $k_2$  än de som ges av

$$k_2 = k_1 \pm \frac{2\pi}{a}$$

Vi känner igen från vår inledande kvalitativa diskussion att frielektrontillstånd med  $k = \pi/a$  kopplades till  $k = -\pi/a$  av potentialen. Vi ser nu att detta första fall är en delmängd av resultatet ovan - inte bara  $k = \pm(\pi/a)$  utan alla par av frielektrontillstånd vars kvärden skiljer sig med  $2\pi/a$  kopplas till nya tillstånd i närvaro av potentialen.

Ekvationssystemet (5.4) kan nu slutligen skrivas som

$$\begin{pmatrix} E_1^0 & -V_0/2 \\ -V_0/2 & E_2^0 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = E \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$$
(5.5)

eller

$$\begin{pmatrix} E_1^0 - E & -V_0/2 \\ -V_0/2 & E_2^0 - E \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = 0$$

För att matrisproblemet ska ha icke-triviala lösningar – eller egenvektorer  $\mathbf{c}$  – krävs det att koefficientmatrisens determinant försvinner. Ur detta krav bestämmer vi då energiegenvärdena E och finner

$$E = \frac{E_1^0 + E_2^0}{2} \pm \sqrt{\frac{(E_1^0 - E_2^0)^2}{4} + \left(\frac{V_0}{2}\right)^2}.$$
(5.6)

#### 5.3.3 Resultat för olika val av k

Vi betraktar ett par olika fall. Först tittar vi lite närmare på fallet där fri-elektronenergierna hos de två frielektrontillstånd vi låter blandas av potentialen är lika, dvs  $E_1^0 = E_2^0$ . Energierna E blir (genom insättning i 5.6)

$$E = E_1^0 \pm \frac{V_0}{2} \qquad (V_0 > 0)$$

Vågvektorerna hos de två kopplade tillstånden är (se figur 5.7)

$$k_1 = \frac{\pi}{a}$$
 och  $k_2 = k_1 - \frac{2\pi}{a} = -\frac{\pi}{a}$ 

- något annat val av  $k_1$  och  $k_2$  finns inte om vi betraktar frielektrontillstånd med samma frielektronenergi. Iom dessa val av k betraktar vi alltså tillstånden

$$\phi(x) = c_1 e^{i\frac{\pi}{a}x} + c_2 e^{-i\frac{\pi}{a}x}$$

För den lägre energin,  $E_1 = E_1^0 - \frac{V_0}{2}$ , följer efter insättning i matrisproblemet

$$c_1 = c_2$$

och därmed vågfunktionen

$$\phi_1(x) = \frac{1}{\sqrt{2L}} \left( e^{i\frac{\pi}{a}x} + e^{-i\frac{\pi}{a}x} \right) = \sqrt{\frac{2}{L}} \cos\left(\frac{\pi}{a}x\right)$$

För den högre energin,  $E_2 = E_1^0 + \frac{V_0}{2}$ , får vi

$$c_2 = -c_1 \operatorname{och}$$

$$\phi_2(x) = \frac{1}{\sqrt{2L}} \left( e^{i\frac{\pi}{a}x} - e^{-i\frac{\pi}{a}x} \right) = \sqrt{\frac{2}{L}} i \sin\left(\frac{\pi}{a}x\right)$$

Båda vågfunktionerna är *stående vågor*. Faktorn som ger tidsberoendet,

$$e^{-i\frac{E}{\hbar}t},$$

kan inte ändra på det faktum att noderna ligger still.

Våra ursprungliga fri-elektrontillstånd ( $\phi_1^0 \operatorname{och} \phi_2^0$ ) var egenfunktioner till rörelsemängdsoperatorn och hade rörelsemängden  $p = \hbar k$ . Hur är det med våra nya approximativa lösningar,  $\phi_1 \operatorname{och} \phi_2$ ? Låter vi rörelsemängdsoperatorn verka på de lösningar vi fann i närvaro av en periodisk potential får vi t ex för funktionen  $\phi_1$ 

$$\mathbf{p}\phi_1 = -i\hbar \frac{d\phi_1}{dx} = \hbar k i \sqrt{\frac{2}{L}} \sin kx = \hbar k \phi_2 \qquad \left(k = \frac{\pi}{a}\right)$$

Vågfunktionerna  $\phi_1$  och  $\phi_2$  är således inte egenfunktioner till **p**. En mätning av rörelsemängden ger då en fördelning av resulterande värden, vars *medelvärde* ges av väntevärdet.

Väntevärdet av operatorn  $\mathbf{p}$  i tillståndet  $\phi_1$  blir

$$\int_0^L \phi_1^* \mathbf{p} \phi_1 dx = \frac{2}{L} \int_0^L \cos(\frac{\pi}{a}x) \mathbf{p} \cos(\frac{\pi}{a}x) dx$$
$$= i\hbar k \frac{2}{L} \int_0^L \cos(\frac{\pi}{a}x) \sin(\frac{\pi}{a}x) dx = 0$$

Vi kallar  $\int_0^L \phi^* \mathbf{A} \phi dx$  väntevärdet av operatorn  $\mathbf{A}$  i tillståndet  $\phi$ .



Figur 5.7: Tillstånden med vågvektorer  $\frac{\pi}{a}$  och  $-\frac{\pi}{a}$  kopplas av potentialen. Bägge tillstånden har samma frielektronenergi  $\frac{\hbar^2 k^2}{2m}$  med  $k = \frac{\pi}{a}$ .



Figur 5.8: Sannolikhetstätheten för våra två lösningar (delfigurer A och B) samt modellpotentialen.

Integranden är periodisk med period a:

$$\frac{1}{L} \int_0^L \cos^2\left(\frac{2\pi}{a}x\right) dx =$$
$$\frac{1}{a} \int_0^a \cos^2\left(\frac{2\pi}{a}x\right) dx$$



Figur 5.9: I närvaro av en periodisk potential plottar vi energierna vid den vågvektor ksom ligger i Brillouin-zonen.

vilket överensstämmer med det faktum att  $\phi_1$  (och  $\phi_2$ ) är stående vågor.

Tittar vi nu närmare på våra lösningar så ser vi att de två vågfunktionerna  $\phi_1$  och  $\phi_2$  har olika sannolikhetstätheter.

Vågfunktionen  $\phi_2 = i \sqrt{\frac{2}{L}} \sin kx$  (streckad linje i figur 5.8 A) ger en sannolikhetstäthet (heldragen linje) som är stor mellan atompositionerna i 0 och  $\pm a$ , dvs en elektron i detta tillstånd kommer med stor sannolikhet att hittas nära *potentialtopparna*.

Vågfunktionen  $\phi_1 = \sqrt{\frac{2}{L}} \cos kx$  (streckad linje i figur 5.8 B) å andra sidan leder till en sannolikhetstäthet som är stor nära de (attraktiva) atomära positionerna, dvs en elektron i detta tillstånd kommer med stor sannolikhet att hittas i *potentialens dalar*.

De två tillstånden  $\phi_1$  och  $\phi_2$  kommer därmed att ha olika potentiell energi i den periodiska potentialen vilket förklarar tillståndens energiskillnad. Denna energiskillnad kallar vi *bandgap*.

För att verifiera vårt resonemang om potentiell energi hos de två tillstånden kan vi beräkna väntevärdet för den potentiella energin

$$\bar{E}_{pot} = \int_0^L \phi^* V(x) \phi dx.$$

För tillstånd  $\phi_1$  får vi

$$\bar{E}_{pot} = \int_0^L \phi_1^* V(x) \phi_1 dx = -\frac{2V_0}{L} \int_0^L \cos^2\left(\frac{\pi}{a}x\right) \cos\left(\frac{2\pi}{a}x\right) dx =$$
$$= -\frac{V_0}{L} \int_0^L \left[1 + \cos\left(\frac{2\pi}{a}x\right)\right] \cos\left(\frac{2\pi}{a}x\right) dx =$$
$$= -\frac{V_0}{L} \int_0^L \cos^2\left(\frac{2\pi}{a}x\right) dx = -\frac{V_0}{2}$$

och för  $\phi_2$  på motsvarande sätt  $\bar{E}_{pot} = +\frac{V_0}{2}$ . Detta bekräftar att det är tillståndens potentiella energi som skiljer dem åt.

Vid vilket k ska vi nu plotta våra energier? Vi kan plotta energierna antingen vid  $k = \frac{\pi}{a}$  eller vid  $k = -\frac{\pi}{a}$  eftersom de två vågfunktionerna,  $\phi_1$  och  $\phi_2$ , som hör till energierna  $E_1$  och  $E_2$ , båda är linjärkombinationer av planvågor med dessa vågvektorer,

$$\phi_1(x) = \frac{1}{\sqrt{2L}} \left( e^{i\frac{\pi}{a}x} + e^{-i\frac{\pi}{a}x} \right) \qquad \text{och}$$
$$\phi_2(x) = \frac{1}{\sqrt{2L}} \left( e^{i\frac{\pi}{a}x} - e^{-i\frac{\pi}{a}x} \right).$$

Alltså är tillstånden inte entydigt karakteriserade av en viss våg-

#### 5.3. NÄSTAN-FRIELEKTRONMODELLEN

vektor som tillstånden var inom FEM.

Vi väljer nu att plotta energierna vid  $k = \frac{\pi}{a}$ , dvs vi väljer vågvektorn som ska representera våra tillstånd att ligga i intervallet

$$-\frac{\pi}{a} < k \le \frac{\pi}{a}$$

Detta intervall (som i tre dimensioner blir en volym) kallar vi Brillouinzon. Intervallet är halvöppet, punkterna  $\frac{\pi}{a}$  och  $-\frac{\pi}{a}$  är ju ekvivalenta, vi inkluderar bara en av dem i Brillouin-zonen.

Nu tittar vi på ett annat fall, som vi väljer sådant att fri-elektronenergierna hos de två kopplade tillstånden är olika,

$$E_1^0 \neq E_2^0.$$

De kan inte vara vad som helst - de måste vara energierna från frielektrontillstånd sådana att dessas k-värden skiljer sig med  $2\pi/a$ . Vi löser alltså Schrödinger-ekvationen för ett k en bit ifrån Brillouinzonkanten vid  $\frac{\pi}{a}$ , se figur 5.10.

Energierna E bestäms av det uttryck vi fann ovan (ekv. 5.6):

$$E = \frac{E_1^0 + E_2^0}{2} \pm \sqrt{\frac{(E_1^0 - E_2^0)^2}{4} + \left(\frac{V_0}{2}\right)^2}$$

 $\operatorname{med}$ 

$$E_1^0 = \frac{\hbar^2 k^2}{2m}$$
 och  $E_2^0 = \frac{\hbar^2 \left(k - \frac{2\pi}{a}\right)^2}{2m}$ .

Vi plottar lösningarna  $E_1$  och  $E_2$  som i figur 5.11. Vi saknar återigen ett entydigt k då de nya lösningarna innehåller såväl k som  $k - \frac{2\pi}{a}$ . Vi gör då samma val som tidigare och placerar båda våra energiegenvärden  $E_1$  och  $E_2$  vid det av våra två k-värden som ligger inom Brillouin-zonen - se figur 5.11. Slutligen efter upprepade räkningar för olika k inom Brillouin-zonen får vi de heldragna kurvorna - bandstrukturen.

#### 5.3.4 Val av vågvektorer

Vi har nu löst Schrödinger-ekvationen för vår enkla modellpotential där vi utgått från ansatsen

$$\phi = c_1 \phi_1^0(x) + c_2 \phi_2^0(x)$$

för vågfunktionen, med

$$\phi_1^0(x) = \frac{1}{\sqrt{L}} e^{ik_1x}$$
 och  $\phi_2^0(x) = \frac{1}{\sqrt{L}} e^{ik_2x}$ 



Figur 5.10: Två fri-elektrontillstånd med olika FEM energier  $E_1^0$  och  $E_2^0$  kopplas av den periodiska potentialen.



Figur 5.11: Bandstrukturen uppvisar bandgap. De heldragna linjerna visar vår resulterande bandstruktur som vi valt att placera helt inom Brillouinzonen - vi ska strax ytterligare diskutera detta sätt att rita E(k).

För en given vågvektor  $k_1$  fann vi att  $k_2$  måste uppfylla

$$k_2 = k_1 \pm \frac{2\pi}{a},$$

vi valde

$$k_2 = k_1 - \frac{2\pi}{a}$$

Därmed försummade vi alla övriga möjliga vågvektorer i vågfunktionens utveckling i plana vågor. Var detta det bästa valet? Är vår lösning trovärdig? Utan att gå in i allt för mycket detalj ska vi försöka besvara dessa frågor med utgångspunkt i analysen ovan.

Potentialen V(x) som kopplade fri-elektrontillstånden dök upp i matriselementet

$$H_{12} = \int_0^L \phi_1^{0^*} \mathbf{H} \phi_2^0 dx = \frac{1}{L} \int_0^L e^{-ik_1 x} V(x) e^{ik_2 x} dx = -\frac{V_0}{2}.$$

Om detta matriselement är noll (vilket det blir om  $k_1$  och  $k_2$  inte följer våra urvalsregler) kan ingen koppling ske.

Vidare påpekades i början av kapitlet att vi utgår från en svag och snäll potential så att frielektronmodellen fortfarande är en rimlig utgångspunkt. Har vi fri-elektrontillstånd som är degenererade eller nästan så, kan även en svag potential koppla dem starkt, vi såg t ex att plana vågor övergick i stående vågor vid Brillouin-zonkanten. Sådana i energi närliggande tillstånd bör vi då ta med i vår ansats, antalet termer i ansatsen bestämmer matrisproblemets dimension.

År däremot energiskillnaden mellan de frielektrontillstånd vi betraktar stor i förhållande till kopplingens matriselement, dvs

 $|E_1^0 - E_2^0| \gg |H_{12}|,$ 

då kan vi approximera vårt resultat (ekv. 5.6);

$$E = \frac{E_1^0 + E_2^0}{2} \pm \sqrt{\frac{(E_1^0 - E_2^0)^2}{4} + H_{12}^2}$$

 $\sqrt{1+x} \approx 1 + \frac{x}{2} \qquad (|x| \ll 1)$ 

på följande sätt ,

$$E = \frac{E_1^0 + E_2^0}{2} \pm \frac{E_1^0 - E_2^0}{2} \sqrt{1 + \frac{4H_{12}^2}{(E_1^0 - E_2^0)^2}}$$
$$E \approx \frac{E_1^0 + E_2^0}{2} \pm \frac{E_1^0 - E_2^0}{2} \left(1 + \frac{2H_{12}^2}{(E_1^0 - E_2^0)^2}\right)$$

och våra lösningar blir

$$E_1 \approx E_1^0 + \frac{H_{12}^2}{E_1^0 - E_2^0}$$
 och  $E_2 \approx E_2^0 - \frac{H_{12}^2}{E_1^0 - E_2^0}$ 

Vi ser att kopplingens effekt blir allt mindre med ökande skillnad i energi - se figur 5.12. Vid k "långt bort" från Brillouinzonskanterna ligger energiegenvärdena mycket nära fri-elektronenergierna, medan avvikelsen är större vid k nära  $\pm \pi/a$ .

#### 5.3.5 Allmän periodisk potential

I verkliga kristaller kommer inte potentialen att vara så enkel som vi har antagit hittills. Även om den är mer komplicerad kommer den dock att vara periodisk. En mera allmän periodisk potential med period a kan skrivas som en Fourierserie

$$V(x) = \sum_{p} V_{p} e^{i\frac{2\pi}{a}px} \qquad (p - \text{heltal})$$

Ur integralen

$$\int_{0}^{L} \phi_{1}^{*} V(x) \phi_{2} dx = = \frac{1}{L} \sum_{p} V_{p} \int_{0}^{L} e^{-ik_{1}x} e^{i\frac{2\pi}{a}px} e^{ik_{2}x} dx =$$
$$= \begin{cases} V_{p} & \text{om } k_{2} = k_{1} - \frac{2\pi}{a}p \\ 0 & \text{annars} \end{cases}$$

följer urvalsregeln

$$k_2 = k_1 - \frac{2\pi}{a}p$$
 (*p* - heltal).

Om vi tänker oss att elektronen befinner sig i ett tillstånd som beskrivs av en planvåg med vågvektor  $k_1$ , så kan den bara spridas till planvågstillstånd med en vågvektor  $k_2$  som skiljer sig från  $k_1$ med någon multipel av  $\frac{2\pi}{a}$ . Vi kan även formulera detta som att potentialen V(x) endast kopplar ihop planvågor vars k skiljer sig med någon multipel av  $\frac{2\pi}{a}$ . Detta innebär i sin tur att vår ansats för  $\phi(x)$  som ges av en linjärkombination av planvågor endast behöver innehålla dessa vågvektorer. Inom intervallet

$$-\frac{\pi}{a} < k \le \frac{\pi}{a}$$
 (Brillouin-zonen)

finns det lika många vågvektorer k som atomer i vår kristall – som vi kommer att se nedan. I stället för att behöva ta med alla dessa i en allmän utveckling av vågfunktionerna räcker det med en av dem!

De "nya" vågfunktionerna i kristallen är superpositioner av frielektronvågfunktioner (planvågor), som endast inkluderar frielektrontillstånd med k-värden separerade med någon multipel av  $2\pi/a$ . De nya tillstånden har alltså inte något entydigt bestämt k, och vi kan välja



Figur 5.12: Ligger tillstånden som potentialen kopplar långt ifrån varandra i energi, blir kopplingen knappt märkbar. Energierna blir då i stort sätt oförändrade, dvs ungefär lika med FEM energierna.



Figur 5.13: Vi kan välja att plotta energierna (i närvaro av den periodiska potentialen, heldragna linjer) så att de ligger nära fri-elektronparabeln (streckad linje) i figur (a) eller inom Brillouin-zonen (b). Pilarna indikerar vår procedur, dvs skifta alla energier sidledes med någon multipel av  $2\pi/a$ , för att samla bandstrukturen inom första Brillouinzonen.

att placera tillståndens energi vid vilket som helst k av de som ingår i summan, se figur 5.13. Konventionen, som vi redan berört, är att man i bandstrukturbilder associerar energin för det nya, blandade, tillståndet med det av de ingående k-värdena som ligger i första Brillouinzonen (se fig 5.13b)

En kontinuerlig kurva E(k) mellan  $-\frac{\pi}{a}$  och  $\frac{\pi}{a}$ , som beskriver ett intervall med tillåtna energier kallar vi ett *energiband*.

Liknande reflektioner som vid  $k = \frac{\pi}{a}$  (och därur följande stående vågor med olika potentiell energi) uppträder vid alla vågvektorer

$$k = \frac{2\pi}{\lambda} = \frac{\pi}{a}n$$
 (*n* - heltal).

här har vi bandgap.

Vi har nu bestämt bandstrukturen som uppstår i närvaro av en enkel endimensionell modellpotential. Att försöka beräkna bandstrukturen för en realistisk kristall är ett oerhört mycket mera omfattande projekt. Men förhoppningsvis har vår introduktion gett en tillräckligt god bakgrund för att våga närma sig de bandstrukturer som finns avbildade i litteraturen.

#### 5.3.6 Bandstuktur i 3D

Verkliga, tredimensionella, kristaller kan ha någon av flera olika kristallstrukturer. Samtliga består dock av något atomarrangemang



Figur 5.14: Bandstruktur i en en-dimensionell kristall.



Figur 5.15: a) Zinkblendestruktur med två olika atomslag (svarta och vita). Om alla atomer är av samma slag kallas strukturen diamantstruktur. b)-d) Kristallplan i några olika riktningar i kristallen. Avstånden mellan planen (periodiciteten) är olika i de olika riktningarna.

som upprepas många gånger för att bygga upp den makroskopiska kristallen.

Som exempel betraktar vi strukturen i figur 5.15(a). Denna är uppbyggd av två olika atomslag - vita eller svarta kulor i figuren. Detta specifika arrangemang av atomer kallas zinkblendestruktur. Är alla atomer i strukturen av samma slag kallas detta istället för diamantstruktur. Diamant (kol i diamantform), kisel och germanium har alla diamantstruktur, medan många binära (dvs byggda av två atomslag) halvledare såsom GaAs, InAs, InP, AlAs, GaP m.fl. kristalliserar i zinkblendestruktur. Många halvledare, däribland våra tekniskt viktigaste sådana, har alltså denna struktur, varför vi tar den som exempel.

Vi försöker nu indikera hur man kan utvidga vår bandstrukturdiskussion från en till tre dimensioner. Nu betraktar vi figur 5.15b)-d) och ser att vi kan hitta kristallplan, i vilka planvågor kan Braggspridas, i flera olika riktningar. I en en-dimensionell modell har vi bara en riktning till vårt förfogande, i tre dimensioner har vi många riktningar att beakta.

Utan att gå in på detaljer anar vi att vi kan hitta kristallplan som upprepas med viss periodicitet i flera olika riktningar. Periodiciteten kan vara olika i olika riktningar. Vi hittar t.ex. i figur 5.15 b) ett kortare avstånd  $a^*$  mellan de två markerade kristallplanen  $(a^* = a/2 \text{ om kubens sidolängd är } a)$  än vi gör i figur 5.15 d) (här blir avståndet  $a^{**} = a/\sqrt{3}$ ). Inom vår en-dimensionella modell är Brillouinzonskanterna viktiga - vid dessa  $k = \pm \pi/a$ , där a är potentialens/kristallens periodicitet, uppstår bandgap. Detsamma gäller i tre dimensioner. Brillouinzonen i k-rummet borde då bli större i riktningen vinkelrät mot planen i fig. b) än i fig. d), då avståndet mellan atomplanen är kortare i fig. b).

I vår endimensionella modell plottade vi bandstrukturen inom in-



Figur 5.16: a) Tredimensionell Brillouinzon för zinkblende/diamantstruktur. Konventionellt betraktar man i bilder över tredimensionella bandstrukturer elektronenergierna längs med de streckade linjerna i figuren. b) Frielektron-bandstruktur för zinkblende/diamantstruktur. Man har flyttat delar av frielektronparabeln in i första Brillouinzonen enligt samma princip som i en dimension - resultatet blir något mer komplicerat i 3D.

tervallet

$$-\frac{\pi}{a} < k \leq \frac{\pi}{a}$$

När vi nu extrapolerar vårt resonemang till tre dimensioner ersätts intervallet med en volym i tre dimensioner. Vi hittar alltså en tredimensionell Brillouinzon, figur 5.16 a), genom att i ett  $(k_x, k_y, k_z)$ rum markera de  $\pi/$ " periodicitet" som är relevanta i olika riktningar. Som exempel kommer Brillouinzonens gräns i  $k_z$ -riktningen att vara vid punkten  $(0,0,\pi/a^*)$ , där  $a^*$  är avståndet mellan två plan i zriktningen i figur 5.15 b).

I tre dimensioner är det inte möjligt att rita upp hela bandstrukturen för alla möjliga riktningar i **k**-rummet. I ett försök att få jämförbara visualiseringar av bandstrukturen för olika kristaller har man bestämt sig för en konvention där man betraktar bandstrukturen i vissa specifika riktningar, indikerade i figur 5.16 a). Som ses i figur 5.16 a) innebär  $\Gamma$ -punkten att man är i Brillouinzonens centrum,  $\mathbf{k} = (0,0,0)$ , och X-punkt(erna) att man betraktar sådana  $\mathbf{k}$  som ligger mellan  $\Gamma$  och X-punkterna, dvs längs en  $k_x$ ,  $k_y$  eller  $k_z$ -riktning.

Gör man en förflyttning av en tre-dimensionell fri-elektronparabel in i denna tre-dimensionella Brillouinzon, på samma sätt som vi tidigare gjorde i en dimension, får man bandstrukturen i figur 5.16 b). Här har ingen hänsyn tagits till effekten av en periodisk potential



Figur 5.17: Bandstrukturer för Si och GaAs. Notera att de är ganska lika varandra (dock med små men avgörande skillnader) samt att de även har tydliga drag av fri-elektronbanden i föregående figur.

- det finns ingen modulering av banden och inga bandgap har öppnats i strukturen. Vi kan tänka oss det som att man har låtit  $V_0$  i V(x) gå mot noll. Denna "empty-lattice-approximation" visar sig vara mycket lik verkliga bandstrukturer för ämnen som kristalliserar i zinkblende- eller diamantform - jämför med bandstrukturerna för GaAs och Si i figur 5.17. De (stora) skillnader i egenskaper som halvedare trots allt uppvisar beror på små olikheter i bandstrukturen. För att modellera dessa krävs dock mer komplicerade beskrivningar.

# 5.4 Rörelseekvationen för elektroner i en kristall

En viktig fråga med tanke på elektroniska tillämpningar återstår: hur beter sig elektronerna när vi t ex utsätter dem för ett elektriskt fält? Kan bandstrukturen förklara skillnaderna mellan metaller och isolatorer? Vi återgår nu till att huvudsakligen titta på en en-dimensionell modellkristall.

Elektronen betraktad som partikel bör kunna lokaliseras. Vi tänker oss därför ett vågpaket som utbreder sig med grupphastigheten

$$v_g = \frac{1}{\hbar} \frac{\partial E^0}{\partial k} = \frac{\hbar k}{m} = \frac{p}{m}.$$

Som i FEM har en elektron i varje tillstånd i bandstrukturen en

konstant grupphastighet:

$$v_g = \frac{\partial \omega}{\partial k} = \frac{1}{\hbar} \frac{\partial E}{\partial k}.$$

Inom FEM blir grupphastigheten

$$v_g = \frac{\hbar k}{m} = \frac{p}{m}.$$

Eftersom E inte längre är en enkel funktion av k kan frielektonresultatet ovan inte längre gälla. Vi ser att för elektroner i en periodisk potential gäller generellt

$$v_g \neq \frac{\hbar k}{m} \text{ och } p \neq \hbar k.$$

De stationära egenfunktionerna till Schrödingerekvationen med periodisk potential är inte längre egenfunktioner till rörelsemängdsoperatorn. Vågfunktionen har ju komponenter med olika vågvektorer. Vi måste beräkna väntevärdet av **p** för att få hastigheten. Denna hastighet är konstant för ett givet k men kan vara noll t ex vid  $k = \frac{\pi}{a}$ .

En lokaliserad elektron beskriver vi ju med ett vågpaket och ett vågpaket förflyttar sig med grupphastigheten  $v_g$ . Man kan visa att grupphastigheten i bandstrukturen är relaterad till rörelsemängdsoperatorns väntevärde och man finner:

$$v_g = \frac{1}{m} \left\langle \phi_k | \mathbf{p} | \phi_k \right\rangle = \frac{1}{\hbar} \frac{\partial E}{\partial k}$$

Man brukar istället kalla  $\hbar k \ kristallrörelsemängd$  eller – något mindre korrekt men inte så otympligt – kristallimpuls (crystal momentum) eftersom  $\hbar k$  ofta förekommer i urvalsregler som om det vore en verklig rörelsemängd. En elektron i ett stationärt tillstånd i bandstrukturen har en konstant hastighet – som kan vara noll – och den kommer inte att spridas av den periodiska potentialen, vars effekt ju redan är inkluderad i bandstrukturen.

Inom FEM hade vi rörelseekvationen

$$\hbar \frac{dk}{dt} = m \frac{dv}{dt} = \frac{dp}{dt} = F$$

där F är kraften som verkar på elektronen.

För bandstrukturen som uppstår på grund av den periodiska potentialen har vi inte längre en enkel relation mellan hastighet och vågvektor,  $\hbar k \neq m v_g$ , se figur 5.18. Vi får nu fram rörelse ekvationen genom att betrakta det arbete  $\delta E$  som en yttre kraft  $F_{ext}$ , t ex på grund av ett elektriskt fält  $\mathscr{E}$ , uträttar på elektronen under en förflyttning  $\delta s$ 

Elektroner som beskrivs av planavågor - som inte är lösningar tillSchrödingerekvationen när den inkluderar en periodisk potential - kan Braggspridas av potentialen.

Urvalsregler talar om för oss om en viss process är tillåten eller ej, vi kommer senare t ex att diskutera urvalsregler för optiska övergångar i halvledare.



hastigheten v(k).



Figur 5.19: Vid T = 0 besätter elektronerna energibanden nerifrån och upp. Blir då det senast besatta, högsta bandet bara delvis fyllt så är kristallen en metall. Är det senast besatta, högsta bandet helt fyllt (och näst högre band tomt) så är kristallen en isolator.

$$\delta E = F_{ext} \cdot \delta s = -e\mathscr{E}v_a \delta t.$$

Energiändringen kan relateras till en ändring av vågvektorn enligt

$$\delta E = \frac{dE}{dk} \delta k = \hbar v_g \delta k$$

och ur dessa två ekvationer följer

$$\hbar \delta k = -e\mathscr{E} \delta t = F_{ext} \delta t.$$

Därmed kommer vi fram till rörelseekvationen

$$\hbar \frac{dk}{dt} = -e\mathscr{E} = F_{ext}$$

Kraften som verkar på elektronen har nu delats upp i två delar. En del,  $F_{ext}$ , som vi lägger på utifrån, och en del som ges av alla krafter som elektroner utsätts för i kristallen, t ex attraktionen till jonerna. Dessa inre krafter har redan inkluderats när vi - åtminstone i princip - bestämde de stationära tillstånden i bandstrukturen.

# 5.5 Metaller - Isolatorer

Vi betraktar en en-dimensionell kristall med  $N_a$  atomer på inbördes avstånd a, vars längd då är

$$L = N_a a$$

Ett band motsvarar ett intervall med längden  $\frac{2\pi}{a}$  på k-axeln, som vi kallade Brillouin-zonen. Avståndet mellan de av periodiska randvillkor tillåtna vågvektorerna är  $\frac{2\pi}{L}$ . Antalet k-vektorer inom ett band är då

Periodiska randvillkor gav  $e^{ikx} = e^{ik(x+L)} \rightarrow e^{ikL} = 1$  $\Rightarrow kL = 2\pi \cdot p \quad k = \frac{2\pi}{L} \cdot p$ 

$$v_g = \frac{1}{\hbar} \frac{\partial E}{\partial k}$$



Figur 5.20: Ett elektriskt fält ligger över en metall och accelererar elektronerna som når ett stationärt tillstånd på grund av spridningsprocesser. I en isolator åstadkommer ett elektriskt fält ingen förändring.

$$\frac{2\pi}{a} / \frac{2\pi}{L} = \frac{L}{a} = N_a$$

Tar vi hänsyn till elektroners spinn, så har vi  $2N_a$  tillstånd i ett energiband.

Om varje atom bidrar med en elektron, har vi  $N_a$  elektroner och det lägsta bandet är halvfyllt. Ferminivån ligger då i det delvis fyllda bandet (se figur 5.19).

Med två elektroner per atom blir det lägsta bandet fyllt. Ferminivån ligger i bandgapet, över det fyllda bandet och under nästa ofyllda band.

Om vi nu lägger på ett elektriskt fält, så rör sig elektronerna i bandstrukturen enligt rörelseekvationen:

$$\hbar \frac{dk}{dt} = F_{ext}$$

Alla elektroner ändrar sin vågvektor i den yttre kraftens riktning. Elektronernas fördelning förskjuts som i figur 5.20 åt höger i kraftens riktning.

För ett delvis fyllt band resulterar detta i en nettoström eftersom fler elektroner nu har hastigheter i kraftens riktning än i kraftens motriktning. Systemet är en *metall. Spridning* på grund av avvikelser från periodicitet bromsar elektronernas acceleration som vi diskuterade i kapitel 4 och leder till ett stationärt tillstånd.

Om bandet är helt fyllt ändras elektronernas fördelning över tillstånden inte, den förblir densamma som i termisk jämvikt, vilket

Vi ser att det är elektronerna närmast  $E_F$  som bär strömmen. Det är dessa elektroner som sprids av vibrationer och defekter, därför använde vi  $v_F$  för att beräkna elektronernas fria medelväglängd enligt  $\ell = v_F \cdot \tau$ . innebär att inga strömmar kan flyta. Systemet är en isolator. Nu kan man ju rätteligen hävda att fördelningen även i en isolator förskjuts på samma sätt som i en metall. Men när en elektron når Brillouin-zonkanten vid  $k = \frac{\pi}{a}$  och verkar trilla ut ur Brillouin-zonen så dyker den ju genast upp vid  $-\frac{\pi}{a}$  igen, dessa två tillstånd var ju helt ekvivalenta. Därmed förblir ju fördelningen ändå oförändrad.

Vi drar slutsatsen:

Ett udda antal elektroner per atom  $\Rightarrow$  delvis fyllda band - metall.

Isolator: De lägsta banden är helt fyllda, högre band helt tomma  $\Rightarrow j \ddot{a} mnt$  antal elektroner per atom.

### 5.6 En en-dimensionell kedja av atomer

I detta avsnitt ska vi betrakta ett alternativ till nästan-frielektronmodellen för att närma oss elektronstrukturen i en kristall. Vi ska nu sammanföra N atomer till en linjär kedja och studera hur de atomära tillstånden kopplas till tillstånd som är delokaliserade över hela kristallen.

Vi söker lösningar till Schrödingerekvationen

$$\mathbf{H}\mathbf{\Phi} = E\mathbf{\Phi}$$

och utvecklar kedjans vågfunktioner i atomära orbitaler  $\phi_i$ , där orbitalen  $\phi_i$  är centrerad kring atom i.

$$\Phi = \sum_{i=1}^{N} K_i \phi_i$$

Vi antar att de atomära vågfunktionerna är reella, normerade och - återigen för enkelhetens skull - ortogonala:

,

$$\int \phi_i \phi_j dV = \delta_{ij}$$

Vi sätter in utvecklingen i Schrödingerekvationen

$$\sum_{i=1}^{N} K_i \mathbf{H} \phi_i = E \sum_{i=1}^{N} K_i \phi_i$$

multiplicerar med  $\phi_j^* = \phi_j$ från vänster, integrerar

$$\sum_{i=1}^{N} K_i \int \phi_j \mathbf{H} \phi_i dV = E \sum_{i=1}^{N} K_i \underbrace{\int \phi_j \phi_i dV}_{\delta_{ji}}$$

I tre dimensioner kan vi inte vända på argumentet. En metall kan ha ett jämnt antal elektroner per atom, och ett system med ett jämnt antal elektroner behöver inte vara en isolator.

I *en dimension* gäller dock även de omvända argumenten.

Vi försummar vågfunktionernas överlapp.

Kronecker delta:

$$\delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

Här utnyttjar vi att vi antagit att vågfunktionerna är ortogonala.



Figur 5.21: Med varje atomär position associerar vi en vågfunktion (eller orbital).

och får

$$\sum_{i=1}^{N} K_i \int \phi_j \mathbf{H} \phi_i dV = EK_j \qquad (j = 1, \dots, N)$$

Vi introducerar nu följande förkortningar för matriselementen (vi tar endast med kopplingen mellan närmsta grannar):

$$\int \phi_i \mathbf{H} \phi_i dV = E_1$$

$$\int \phi_i \mathbf{H} \phi_{i+1} dV = \int \phi_{i+1} \mathbf{H} \phi_i dV = -A \quad (A > 0)$$

vilket ger

$$EK_j = -K_{j-1}A + K_jE_j - K_{j+1}A$$
  $(j = 0, \dots N)$ 

För denna atomära kedja med N atomer har vi nu kommit fram till ett system av ekvationer ur vilket vi kan bestämma koefficienterna  $K_j$  och energiegenvärdena E när vi väl har bestämt oss för vilka randvillkor vi vill använda.

#### 5.6.1 Randvillkor: ändlig kedja - stående vågor

Vi kan lägga till två (imaginära) atomer 0 and N + 1 till vår kedja, vilket inte ändrar något om vi samtidigt kräver följande randvillkor

$$K_0 = K_{N+1} = 0.$$

Vi har då ekvationssystemet

För en ändlig kedja med Natomer bör termen med  $K_0$ utelämnas i den första ekvationen och termen med  $K_{N+1}$  i den sista.

 $E_1$  är den atomära orbitalens energi och A är den så kallade

hopping-parametern

Termer i boxar i ekvationssystemet försvinner med randvillkoren  $K_0 = K_{N+1} = 0.$ 



Figur 5.22: Kedja där vi har fixerat ändpunkterna - ändlig kedja.

Med detta knep får alla ekvationer samma form vilket gör lösningen mycket enklare. Vi utgår från ansatsen

 $K_j = \sin k x_j \qquad (x_j = j \cdot a)$ 

som vi sätter in i t ex den j:te ekvationen:

$$-AK_{j-1} + E_1K_j - AK_{j+1} = EK_j.$$

Vi får

$$-A\sin k(x_j - a) + E_1\sin kx_j - A\sin k(x_j + a) = E\sin kx_j$$

och därmed

$$-A \cdot 2\sin kx_i \cdot \cos ka + E_1 \sin kx_i = E\sin kx_i$$

eller slutligen

$$E = E_1 - A \cdot 2\cos ka \qquad (A > 0)$$

• Randvillkor:  $K_0 = K_{N+1} = 0$ 

Randvillkoret  $K_0 = 0$  uppfylldes redan av ansatsen, randvillkoret vid kedjans högra ände leder till

$$K_{N+1} = \sin kx_{N+1} = \sin k(N+1)a = 0$$
  
 $k = \frac{\pi}{(N+1)a}p = \frac{\pi}{L}p \qquad p = 1, \dots N$ 

där L = (N + 1)a. Vi väljer k att vara positiva, ett teckenbyte ger inget nytt tillstånd.

• Exempel: 5 atomer i fig 5.24

Ansatsen är vald att uppfylla randvillkoret  $K_0 = 0$  och leder till **stående vågor**.

Kom ihåg:  $\sin k(x_j \pm a) =$   $= sinkx_j \cdot \cos ka$  $\pm \cos kx_j \cdot \sin ka$ 



Figur 5.23: Bandet av energier E(k) för en ändlig kedja av atomer. Vågfunktionerna är stående vågor, vågtalen väljs så att k > 0.



Figur 5.24: A) Atomära vibrationer hos en kedja av 5 atomer, B) molekylorbitaler hos en kedja av 5 atomer, C) bandet E(k) med de 5 av randvillkoren tillåtna vågvektorerna. För två vågfunktioner har våglängderna angetts och motsvarande k-vektorer markerats i bandet E(k).



#### 5.6.2 Periodiska randvillkor - fortskridande vågor

Även i den atomära kedjan föredrar vi ofta att beskriva elektronerna med hjälp av fortskridande vågor. Vi tänker oss som tidigare att vi sluter kedjan till en ring, dvs lägger till termer som kopplar de atomära vågfunktionerna  $\phi_1$  och  $\phi_N$  och skriver



Figur 5.25: Två exempel på vågor som uppfyller periodiska randvillkor.



Vi söker nu lösningar som motsvarar fortskridande vågor och startar därför med ansatsen

$$K_s = K e^{ikx_s} \qquad (x_s = s \cdot a)$$

och lägger på periodiska randvillkor:

$$K_s = K_{s+N} \qquad (Na = L)$$

Därur följer

$$e^{iksa} = e^{ik(s+N)a}$$
 vilket ger  $e^{ikNa} = 1$ 

eller

$$kNa = 2\pi \cdot p$$
  $k = \frac{2\pi}{Na} \cdot p = \frac{2\pi}{L} \cdot p$ 

Avståndet mellan tillåtna vågvektorer är nu dubbelt så stort som var fallet för stående vågor, men å andra sidan har vi nu både positiva och negativa vågvektorer. Antalet k i ett band och därmed antalet tillstånd i ett band är oförändrat.

Vi sätter in vår ansats

$$K_s = K e^{ikx_s}$$

i t ex den j:te ekvationen:

$$-Ae^{ik(j-1)a} + E_1e^{ikja} - Ae^{ik(j+1)a} = Ee^{ikja}$$
$$-Ae^{-ika} + E_1 - Ae^{ika} = E$$
$$E = E_1 - 2A \cdot \cos ka$$

Energibanden i de bägge fallen har samma form! Med periodiska randvillkor blir vågfunktionerna *fortskridande*e vågor, tecknet av k bestämmer nu vågens utbredningsriktning. Med positivt k rör sig vågen i positiv x-led, med negativt k i motsatt riktning.

Ur ett atomärt tillstånd per atom har det bildats ett band av lika många tillstånd som vi har atomer.



Figur 5.26: Energibandet är halvfyllt om varje atom i kedjan bidrar med en elektron

Två vågvektorer som skiljer sig med  $\frac{2\pi}{a}$  är ekvivalenta, dvs de beskriver samma våg:

$$e^{i\left(k+\frac{2\pi}{a}\right)x_j} = e^{ikx_j} \cdot e^{i\frac{2\pi}{a}x_j} = e^{ikx_j} \cdot e^{i2\pi j} = e^{ikx_j}$$

där vi använt  $x_j = j \cdot a.$  Det räcker då att betrakta vågvektorer inom intervallet

$$-\frac{\pi}{a} < k \le \frac{\pi}{a},$$

som vi kallade Brillouin-zonen ovan. I detta intervall finns

$$\frac{2\pi}{\frac{a}{L}} = \frac{L}{a} = N_a$$
 tillåtna *k*-vektorer,

där  $N_a$  är antalet atomer i kedjan.

#### 5.6.3 Antal tillstånd per band

Vi har  $N_a$  tillåtna k-vektorer inom Brillouin-zonen.

Därmed finns det  $2N_a$  tillstånd i energibandet på grund av elektroners spinn.

Med en elektron per atom blir energibandet halvfyllt, och med två elektroner per atom blir energibandet helt fyllt.

#### Analogi: N atomer i en kedja i fig. 5.27



Figur 5.27: Vågkaraktären hos atomära svängningar liknar vågkaraktären hos elektroniska tillstånd.

# 5.7 Bundna tillstånd

Antag nu att det sitter en störatom i den linjära atomära kedjan på plats 0. "On-site" matriselementet i hamiltonianen skriver vi nu

$$\langle 0|\mathbf{H}|0\rangle = E_1 - F$$

och vi får ett "stört" ekvationssystem

$-AK_{-3}$	+	$E_1 K_{-2}$	_	$AK_{-1}$	=	$EK_{-2}$
$-AK_{-2}$	+	$E_1K_{-1}$	_	$AK_0$	=	$EK_{-1}$
$-AK_{-1}$	+	$(E_1 - F) K_0$	_	$AK_1$	=	$EK_0$
$-AK_0$	+	$E_1K_1$	_	$AK_2$	=	$EK_1$
$-AK_1$	+	$E_1K_2$	_	$AK_3$	=	$EK_2$

Vi räknar inte med att en störatom märkbart kommer att kunna ändra energibandet hos kristallen, men lokalt kring störatomen, där periodiciteten är bruten, kan vi få nya tillstånd.

. . .

Om t exF>0, dvs störatomens orbital har lägre energi, förväntar vi oss att en elektron kan bli infångad av störatomen och att ett nytt, lokaliserat tillstånd kan uppstå under bandet av tillstånd hos det ostörda periodiska systemet.



Figur 5.28: N kopplade atomära orbitaler ger upphov till Ntillstånd i ett band.

I halvledare ersätter man t ex enstaka kiselatomer med störatomer för att skräddarsy materialegenskaperna.

Detta egenvärdesproblem kan lösas för hand, här visar vi dock bara resultatet. F<0:elektroner undviker störatomen, ett lokaliser<br/>at tillstånd dyker upp $\ddot{o}ver$  bandet. Denna situation möter vi<br/> hos acceptorer.



F>0:elektroner attraheras till störatomen, ett bundet tillstånd dyker uppunder bandet. Denna situation förekommer hos donatorer.

I halvledare ger donatorer och acceptorer upphov till sådana störnivåer utanför energibanden i bandgapet.
#### 5.7. BUNDNA TILLSTÅND





Fig. 16.30. The overlap of s-orbitals gives rise to an s-band, and the overlap of p-orbitals gives rise to a p-band. In this case the s- and p-orbitals of the atoms are so widely spaced that there is a band gap. In many cases the separation is less, and the bands overlap.

Figur 5.29: Energiband från atomära s och p-orbitaler.

#### • s och p band

Varje atomär orbital genererar ett band.



Reduced Wave Vector

Figur 5.30: För att närma sig den verkliga bandstrukturen i en kristall kan vi antingen utgå ifrån atomära elektrontillstånd som växelverkar när atomerna sammanförs (vänster) eller från fria elektroner som påverkas av kristallens periodiska potential

# 5.8 Modeller för bandstruktur - sammanfattning

I nästan-frielektronmodellen startade vi från FEM – till höger i figur 5.30 – och betraktade effekten av en svag periodisk modellpotential. Spridning av elektronvågor ledde till energiband och bandgap. Vi gick åt vänster i figuren.

I vårt andra tillvägagångssätt startade vi från atomer – till vänster i figuren – som vi satte samman till en atomär kedja. Vi såg hur kopplingen mellan elektroniska tillstånd ledde till band av energier och bandgap. Vi gick åt höger.

Sanningen ligger, som alltid, någonstans däremellan.

# Kapitel 6

# Halvledare

En ren halvledare leder ström om ett tillräckligt antal elektroner har exciterats till ledningsbandet, vilket kan ske t ex genom uppvärmning eller belysning. Elektronerna i det nästan tomma ledningsbandet kan beskrivas som om de vore klassiska partiklar med en effektiv massa skild från vakuum-massan. De tömda tillstånden i det nästan fulla valensbandet kan vi betrakta som hål som beter sig som positivt laddade partiklar, även dessa med sin effektiva massa. Ett elektriskt fält kan förändra elektronernas och hålens fördelning på tillstånden i respektive band, därmed bidrar både elektronerna och hålen till strömmen. Dopning är det viktigaste sättet att förse halvledaren med laddningsbärare. Här ersätts en del av halvledarens atomer med främmande atomer som skiljer sig från halvledarens atomer i antal valenselektroner. Vid låga temperaturer är bundna till störatomerna, men redan vid rumstemperatur brukar störatomerna vara fullständigt joniserade, dvs laddningsbärarna är fria att röra sig i banden.

### 6.1 Elektroner i bandstrukturen

Bandstrukturen för kisel – framtagen med hjälp av experimentella data och teoretiska beräkningar – visas i figur 6.1. Elektronerna rör sig här i tre dimensioner och figuren visar energin som funktion av vågvektorns längd i några speciella riktningar i k-rummet. Bandstrukturen är betydligt mera komplicerad än i en dimension och vi ska inte gå in på några detaljer här – se kapitel 5 för en ytlig diskussion om tre-dimensionell bandstruktur.

Vid T = 0 är valensbanden i kisel (figur 6.1) mellan ungefär -12 eV och 0 helt fyllda och separerade från de högre liggande, tomma ledningsbanden genom ett bandgap  $E_g$ , se figur 6.2. Bandgapet  $E_g$ anger energiskillnaden mellan botten av ledningsbandet,  $E_C$ , och toppen av valensbandet,  $E_V$ , och brukar i halvledare vara någon eller några eV stort – i Si är  $E_g \approx 1,1$  eV. Energins nollnivå har i figur 6.1 lagts vid toppen av valensbandet, vilket egentligen är helt godtydligt men mycket vanligt i bandstrukturbilder. One should not work on semiconductors, they are a mess - who knows if they exist at all? Wolfgang Pauli, 1931



Figur 6.1: Till vänster: Kisels bandstruktur  $E(\mathbf{k})$  plottad i olika riktningar i **k**-rummet. L,  $\Gamma$ , X, U och K är speciella punkter i Brillouin-zonen. I mitten: Fermi-Dirac fördelningen F(E). Till höger: Brillouin-zonen är det område i **k**-rummet som i tre dimensioner motsvarar intervallet  $-\frac{\pi}{a} < k \leq \frac{\pi}{a}$  som vi mötte i en dimension.

Vi noterar här att i kisel ligger toppen av valensbandet vid en punkt i k-rummet ( $\Gamma$  i figur 6.1) och botten av ledningsbandet vid en annan punkt (ganska nära X i samma figur). Detta kallas för indirekt bandgap. Halvledare kan också ha direkt bandgap – se t.ex. bandstrukturen för GaAs i kapitel 5 eller figur 6.2.

Vid T = 0 är Fermi-Dirac fördelningen en stegfunktion, Ferminivån  $E_F$  ligger då i bandgapet, över  $E_V$  och under  $E_C$  – precis i mitten som vi kommer att se.

Vid högre temperatur kan elektroner exciteras från valensbanden till ledningsbanden. Detta visar även Fermi-Dirac fördelningen i figur 6.1 där sannolikheten att ett tillstånd i botten av ledningsbandet är besatt har ett ändligt om än mycket litet värde (sannolikheten var noll vid T = 0), och på motsvarande sätt har sannolikheten att ett tillstånd vid toppen av valensbandet är tomt antagit ett värde skilt från noll.

Ett tomt tillstånd i valensbandet innebär att elektronernas fördelning kan förändras av det elektriska fältet, det tomma tillståndet som vi kallar "hål" kan flyttas och därmed bidra till strömmen. Vid T = 0 K är halvledaren en isolator då ett elektriskt fält inte kan ändra på elektronernas fördelning över tillstånden. Det fulla



Figur 6.2: Förstoring av bandstrukturen runt bandgapet, dvs runt botten av ledningsbandet och toppen av valensbandet, hos en halvledare med direkt bandgap vid T = 0 K. Valensbandet är fullt, ledningsbandet är tomt.

(valens-)bandet, se fig. 6.2, medger ingen sådan förändring. (Ett mycket högt elektriskt fält,  $\mathscr{E} \sim 10^8 \text{ V/m}$ , skulle kunna excitera elektroner till ledningsbandet, en process som endast undantagsvis är av betydelse.)

Så länge halvledaren är i termisk jämvikt – eller inte har rubbats ur jämvikt alltför mycket av det elektriska fältet – kommer de flesta elektroner och hål att finnas nära respektive bands extrempunkt (gråa cirklar i figur 6.1) där enligt Fermi-Dirac fördelningen sannolikheten är störst att hitta elektroner och hål. Dessa områden och bandgapet däremellan är därför ofta den enda del av bandstrukturen vi behöver känna till.

För att få ett ungefärligt uttryck för bandens form E(k) kan vi, oavsett bandens exakta form, göra en Taylor-utveckling runt bandens extrempunkter. För botten av ledningsbandet ger Taylor-utvecklingen

$$E(k) = E_C + \frac{1}{2} \frac{\partial^2 E}{\partial k^2} \Big|_{k_0} (k - k_0)^2 \qquad \left( \frac{\partial E}{\partial k} \Big|_{k_0} = 0 \right)$$



Figur 6.3: I områden kring extrempunkter i bandstrukturen kan banden beskrivas av anpassade parabler.



Figur 6.4: Parabler har anpassats till bandstrukturen. Nära ett minimum är  $m^*$  positiv, nära ett maximum negativ!

där  $E_C$  och  $k_0$  betecknar energin resp. vågvektorn vid ledningsbandets botten. Endast den kvadratiska termen finns med – högre ordningens termer bortser vi från och den linjära termen försvinner när vi betraktar extrempunkten.

Kring extrempunkterna kan bandstrukturen alltså approximeras med lämpligt valda parabler (se figur 6.3), som då innehåller den information om elektronstrukturen vi behöver för att t ex beräkna tillståndstäthet eller studera laddningstransport. Approximationen är relevant så länge vår förutsättning, att elektroner och hål befinner sig nära extrempunkterna, är uppfylld.

Genom att definiera

$$\frac{1}{m_e^*} = \frac{1}{\hbar^2} \frac{\partial^2 E}{\partial k^2} \tag{6.1}$$

kan vi teckna ledningsbandets parabel som

$$E(k) = E_C + \frac{\hbar^2 (k - k_0)^2}{2m_e^*}.$$
(6.2)

 $m_e^*$ är alltså en parameter som bestäms ur den andra termen i Taylor-serien. Uttrycket 6.2 blir då mycket likt fri-elektronenergin och vi kallar därför  $m_e^*$ för elektronens (eller tillståndets) effektiva massa.

Elektronens massa har naturligtvis inte förändrats när vi placerar den i en halvledare, men när den befinner sig i en halvledare påverkas den av kristallens potential och beter sig *som om* den hade en effektiv massa skild från vakuummassan. Elektronens växelverkan med den periodiska potentialen påverkar dess rörelse, och vår analys ovan tillåter oss att kondensera denna påverkan till en justerad massa som ges av bandens kurvatur.

Experimentellt uppmätta effektiva massor för olika halvledare finns i tabeller där de ofta anges i form av en faktor som multipliceras med vakuum-massan, t ex  $(m_e^*/m_0) = 0.067$  för elektroner i ledningsbandet i GaAs.

Elektronens hastighet i ett visst tillstånd, eller mer korrekt uttryckt hastigheten hos ett vågpaket centrerat runt ett visst k, ges av grupphastigheten

$$v_g = \frac{1}{\hbar} \frac{\partial E}{\partial k} = \frac{\hbar (k - k_0)}{m_e^*} \tag{6.3}$$

hos det vågpaket som underförstått beskriver en någotsånär lokaliserad elektron, även om vi framöver aldrig explicit kommer att bilda sådana vågpaket. När yttre krafter verkar på elektronerna beskriver rörelseekvationen (som vi tog fram i kapitel 5)

$$\hbar \frac{d(k-k_0)}{dt} = F_{ext}$$

hur vågvektorn ändras med tiden. Vi har här valt att ange vågvektorn relativt bandets minimum vid  $k_0$ . För detta paraboliska ledningsband kan vi nu direkt relatera elektronens acceleration till den yttre kraften:

$$\frac{dv_g}{dt} = \frac{1}{\hbar^2} \frac{\partial^2 E}{\partial k^2} \cdot \hbar \frac{dk}{dt} = \frac{1}{\hbar^2} \frac{\partial^2 E}{\partial k^2} \cdot F_{ext} = \frac{1}{m_e^*} F_{ext}$$

Vi jämför med Newtons andra lag

$$\frac{dv}{dt} = \frac{F}{m}$$

och finner att införandet av den effektiva massan och grupphastigheten tillåter oss att beskriva elektronens rörelse i ledningsbandet som om den vore en partikel. Vi behöver bara ersätta rörelsemängden  $\hbar k$  i FEM med kristallimpulsen  $\hbar (k - k_0)$  och frielektronmassan m med den effektiva massan  $m_e^*$ .

Vi kan skriva om parabeln (ekv 6.2), som beskriver botten av ledningsbandet, med hjälp av  $v_g$  (ekv 6.3) så att

$$E = E_C + \frac{\hbar^2 (k - k_0)^2}{2m_e^*} = E_C + \frac{m_e^* v_g^2}{2}$$

Här tolkar vi den energi som elektronen har i vila,  $E_C$ , som elektronens potentiella energi och den andra termen, i analogi med den

Kom ihåg att så snart vi har en icke-konstant potential så har vi en kraft:  $F = -\frac{dV}{dx}$ 

Kom ihåg:

 $v_g = \frac{1}{\hbar} \frac{\partial E}{\partial k}$ 

partikelbild av elektronen som vi har (åter)fått, som elektronens kinetiska energi. Vid botten av ledningsbandet blir  $v_g = 0$ , vilket harmonierar med vår beskrivning av elektroniska tillstånd intill bandgapet som stående vågor med grupphastighet noll.

Som vi senare kommer att visa är antalet elektroner i ledningsbandet i regel så litet att sannolikheten att två elektroner skulle vilja besätta samma tillstånd är helt försumbar, vi kan bortse från Pauliprincipen. Därmed kan elektronerna beskrivas som en klassisk gas och deras fördelning på olika tillstånd bestäms av temperaturen. De kommer då att besätta tillstånd så att deras kinetiska energi ges av likafördelningslagen och i medel är

$$\overline{E_{kin}} = \frac{3}{2}kT.$$

#### 6.2 Hål

I valensbandet kommer vi att få en del tomma tillstånd då elektroner exciteras t.ex. termiskt över bandgapet. Vi behöver nu ta reda på hur vi kan beskriva dessa tomma tillstånd.

Ett fullt valensband bidrar inte till strömmen när ett elektriskt fält läggs över en halvledare. När vi lägger på fältet förändras inte elektronernas fördelning på tillstånden, alla tillstånd i valensbandet förblir fyllda (som i grundtillståndet), ingen ström flyter.

När en elektron tas bort från det fyllda bandet kan bandet bidra till laddningstransport. Summan av alla elektroners hastigheter är inte längre noll, laddning kan förflyttas.

Det är nu uppenbart lättare att fokusera på det enda tomma tillståndet än på det oerhört stora antalet tillstånd fyllda av elektroner. Vi kan då tänka oss att ett fyllt band där en elektron saknas är ekvivalent med ett helt fyllt band till vilket vi lagt till en positiv laddning, se figur 6.7. Vi lägger alltså tillbaka elektronen på sin plats, tillsammans med en positiv laddning, och kräver att den positiva laddningen hela tiden precis neutraliserar elektronens laddning.

Vi måste då se till att den positiva laddningen (med hastighet  $v_h$ ) rör sig med samma hastighet som den tillbakalagda elektronen har,  $v_e$ , dvs  $v_h = v_e$ . Då tar den positiva laddningen precis ut effekten av den elektron vi lagt tillbaka för att fylla valensbandet. Eftersom det fulla bandet inte påverkas av ett elektriskt fält behöver vi nu bara se hur fältet påverkar den positiva laddningen.

Ett elektriskt fält  $\mathbf{E}$  leder till en kraft som verkar på laddningarna. På grund av Fermi-fördelningen finns de tomma tillstånden nära



Figur 6.5: I ett fullt band tar elektronernas hastigheter precis ut varandra.



Figur 6.6: Fattas en elektron i bandet så blir summan av alla elektroners hastigheter skild från noll.

toppen av valensbandet. En elektron *nära toppen* av bandet har en negativ effektiv massa  $m^*$  enligt resonemanget i förra stycket (se ekv. 6.1 och fig 6.4), vilket innebär att elektronen hade accelererats i motsatt riktning mot kraftens riktning, dvs i det elektriska fältets riktning:

$$m^* \mathbf{a}_e = \mathbf{F} = -e\mathbf{E}$$
 eller  $\mathbf{a}_e = \frac{-e}{m^*}\mathbf{E} = \frac{e}{|m^*|}\mathbf{E}$ 

där  $\mathbf{a}_e$  är elektronens acceleration. Den positiva laddningen måste accelereras på samma sätt för att fortfarande ta ut effekten av elektronen, dvs dess acceleration

$$\mathbf{a}_h = \frac{e}{m_h^*} \mathbf{E}$$

måste vara lika med elektronens. Detta innebär att den positiva laddningen måste ha en effektiv massa

$$m_h^* = |m^*| = -m^*$$

som är positiv nära toppen av valensbandet, där  $m^* < 0$ . Hålets rörelseekvation blir slutligen

$$m_h^* \frac{dv_h}{dt} = F_{ext} = e\mathcal{E}$$

återigen liknande Newtons andra lag. Vi beskriver valensbandet nära extrempunkten (toppen) som

$$E(k) = E_V + \frac{\hbar^2 (k - k_0)^2}{2m^*}$$

Då  $m^* < 0$ , skriver vi om som

$$E(k) = E_V - \frac{\hbar^2 (k - k_0)^2}{2|m^*|} = E_V - \frac{\hbar^2 (k - k_0)^2}{2m_h^*}$$

Återigen utnyttjar vi $\boldsymbol{v}_g$  för att skriva om detta som

$$E(k) = E_V - \frac{m_h^* v_g^2}{2}$$



borttagen elektron

Figur 6.7: Modell för hål: Ett fullt band med en borttagen elektron modellerar vi som ett helt fullt band samt en tillagd positiv laddning, som hela tiden följer och "tar ut" den elektron som ju egentligen saknas.



Figur 6.8: Hålens kinetiska energi ökar i den negativa E - axelns riktning, de bubblar gärna upp till toppen av valensbandet. Elektroner lägger sig gärna i botten av ledningsbandet.

I analogi med den tidigare diskussionen om elektroner i ledningsbandet tolkar vi  $E_V$  som hålets potentiella energi och den andra termen som hålets kinetiska energi. Vid toppen av valensbandet är  $v_g = 0$  och hålet har då enbart potentiell energi. För  $k \neq k_0$  (vilket innebär att hålet är i ett tillstånd längre ner i bandet) blir  $v_g$  ändlig och hålet får en kinetisk energi. Hålets energi ökar alltså neråt i valensbandet, hålen "bubblar" upp mot valensbandets topp för att minska sin energi.

Finns det tomma tillstånd nära toppen av bandet betraktar vi alltså dessa som positiva partiklar med positiv massa och kallar dem hål.

Den effektiva massan uppvisar kraftiga variationer när vi plottar den som funktion av vågvektorn längs ett band, se figur 6.9. Inom begränsade intervall är effektiva massan approximativt konstant och därmed en meningsfull parameter. En laddningsbärare som tillbringar sitt liv nära en extrempunkt i bandstrukturen kommer då att ha en nästan konstant effektiv massa. Även om den accelereras av ett yttre fält kommer spridningsprocesser på grund av avvikelser från periodicitet att förhindra alltför vidlyftiga exkursioner.

## 6.3 Optiska övergångar

Det finns situationer då vi måste tänka på hur lednings-och valensband är placerade i k-rummet i förhållande till varandra. Figur 6.10 visar bandstrukturen för energier nära bandgapet dels för en halvledare med direkt bandgap, där valensbandets topp ligger och ledningsbandets botten ligger vid samma vågvektor  $\mathbf{k}$ , och dels för en halvledare med indirekt bandgap, där bandens extrempunkter



Figur 6.9: Ett band E(k), hastigheten v(k) och effektiva massan  $m^*(k)$ . Endast nära k = 0och  $k = \pm \pi/a$  är den effektiva massan någotsånär konstant och en meningsfull parameter.



Figur 6.10: Absorption av ljus i en halvledare sker normalt genom elektroniska övergångar från valensbandet till ledningsbandet. Indirekta övergångar kräver medverkan av en *fonon* (vibration hos atomerna i kristallen). Om halvledarna belyses med fotoner motsvarande bandgapsenergin kan en direkt övergång ske i ett direkt bandgap, men inte i ett indirekt. Är fotonenergin tillräckligt stor kan direkta övergångar ske även i indirekta halvledare.

ligger vid olika vågvektorer **k**. Kisel, se figurer 6.1 och 6.3), har indirekt bandgap.

I bandstrukturen hos en kristall är vågvektorn inte längre enkelt relaterad till elektronens rörelsemängd. De elektroniska tillståndens vågfunktioner är ju nu, i närvaron av en periodisk potential, en blandning av många planvågor med olika vågtal. I konserveringslagar, däremot, dyker  $\hbar \mathbf{k}$  upp som om det vore rörelsemängden. För att särskilja  $\hbar \mathbf{k}$  från den sanna rörelsemängden kallar vi  $\hbar \mathbf{k}$ *kristallimpuls*, som vi nämnt i föregående kapitel.

Vi betraktar nu optiska övergångar mellan banden. Genom att absorbera eller emittera en foton kan en elektron övergå från ett tillstånd med ( $\mathbf{k}_{el}, E$ ) till ett tillstånd med ( $\mathbf{k}'_{el}, E'$ ). Konservering av kristallimpuls ger

$$\mathbf{k}_{el}' = \mathbf{k}_{el} \pm \mathbf{k}_{foton}$$

och konservering av energi ger

$$E' = E \pm \hbar \omega_{\text{foton}}.$$

Energin levereras eller tas upp av en foton. Fotonens vågvektor är dock försumbar i förhållande till typiska elektroniska vågvektorer. Halvledares bandgapsenergi är av samma storlek som fotonenergier inom det synliga området och fotonens våglängd är då

$$\lambda \sim 1000 \text{ nm} \qquad \left(hf = \frac{hc}{\lambda} \sim 1 \text{ eV}\right)$$

och typiska atomavstånd är

 $a \sim 0,1$  nm.

Därmed följer

$$k_{\text{foton}} = \frac{2\pi}{\lambda} \ll \frac{\pi}{a},$$

där vi ser att absorption eller emission av en foton inte kan ändra elektronens vågvektor särskilt mycket i förhållande till Brillouinzonens storlek. En *direkt övergång* (vertikala pilar i figur 6.10) ändrar alltså inte elektronens vågvektor märkbart.

Hur kan nu en elektron nära toppen av valensbandet optiskt exciteras till botten av ledningsbandet om bandgapet är indirekt? Vi behöver då blanda in en tredje "partikel" som kan bidra med en vågvektor som kompenserar för elektronens ändring i vågvektor. Denna tredje part är en *fonon*, ett kvantum av en svängningsvåg i kristallen (ljudvågor är ett exempel på en sådan svängning). Vi har då

$$\mathbf{k}_{el}' = \mathbf{k}_{el} \pm \mathbf{k}_{foton} \pm \mathbf{k}_{fonon},$$

 $\operatorname{och}$ 

$$E' = E \pm \hbar \omega_{foton} \pm \hbar \omega_{fonon}$$

Fononer är oftast termiskt exciterade svängningar och har därmed en energi i storleksordningen kT;  $\hbar\omega_{fonon} \sim kT$ . Denna energi är oftast liten nog att kunna försummas i relativt fotonenergier motsvarande bandgapet eller mer. Fononens vågvektor kan däremot vara stor. Vi kan uppskatta den genom att fundera på möjliga våglängder för en atomär svängning sådan som den i figur 6.11. Den kortaste möjliga våglängden är 2a, dvs varannan atom svänger upp och varannan svänger ner, och vågvektorn blir då

$$k_{\text{fonon}} = \frac{2\pi}{\lambda} = \frac{\pi}{a},$$

vilket inte alls är försumbart jämfört med Brillouinzonens storlek. Även en längre våglängd som den i figur 6.11 ger en ganska stor vågvektor.

Fotonen bidrar således med energin, fononen med vågvektor. Men som alltid när flera parter är inblandade i en process kommer sannolikheten för indirekta övergångar att vara betydligt mindre än för direkta övergångar där elektronen bara behöver komma överens med en foton. Som ses i figur 6.10 finns det inget som hindrar en direkt övergång mellan valensband och ledningsband i halvledare



Figur 6.11: De svarta prickarna är atomer vars kollektiva rörelse utgör en våg i kristallen. En fonon är ett kvantum av en sådan vågliknande atomär svängning som t ex en ljudvåg.

med indirekt bandgap om halvledaren belyses med fotoner som har tillräckligt stor energi.

Pilarna i figur 6.10 visar excitationsprocesser – dvs elektroner växelverkar med en foton och ibland en eller flera fononer för att kunna exciteras över bandgapet. Det finns förstås också rekombinationsprocesser – en elektron i ledningsbandet rekombinerar med ett hål i valensbandet och emitterar en foton och ibland fononer.

En elektron som exciterats till ett tillstånd vid en energi en bit över ledningsbandskanten kommer snabbt att göra sig av med energi; typiskt genom att emittera ett antal fononer med låg energi. På så sätt hamnar den i ett tillstånd närmare botten av bandet – se figur 6.12. Vi kan se detta som att systemet strävar mot termisk jämvikt, dvs mot att elektroner är fördelade på tillstånden i bandet enligt Fermifördelningen. Även hålet som bildats i valensbandet gör sig av med energi och hamnar i ett tillstånd närmare toppen av valensbandet. Processerna kallas termisk relaxation och sker med någon typisk relaxationstid, som beror på temperatur och materialparametrar.

Denna termiska relaxationstid är ofta mycket kort, och typiskt kortare än tiden för elektroner och hål att hitta varandra för direkt rekombination över bandgapet. I en situation sådan som i figur 6.12 relaxerar elektronen och hålet till bandkanterna innan de rekombinerar.

Åven i fallet med direkta bandgap sker termisk relaxation. I situationen i figur 6.10 kommer elektronen och hålet från den mer energirika excitationen typiskt att relaxera till tillstånd nära bandkanterna innan de rekombinerar. Fotonerna som emitteras har då i regel energier nära bandgapsenergin, även om excitationsenergin var högre än så.

Emissionsprocesser sker alltså i stort sett uteslutande från botten av ledningsbandet till toppen av valensbandet. I halvledare med direkt bandgap sker fotonemission då snabbt, medan processen i halvledare med indirekt bandgap är mycket långsammare pga behovet av en fonon för att ta hand om ändringen i vågvektor elektronen måste undergå för att ta sig från ledningsbandet till valensbandet.

I många material ger absorptionen över bandgapet upphov till färg hos materialet när delar av det synliga spektrumet absorberas (se fig 6.13). Bandgapsfärger tenderar att ligga i det röd-gula området eftersom bandgapet om något absorberar fotoner med högre energi.



E

Figur 6.12: Elektronen och hålet, som skapats genom absorption av en foton, relaxerar snabbt ner mot tillstånd nära bandens extrempunkter.



Figur 6.13: CdS absorberar blått och grönt ljus och ser därmed gult ut. CdSe absorberar alla färger och ser svart ut. Blandkristaller med varierande bandgap övergår från gul över röd till svart färg.

#### 6.4 Laddningsbärarkoncentrationer

En halvledares konduktivitet bestäms framförallt av laddningsbärarnas koncentrationer, dvs koncentrationen n av elektroner i ledningsbandet och koncentrationen p av hål i valensbandet. För att bestämma elektronkoncentrationen behöver vi veta antalet besatta tillstånd i ledningsbandet per volymsenhet. För hålkoncentrationen behöver vi veta antalet *icke* besatta tillstånd i valensbandet per volymsenhet. Dessa antal kan beräknas med hjälp av tillståndstätheterna i lednings- och valensband och Fermi-Dirac fördelningen. Vi förutsätter att både ledningsband och valensband kan beskrivas med paraboler, men med olika effektiva massor, som vi diskuterade i början av kapitlet. Då E(k) är paraboler kan vi smidigt nog återanvända uttrycken för tillståndstäthet som vi tog fram för frielektronfallet i kapitel 4 - vi behöver bara byta ut elektronmassan mot den effektiva massan.

Tillståndstätheterna lyder då:

$$Z_{C}(E) = \underbrace{4\pi \left(\frac{2m_{e}^{*}}{h^{2}}\right)^{3/2}}_{K_{C}} \sqrt{E - E_{C}} \qquad \text{(ledningsband)}$$
$$Z_{V}(E) = \underbrace{4\pi \left(\frac{2m_{h}^{*}}{h^{2}}\right)^{3/2}}_{K_{V}} \sqrt{E_{V} - E} \qquad \text{(valensband)}$$

Eftersom vi vill beräkna koncentrationerna av elektroner och hål utelämnar vi faktorn  $V = L^3$  i  $K_C$  och  $K_V$ . Tillståndstätheterna blir då per volymsenhet.

elektroner:  $E = E_C + \frac{\hbar^2 \mathbf{k}^2}{2m_e^*}$ hål:  $E = E_V - \frac{\hbar^2 \mathbf{k}^2}{2m_h^*}$  $E_C$  $E_F$  $E_V$ k För att hitta koncentrationerna av laddningsbärare behöver vi inte bara veta något om hur många möjliga tillstånd det finns för dem att besätta, utan också hur många av dessa tillstånd som faktiskt är besatta av elektroner eller hål. För detta använder vi Fermi-Dirac fördelningen:

$$F(E) = \frac{1}{e^{(E-E_F)/kT} + 1}$$

Kom ihåg att  $E_F$  i F(E) betyder Ferminivå, som kan vara temperaturberoende.

Ferminivån ligger i de flesta fall vi betraktar i bandgapet, inte alltför nära bandkanterna. Begränsar vi analysen till dessa fall, dvs kräver att

$$E_C - E_F \gg kT$$
 och  
 $E_F - E_V \gg kT$ ,

så kan vi införa en approximation till F(E) som avsevärt förenklar räkningarna. För tillstånd i ledningsbandet hittar vi sannolikheten att dessa är besatta av elektroner genom att betrakta F(E) för energier  $E > E_C$ :

$$F(E) = \frac{1}{e^{(E-E_F)/kT} + 1} \Rightarrow F(E) = e^{-(E-E_F)/kT}$$
 om  $E_C - E_F \gg kT$ 

För tillstånd i valensbandet hittar vi sannolikheten att dessa är tomma, dvs besatta av hål, genom att betrakta 1-F(E) för energier  $E < E_V$ :

$$1 - F(E) = \frac{1}{e^{-(E - E_F)/kT} + 1} \Rightarrow 1 - F(E) = e^{(E - E_F)/kT} \quad \text{om} \quad E_F - E_V \gg kT$$

Eftersom vi under förutsättningarna ovan har  $F(E) \ll 1$  för  $E > E_C$  (elektroner) och  $1 - F(E) \ll 1$  för  $E < E_V$  (hål) blir det få elektroner och hål i banden; banden blir nästan tomma.

Vi kan nu beräkna elektronkoncentrationen n i ledningsbandet som

$$n = \int_{E_C}^{\infty} Z_C(E) F(E) dE = K_C \int_{E_C}^{\infty} \sqrt{E - E_C} e^{-(E - E_F)/kT} dE.$$

För den intresserade läsaren visar vi här hur man kan lösa den här integralen - resultatet återfinns lite längre ner (ekv. 6.4).

$$n = K_C (kT)^{3/2} e^{-(E_C - E_F)/kT} \int_0^\infty \sqrt{x} e^{-x} dx$$
$$\int_0^\infty \sqrt{x} e^{-x} dx = \underbrace{-x^{1/2} e^{-x} \Big|_0^\infty}_0 + \frac{1}{2} \int_0^\infty x^{-1/2} e^{-x} dx$$



Figur 6.14: Vi förutsätter att Ferminivån ligger minst några kTifrån bandkanterna.

Variabelsubstitution:

$$\frac{E - E_C}{kT} = a$$

Partialintegration



159

Figur 6.15: En halvledares paraboliska band, motsvarande tillståndstätheter Z(E) och Fermi-Dirac fördelningsfunktioner F(E). Längst till höger antalet besatta tillstånd per energi<br/>intervall och volymsenhet. Ytan under kurvorna ger elektron- resp. hålkon<br/>centrationen.

$$\int_{0}^{\infty} x^{-1/2} e^{-x} dx = 2 \int_{0}^{\infty} e^{-y^{2}} dy \qquad (x = y^{2})$$

$$\int_{-\infty}^{\infty} e^{-x^{2}} dx \cdot \int_{-\infty}^{\infty} e^{-y^{2}} dy = \int_{0}^{\infty} \int_{0}^{2\pi} e^{-r^{2}} r dr d\theta = 2\pi \int_{0}^{\infty} r e^{-r^{2}} dr =$$

$$= -\pi \int_{0}^{\infty} \frac{d}{dr} \left( e^{-r^{2}} \right) dr = -\pi e^{-r^{2}} \Big|_{0}^{\infty} = \pi$$

$$\int_{-\infty}^{\infty} e^{-y^{2}} dy = \sqrt{\pi} \quad \rightarrow \quad \int_{0}^{\infty} e^{-y^{2}} dy = \frac{\sqrt{\pi}}{2} \qquad \text{och slutligen}$$

$$\int_{0}^{\infty} \sqrt{x} e^{-x} dx = \frac{1}{2} \cdot 2\frac{\sqrt{\pi}}{2} = \frac{\sqrt{\pi}}{2}$$

Integralen kan lösas genom att utgå från produkten av två integraler och gå över till integration i ett plan.

Elektronkoncentrationen blir:

$$n = K_C \frac{\sqrt{\pi}}{2} (kT)^{3/2} e^{-(E_C - E_F)/kT} = N_C e^{-(E_C - E_F)/kT}$$
(6.4)

där den effektiva tillståndstätheten  $N_{C}$  för ledningsbandet är

$$N_C = 2\left(\frac{2\pi m_e^* kT}{h^2}\right)^{3/2} = 2.5 \cdot 10^{25} \cdot \left(\frac{m_e^*}{m} \frac{T}{300K}\right)^{3/2} \mathrm{m}^{-3}$$

För hålen skriver vi

$$p = \int_{-\infty}^{E_V} Z_V(E)(1 - F(E))dE = K_V \int_{-\infty}^{E_V} \sqrt{E_V - E} e^{(E - E_F)/kT} dE =$$
$$= K_V \int_0^\infty \sqrt{E} e^{(-E - E_F)/kT} dE$$

Med en liknande procedur som för elektronerna:

$$\frac{E}{kT} = x \qquad \text{och} \qquad \int_0^\infty \sqrt{x} e^{-x} dx = \frac{\sqrt{\pi}}{2}$$
$$p = K_V (kT)^{3/2} e^{-(E_F - E_V)/kT} \int_0^\infty \sqrt{x} e^{-x} dx = K_V (kT)^{3/2} e^{-(E_F - E_V)/kT} \frac{\sqrt{\pi}}{2}$$

får vi

$$p = K_V \frac{\sqrt{\pi}}{2} (kT)^{3/2} e^{-(E_F - E_V)/kT} = N_V e^{-(E_F - E_V)/kT}$$
(6.5)

där den effektiva tillståndstätheten  $N_V$  för valensbandet är

$$N_V = 2\left(\frac{2\pi m_h^* kT}{h^2}\right)^{3/2} = 2.5 \cdot 10^{25} \cdot \left(\frac{m_h^*}{m} \frac{T}{300K}\right)^{3/2} \mathrm{m}^{-3}$$

Notera att uttrycken för n och p fortfarande innehåller Ferminivån  $E_F$  vars läge inte har bestämts än. Multiplicerar vi dessa uttryck (ekv 6.4 och 6.5) med varandra försvinner  $E_F$  och vi finner

### 6.5 Massverkans lag

$$n \cdot p = N_C N_V e^{-E_g/kT} = n_i^2$$

Detta viktiga samband talar om för oss att produkten  $n \cdot p$  (och därmed den här definierade storheten  $n_i$ ) vid termisk jämvikt endast beror av materialparametrar  $E_g$ ,  $m_e^*$ ,  $m_h^*$  och temperaturen T, dvs är *oberoende* av en möjlig dopning (dvs medveten förorening av kristallen, se nedan)!

#### 6.6 Intrinsiska halvledare

Är halvledaren ren, dvs störatomernas koncentration är försumbar, kallas den *intrinsisk* eller*egenledande* och elektroner och hål uppstår endast genom excitationer över bandgapet, dvs

$$n = p = n_i$$

där vi nu identifiera<br/>r $n_i$ som den intrinsiska laddningsbärarkon<br/>centrationen. Ferminivåns läge bestäms till

$$E_{Fi} = \frac{E_C + E_V}{2} - \frac{kT}{2} \ln\left(\frac{N_C}{N_V}\right) = \frac{E_C + E_V}{2} + \frac{3}{4}kT \ln\left(\frac{m_h^*}{m_e^*}\right)$$

$$\begin{split} N_C e^{-(E_C-E_F)/kT} &= \\ &= N_V e^{-(E_F-E_V)/kT} \rightarrow \\ &\rightarrow \frac{E_F}{kT} = -\frac{1}{2} \ln \frac{N_C}{N_V} + \frac{E_g}{2kT} \end{split}$$

Denna *intrinsiska Ferminivå*  $E_{Fi}$  ligger mitt i bandgapet så när som på några kT.

### 6.7 Dopning

En förutsättning för halvledarnas användbarhet inom elektronik – och därmed för hela den mikroelektroniska revolutionen – är möjligheten att genom dopning förändra halvledarnas egenskaper. Dopning innebär att vi medvetet förorenar mycket rena kristaller med störatomer. Vanliga dopatomer i kisel är fosfor (P) ur grupp V i det periodiska systemet, och aluminium (Al) ur grupp III. När vi dopar kan vi bestämma om halvledaren ska ha hål eller elektroner som huvudsakliga laddningsbärare, och hur stor koncentrationen ska vara.

Man ersätter alltså ett fåtal Si-atomer, om det är kisel man dopar, med något atomslag som har en valenselektron mer eller en mindre än vad kisel har (4 st). Det är viktigt att dessa atomer, störatomerna, ersätter kiselatomer på deras platser i kristallen för att bli elektriskt aktiva, dvs för att bidra med laddningsbärare, elektroner eller hål, till halvledaren. Hamnar störatomen på en godtycklig plats i kristallen, t.ex. inklämd mellan två Si-atomer som sitter där de ska, kan det hända att den inte gör någon nytta alls för att öka konduktiviteten. Vidare gäller allmänt att när vi bryter kristallens periodicitet så är det inte längre förbjudet för elektroniska tillstånd att finnas i bandgapet lokalt nära störatomen.

Vilken effekt har det nu att ersätta en Si-atom med en P-atom vid T = 0? Vi kan tänka oss – just det, bara tänka – att vi gör detta inte genom att faktiskt byta ut atomerna utan genom att lägga till en proton i en kiselkärna – vi får en P<sup>+</sup>-jon – och att samtidigt lägga en extra elektron i halvledaren. Valensbandet är fortfarande fullt, elektronen måste – i en första approximation – hamna i ledningsbandet där den har effektiva massan  $m_e^*$ . Elektronen i ledningsbandet ser dock fortfarande P<sup>+</sup>-jonens positiva laddning och kan bli bunden. Vi har då en ensam elektron som kan bli bunden till en jon (P<sup>+</sup>) som har en positiv elementarladdning. Detta liknar situationen i en väteatom. Problemet har alltså reducerats till ett väteproblem inbäddat i en kiselkristall.

Grundtillståndet i en väteatom, där elektronen rör sig i Coulombpotentialen

$$V(r) = -\frac{e^2}{4\pi\varepsilon_0 r},$$

har energin

$$E_{1s} = -\frac{e^2}{8\pi\varepsilon_0 a_0} = -\frac{me^4}{2(4\pi\varepsilon_0)^2\hbar^2} = -13.6 \text{ eV}$$

Även en mycket ren kiselkristall har en (oavsiktlig) störatom på ungefär  $10^{12}$  kiselatomer. När man dopar introducerar man en dopatom på ungefär  $10^6$  kiselatomer. Dopkoncentrationen är alltså mycket liten jämfört med koncentrationen av kiselatomer, kristallen är fortfarande kisel!

Man kan faktiskt dopa kisel genom neutron-bestrålning som leder till transmutation, dvs Si kärnor omvandlas till P kärnor.



Figur 6.16: Vänster: En elektron bunden till den positiva fosforjonen rör sig på en bana med stor radie. Höger: Polarisationen av kiselkristallen på grund av fältet mellan  $P^+$  och  $e^-$  leder till förskjutning av laddning. Attraktionen mellan  $P^+$  och elektronen skärmas och reduceras därmed.

där Bohr-radien är

$$a_0 = \frac{4\pi\varepsilon_0\hbar^2}{me^2} = 0.0528 \text{ nm}$$

Blir elektronen bunden med 13,6 eV till P<sup>+</sup>-jonen kommer den aldrig att kunna bidra till laddningstransport, det vore omöjligt att excitera den termiskt till ledningsbandet. Även om vi ersätter elektronmassan m med effektiva massan reduceras bindningsenergin bara till ungefär en tredjedel.

Nu måste vi också ta hänsyn till att vårt väteliknande problem är inbäddat i kisel. Om vi antar att elektronen endast blir svagt bunden, dvs rör sig på en bana med stor radie som i figur 6.16, kommer det att finnas kiselatomer mellan P<sup>+</sup> och elektronen och dessa kommer att *polariseras* av det elektriska fältet mellan P<sup>+</sup> och  $e^-$ . Detta tar vi hänsyn till genom att ersätta  $\varepsilon_0 \mod \varepsilon_0 \cdot \varepsilon_r$  där  $\varepsilon_r$ är den relativa dielektricitetskonstanten. I t ex kisel är  $\varepsilon_r = 11.4$ vilket leder till att Coulombpotentialen blir betydligt svagare,

$$V(r) = -\frac{e^2}{4\pi\varepsilon_0\varepsilon_r r},$$

den effektiva Bohr-radien blir betydligt större,

$$a_0^* = \frac{4\pi\varepsilon_0\varepsilon_r\hbar^2}{m_e^*e^2} = a_0 \cdot \frac{\varepsilon_r m}{m_e^*}$$

och bindningsenergin i grundtillståndet blir betydligt mindre än i väte,

$$E_C - E_D = \frac{me^4}{2(4\pi\varepsilon_0)^2h^2} \cdot \frac{m_e^*}{m\varepsilon_r^2} = 13.6\frac{m_e^*}{m\varepsilon_r^2} \text{ eV}.$$

Hur stor är den termiska energin kT vid rumstemperatur?

Enligt ny standard heter  $\varepsilon_0$  permittiviteten i vakuum och  $\varepsilon_r$  permittivitetstalet.

 $E_D$  är den bundna elektronens energinivå och elektronens bindningsenergi (den energi som måste tillföras för att elektronen ska kunna bli fri, dvs exciteras till ledningsbandet) betecknar vi med  $E_d$ :

$$E_d = E_C - E_D$$

Faktorn  $m_e^*/(m\varepsilon_r^2)$  är av storleksordningen 10<sup>3</sup>, bindningsenergin brukar för olika atomer ur grupp V hamna kring 0,05 eV. Då är det fullt möjligt att elektronen är termiskt exciterad upp till ledningsbandet vid rumstemperatur. Störatomen bidrar alltså med en rörlig elektron och kallas *donator*. Ledningsförmågan ökar.

För att studera en störatom ur grupp III förfar vi helt analogt. Vi gör om en Si-atom till en Al-atom genom att lägga till en (tänkt) negativ laddning till kärnan (denna förintar då en proton så att vi får en Al<sup>-</sup>-jon) och tar samtidigt bort en elektron ur halvledaren, det skapas därmed ett hål i valensbandet. Detta positiva hål binds till den negativa jonen, bindningsenergin blir av samma storleksordning som för donatorn. Acceptorns bindningsenergi betecknar vi med  $E_a = E_A - E_V$  (se figur 6.18). Vid rumstemperatur är acceptorn joniserad och bidrar till ledningsförmågan genom att ge upphov till ett hål i ledningsbandet, vi har en *acceptor*. Termen antyder att när störatomens hål exciteras till valensbandet, kan vi också se det som att störatomen accepterar en elektron.

#### 6.8 Extrinsiska halvledare

Förorenar vi en halvledare med acceptorer eller donatorer för att öka laddningsbärarkoncentrationen kallar vi den *extrinsisk* eller *störledande*. Då bestäms ledningsförmågan inte av värdkristallen (t ex Si) utan av dopningen.

Kristaller dopade med donatorer kallas n-typ halvledare och de dopade med acceptorer kallas p-typ. Generellt kan båda typerna av dopatomer förekomma samtidigt, även om den ena då ofta dominerar. När vi dopar tillför vi neutrala atomer till materialet – vi tillför alltså inte någon nettoladdning. Om en donator joniseras har vi en positiv donatorjon och en negativ elektron, och kristallen är fortfarande neutral. Kristallens *laddningsneutralitet* kan skrivas:

$$n + N_A^- = p + N_D^+$$

Här är  $N_D$  och  $N_A$  donatorernas resp. acceptorernas koncentrationer och  $N_D^+$  och  $N_A^-$  koncentrationerna av joniserade donatorer och

$$\frac{m_{\rm e}^*}{E_{\rm D}} \sim E_{\rm C}$$

Figur 6.17: I en donators grundtillstånd är en elektron svagt bunden till donatorjonen. Exciteras elektronen till ledningsbandet blir donatorn tom och elektronen en (fri) ledningselektron.



Figur 6.18: I en acceptors grundtillstånd är ett hål svagt bundet till acceptorjonen. Att hålet exciteras till valensbandet är ekvivalent med att en elektron exciteras till acceptornivån från valensbandet.

Positiva laddningar: hål och donatorjoner Negativa laddningar: elektroner och acceptorjoner



Figur 6.19: a) n-typ halvledare vid T = 0 K. Alla elektroner är bundna till sina donatoratomer. Fermifördelningen F(E) är en stegfunktion. b) n-typ halvledare vid låga T. Några elektroner har exciterats från donatorerna till ledningsbandet och F(E) har "mjukats upp" något.

acceptorer,

$$N_D^+ = N_D \left(1 - F(E_D)\right)$$
 och  $N_A^- = N_A F(E_A)$ 

 $1-F(E_D)$  ger sannolikheten att donatorn är tom, dvs joniserad och därmed positivt laddad.  $F(E_A)$  ger sannolikheten att acceptornivån är besatt med en elektron, hålet har då exciterats och acceptorn är joniserad, dvs negativt laddad.

När vi nu ska undersöka laddningsbärarkoncentrationernas temperaturberoende väljer vi som exempel en n-typ halvledare; dvs fallet att  $N_D \gg N_A$ . Vi antar då att acceptorkoncentrationen kan försummas:

$$N_A \approx 0$$
 och  $n = p + N_D^+$ 

Vidare antar vi, vilket normalt är fallet i komponenter, att donatorerna är fullständigt joniserade vid rumstemperatur,  $N_D^+ = N_D$ , och att donatorkoncentrationen är mycket större än den intrinsiska laddningsbärarkoncentrationen vid rumstemperatur,

$$N_D \gg n_i (300 \text{ K})$$

I det här fallet är n mycket större än p. Elektronerna kallas då majoritetsladdningsbärare och hålen minoritetsladdningsbärare. I kommande kapitel kommer vi ofta att skriva  $n_{n0}$  och  $p_{n0}$  för att beteckna elektron- respektive hålkoncentrationen i en n-typ halv-ledare.

Den kommade diskussionen fokuserar på en halvledare av n-typ. På helt analogt sätt kan man härleda motsvarande samband för en p-typ halvledare.

 $\mathbf{T}=\mathbf{0}~\mathbf{K}:$  figur 6.19a. Vid absoluta nollpunkten har alla donatorer kvar sina elektroner - inga elektroner finns i ledningsbandet. Fer-

 $p = n_i^2/n$ enligt massverkans lag. minivån, som vid T = 0 sammanfaller med Fermienergin, måste då ligga över donatornivån  $E_D$  men under ledningsbandskanten  $E_C$ .

Låga T: figur 6.19b. Låter vi temperaturen öka förväntar vi oss att donatorelektronerna börjar exciteras termiskt till ledningsbandet, men excitationer kommer ännu inte att kunna ske över bandgapet i märkbar utsträckning, och därmed har vi fortfarande  $p \ll n = N_D^+$ . Steget i Fermi-Dirac fördelningen kommer att bli något mjukare och  $F(E_D) < 1$  då en del elektroner lämnat donatorerna.  $E_F$  ligger kvar i närheten av  $E_D$  så länge donatornivåerna delvis är fyllda.

Då vi har försumbar excitation över bandgapet och därmed att  $p \ll n$  kan vi försumma p i neutralitetsvillkoret och får

$$n = N_D^+$$

Vi bestämmer först n genom att eliminera  $E_F$  ur sambanden

$$n = N_C e^{-(E_C - E_F)/kT} = N_C e^{-(E_C - E_D)/kT} e^{-(E_D - E_F)/kT}$$

 $\operatorname{och}$ 

$$N_D^+ = N_D \left(1 - F(E_D)\right) = \frac{N_D}{e^{-(E_D - E_F)/kT} + 1}$$

Vi får då

$$n = \frac{N_D}{\frac{n}{N_C} e^{(E_C - E_D)/kT} + 1},$$
$$\frac{n^2}{N_C} e^{(E_C - E_D)/kT} + n = N_D \quad \text{och slutligen}$$
$$n^2 = (N_D - n)N_C e^{-(E_C - E_D)/kT}$$

 $Om \ N_D \gg n$ , dvs  $N_D \gg N_D^+$ , vilket gäller så länge de flesta donatorer har kvar sin elektron, följer

$$n = \sqrt{N_D N_C} e^{-(E_C - E_D)/2kT}$$
(6.6)

 $\operatorname{och}$ 

$$E_F = \frac{E_C + E_D}{2} - \frac{kT}{2} \ln\left(\frac{N_C}{N_D}\right) \tag{6.7}$$

Notera att ekv. 6.6 bara gäller under omständigheter då villkoret  $n \ll N_D$  är uppfyllt - det var under detta antagande vi härledde sambandet.

Vi noterar även, från ekv. 6.7, att Ferminivån  $E_F$  vid T = 0 har ett läge mellan donatornivån och ledningsbandskanten för att med ökande temperatur sjunka något i energi. Vi gör förenklingen  $N_D \gg n$ , vilket medför att vårt resultat stämmer dåligt om detta inte är uppfyllt.

När n är bestämd, följer Ferminivåns läge ur sambandet

$$n = N_C e^{-(E_C - E_F)/kT}.$$

som vi tidigare härlett.



Figur 6.20: a) n-typ halvledare vid så hög temperatur att i princip alla donatorer släppt sin elektron till ledningsbandet (mättnad).  $E_F$  ligger nu under  $E_D$ . b) n-typ halvledare vid så höga T att excitationen över bandgapet är påtaglig.  $E_F$  ligger nu mycket nära mitten av bandgapet.

**Mättnadsområdet:** figur 6.20a. När temperaturen stiger ytterligare kommer i stort sett alla donatorer att bli joniserade,  $N_D^+ = N_D$ . Fortfarande är excitationen över bandgapet svag. Vi har alltså

$$n = p + N_D^+ = p + N_D \ge N_D \gg n_i$$

 $(N_D \gg n_i$  enligt förutsättningarna ovan). Fortfarande gäller massverkans lag så att

 $p = \frac{n_i^2}{n} \ll n_i$  och därmed  $n = N_D$ .

Ur  $n = N_D = N_C e^{-(E_C - E_F)/kT}$  följer att

$$E_F = E_C - kT \, \ln\left(\frac{N_C}{N_D}\right).$$

Ferminivån ligger nu närmare mitten av bandgapet än tidigare.

**Höga T:** figur 6.20b. Donatorerna är väsentligen alla joniserade och excitationer över bandgapet har ökat så kraftigt att antalet elektroner exciterade från valensbandet vida överstiger antalet som exciterats från donatorerna. Halvledaren beter sig intrinsiskt,  $n_i \gg N_D$ , dess ledningsförmåga beror inte längre på dopningen och störatomkoncentrationen kan försummas.

$$n=p=n_i.$$

Ferminivån  $E_F$ sammanfaller med den intrinsiska Ferminivån  $E_{Fi}$ 

$$E_F = E_{Fi} = \frac{E_C + E_V}{2} - \frac{kT}{2} \ln\left(\frac{N_C}{N_V}\right) = \frac{E_C + E_V}{2} + \frac{3}{4}kT \ln\left(\frac{m_h^*}{m_e^*}\right)$$



**Elektroniska komponenter** ligger normalt i detta intervall där majoritetsladdningsbärarkoncentration är konstant.

**Exempel:** Si:  $n_i(300 \ K) = 1,0 \cdot 10^{16} \ m^{-3}$ . Vid rumstemperatur och med ett typiskt värde  $N_D = 10^{22} \ m^{-3} \ (N_A = 0)$  får vi

$$N_D \gg n_i \Rightarrow n = p + N_D \gg n_i \Rightarrow p = \frac{n_i^2}{n} \ll n_i$$
$$\Rightarrow n = N_D = 10^{22} \text{ m}^{-3} , p = \frac{n_i^2}{N_D} \sim 10^{10} \text{ m}^{-3}$$

Alla donatorer är fullständigt joniserade, elektronkoncentrationen är många storleksordningar större än hålkoncentrationen. Detta är normalfallet i de n-typ halvledare vi kommer att möta framöver.

Vi tar till sist fram två variationer på uttrycken för n och p som kommer att vara till nytta framöver. För elektronkoncentrationen gäller allmänt, dvs även om halvledaren är dopad (vilket påverkar Ferminivåns läge)

$$n = N_C e^{-(E_C - E_F)/kT}$$

För en intrinsisk halvledare, där  $n = p = n_i$ , gavs elektron- och hålkoncentrationen av

$$n_i = N_C e^{-(E_C - E_{Fi})/kT}.$$

Därmed kan vi relatera n till  $n_i$  genom

$$n = n_i e^{(E_F - E_{Fi})/kT}$$

och analogt för hålkoncentrationen

$$p = n_i e^{(E_{Fi} - E_F)/kT}$$

Ökar vi genom dopning t<br/> exnjämfört med det intrinsiska fallet, dv<br/>s $n\gg n_i$ , så måste enligt sambanden ovan gälla

$$e^{(E_F - E_{Fi})/kT} \gg 1,$$

dv<br/>s $E_F$ måste ligga över  $E_{Fi}$ . Samtidigt är då

$$e^{(E_{Fi}-E_F)/kT} \ll 1$$

vilket innebär att p måste minska jämfört med  $n_i$ . Detta ser vi även explicit genom massverkans lag,  $np = n_i^2$ . Ökar vi n, måste p minska förutsatt att vi är i termisk jämvikt.

#### 6.9 Elektron- och hålströmmar i halvledare

Vi har sett att vi verkar kunna betrakta elektroner och hål i en halvklassisk bild som "partiklar" med en effektiv massa och en hastighet som ges av grupphastigheten (för de vågpaket vi egentligen vet att elektroner och hål utgörs av). Laddningsbärare i halvledaren är elektroner i ledningsbandet och hål i valensbandet. Både elektroner och hål kan accelereras av ett elektriskt fält och bidra till strömmen. Konduktiviteten  $\sigma$  ges av summan av två termer

$$\sigma = \frac{e^2 \tau_e n}{m_e^*} + \frac{e^2 \tau_h p}{m_h^*} = e \mu_e n + e \mu_h p$$

där n och p är koncentrationen av elektroner resp. hål och  $\tau_e$  och  $\tau_h$  är kollisionstiderna. Elektroner och hål skapas vid uppvärmning, dvs tillförsel av termisk energi, men även t ex belysning kan excitera elektroner över bandgapet vilket vi utnyttjar i fotoledare.

Vi införde elektronens effektiva massa för att kunna beskriva elektronerna i en halv-klassisk "partikel"-bild med rörelseekvationen

$$m_e^* \frac{dv_e}{dt} = F_{ext} = -e\mathscr{E}$$
  $(m_e^* > 0 \text{ i ledningsbandet}).$ 

Denna bild föredrar vi ofta då den ligger närmast vår vardagliga erfarenhet. Slutligen kompletterade vi denna halv-klassiska bild med hål i valensbandet som beskrivs av en motsvarande rörelseekvation

$$m_h^* \frac{dv_h}{dt} = F_{ext} = e\mathscr{E}$$
  $(m_h^* > 0 \text{ i valensbandet}).$ 

I kapitel 2 tog vi fram uttryck för strömtätheter givet att det fanns både elektroner och positiva laddningsbärare. Vi vet nu att dessa är hål i valensbandet. Uttrycken innehöll både *driftström* och *diffusionsström*:

$$J_e = eD_e \frac{dn}{dx} + \frac{e^2 \tau_e n}{m_e^*} \mathscr{E}$$
$$J_h = -eD_h \frac{dp}{dx} + \frac{e^2 \tau_h p}{m_h^*} \mathscr{E}$$

### 6.10 Spridningsprocesser

I mobiliteten ingår spridningstiderna  $\tau_e$  (för elektroner) och  $\tau_h$  (för hål). När vi studerade bandstrukturer fann vi att den perfekt periodiska kristallen inte kan sprida laddningsbärarna. Eftersom Ohms lag kräver friktion eller spridningsprocesser drog vi slutsatsen att

n - elektronkoncentration p - hålkoncentration  $D_e, D_h$  - diffusionskonstanter



Figur 6.21: Mobilitet  $\mu$  som funktion av temperatur T. Vid låga T begränsas mobiliteten av spridning mot orenheter/störatomer, vid höga Tbegränsas den istället av spridning mot atomära vibrationer. För metaller är spridning mot joniserade orenheter dominerande bara vid mycket låga T, varför bilden ovan kan sägas representera halvledare. I figuren är två olika dopnivåer representerade,  $N_1 < N_2$ . Den lägre dopkoncentrationen  $N_1$  ger en högre mobilitet vid låga T eftersom färre störatomer ger mindre spridning. De heldragna linjerna visar den sammanlagda effekten av båda spridningsmekanismerna.

det måste vara avvikelser från periodiciteten, dvs atomära svängningar och störställen, som sprider laddningsbärarna.

I. Atomers termiska svängningar: Atomernas termiska svängningar ökar med temperaturen. De stör kristallens perfekta periodicitet och reducerar mobiliteten vid högre T (vid låga T bidrar de mycket lite till spridningen). En sådan svängning kallar vi, som tidigare, fonon (eng. phonon). För halvledare gäller att (visas strax)

 $\mu_{ph} \sim \tau \sim T^{-3/2}$ 

II. Joniserande störatomer: Grunda störställen som introduceras vid dopningen är normalt joniserade vid rumstemperatur och deras Coulombpotential dominerar spridningen vid störställen. Vid högre T ökar partiklarnas kinetiska energi och de böjs av mindre i Coulombpotentialen. Spridningen mot joniserade störatomer är således mindre effektiv vid höga T. I halvledare gäller att (även detta visas strax)

 $\mu_{imp} \sim \tau \sim T^{3/2}/N$ 

(index imp. från impurity) där N är störatomkoncentrationen.

Båda effekterna bidrar till att begränsa mobiliteten. För låga T dominerar spridning mot stör- eller dopatomer och vid högre T

dominerar spridning mot atomära vibrationer. Mobilitetens temperaturberoende sammanfattas schematiskt i figur 6.21. Den totala mobiliteten ges som tidigare av

$$\frac{1}{\mu} = \frac{1}{\mu_{ph}} + \frac{1}{\mu_{imp}}$$

I halvledare är det inte enbart spridningsprocesserna som bestämmer konduktivitetens temperaturberoende utan även laddningsbärarkoncentrationerna som uppvisar ett exponentiellt beroende av temperaturen.

Båda typerna av spridningsprocesser förekommer i både metaller och halvledare, men temperaturberoendet är olika vilket vi ska försöka troliggöra i följande avsnitt (baserat på "Principles of Electronic Materials and Devices" av S O Kasap).

#### 6.10.1 Spridning mot atomära vibrationer

Varje vibrerande atom i kristaller har i medel, per rumsdimension, energin:

$$E_x = \frac{p_x^2}{2M} + \frac{M}{2}\omega^2 x^2 = kT$$

där de kinetiska och potentiella bidragen är lika stora enligt ekvipartitionsteoremet. Atomerna vibrerar med en amplitud x från sin utgångsposition (figur 6.22) och rör sig då över en yta  $S = \pi x^2 \sim kT$ . Om elektronen hamnar inom avståndet x från atomen, kan vi tänka oss att den sprids.

I metaller, inom fri-elektronmodellen (FEM), har endast elektroner nära Ferminivån möjlighet att spridas. När vi diskuterar spridning i metaller fokuserar vi då på dessa elektroner, som alla har ungefär samma fart  $v_F$ .

Figur 6.22 visar elektronens väg mellan två spridningar. En elektron vid Fermienergin med hastighet  $v = v_F$  tillryggalägger en sträcka



Figur 6.22: En elektron med hastighet v färdas mellan två spridningar en sträcka  $\ell = v\tau$ , där  $\tau$  är medeltiden mellan kollisioner och  $\ell$  är fria medelväglängden.

 $\ell = v_F \tau$  mellan två spridningshändelser. Inom volymen  $Sv_F \tau$  finns då precis ett spridningscentrum. Om spridningscentra förekommer i en koncentration  $N_s$ , måste  $N_s \times Sv_F \tau$  vara 1. Vi får att

$$\tau = \frac{1}{Sv_F N_s} \sim \frac{1}{S} \sim T^{-1}$$

När vi istället betraktar spridning av laddningsbärare i halvledare mot vibrerande atomer måste vi tolka hastigheten v i fig. 1 som den termiska hastigheten  $v_{th}$  vilken beror på T som  $v_{th} \sim T^{1/2}$  (se kap. 2):

$$\tau = \frac{1}{Sv_{th}N_s} \sim \frac{1}{T \times T^{1/2}} = T^{-3/2}$$

Temperaturberoendet blir alltså inte detsamma för metaller och halvledare pga att elektroner i metaller fördelas på ett sätt som är hårt styrt av Pauliprincipen, medan elektroner/hål i halvledare fördelas termiskt. Spridning mot atomära vibrationer är oftast den dominerande spridningsmekanismen i halvledare vid rumstemperatur (och även vid höga T).

#### 6.10.2 Spridning mot joniserade orenheter

För spridning mot joniserade orenheter får vi hitta en annan modell. Vi tänker oss en elektron som är på väg mot en positiv jon, t.ex. en joniserad donator. Den negativa elektronen har en potentiell energi relativt jonen som vi kan skriva som

$$E_{pot} = \frac{e^2}{4\pi\varepsilon_0\varepsilon_r r}$$

där r är avståndet mellan jonen och elektronen.

Vi tänker oss nu att om elektronens kinetiska energi är större än den potentiella,  $E_{kin} > E_{pot}$ , så fortsätter elektronen utan att spridas, medan om  $E_{kin} \leq E_{pot}$  så har vi spridning. Vi kan då hitta ett kritiskt avstånd  $r_c$  mellan jonen och elektronen.

$$E_{kin} = E_{pot} \quad \Leftrightarrow \quad \frac{3}{2}kT = \frac{e^2}{4\pi\varepsilon_0\varepsilon_r r_c}$$

vilket ger att  $r_c \sim T^{-1}$ .

Vi får då ett tvärsnitt för spridning  $S = \pi r_c^2$  och vi ser att  $S \sim T^{-2}$ . När vi nu skriver ner vårt uttryck för  $\tau$  på samma sätt som tidigare finner vi att

$$\tau = \frac{1}{Sv_{th}N} \sim \frac{1}{T^{-2}T^{1/2}N} = \frac{T^{3/2}}{N}$$

#### 6.11. GENERATION OCH REKOMBINATION – LIVSTIDER173



Figur 6.23: Kommer elektronen innanför den kritiska radien  $r_c$  som innebär att kinetisk och potentiell energi är lika stora har vi en spridning.

där N är koncentrationen av joniserade störatomer, t.ex. joniserade acceptorer eller donatorer. Är N stort minskar  $\tau$ , dvs spridningsprocessen bidrar mer till den totala spridningen. Denna spridningsmekanism är effektivast och dominerande vid låga T.

Hur blir detta i en metall? Här har elektronerna mycket hög kinetisk energi jämfört med den i en halvledare och den kinetisk energin beror dessutom inte på temperaturen, utan ges av Fermienergin. De kommer alltså att behöva komma mycket nära jonen innan de påverkas av den (dvs  $r_c$  blir mycket liten). Vi får att  $r_c \sim v_F^{-2}$ . Vi får då ett mycket litet  $S (S \sim r_c^2 \sim v_F^{-4})$  - att hastigheten  $v_F$  är mycket högre än  $v_{th}$  i en halvledare hjälper inte - och alltså en mycket lång spridningstid med avseende på spridning mot joner om vi betraktar processen i en metall. Spridning mot atomära vibrationer kommer att dominera utom vid mycket låga temperaturer.

# 6.11 Generation och rekombination – livstider

Även om halvledaren befinner sig i jämvikt exciteras hela tiden elektroner från valensbandet till ledningsbandet (generation av elektronhål par) och elektroner i ledningsbandet rekombinerar med hål i valensbandet. Men i jämvikt måste både elektronkoncentrationen och hålkoncentrationen vara konstant. Rubbar vi jämvikten genom t ex belysning eller på elektrisk väg i en komponent kan lokala överskott av laddningsbärare skapas i halvledaren. Generations- och rekombinationsprocesser bestämmer då hur lång tid överskottet består, dvs överskottsladdningsbärarnas *livstid*. Vi kan nöja oss med att betrakta denna livstid som en given parameter när vi behöver den, en mera detaljerad diskussion finns dock i ett appendix i slutet av kompendiet



Figur 6.25: Diamanter är pga sitt stora bandgap genomskinliga. Kväveföroreningar ger dem gulaktig färg, grönt och blått ljus absorberas starkast. Bor fungerar som acceptor och ger blå färg vilket beror på absorptionens spektralfördelning och inte kan inses på enkelt sätt.

## 6.12 Djupa störställen

De störställen som används vid dopning ger upphov till grunda störnivåer, dvs nivåerna ligger nära bandkanterna. Detta är nödvändigt för att störatomerna ska kunna vara joniserade vid rumstemperatur. Det förekommer dock många andra störställen som leder till djupa nivåer närmare mitten av bandgapet. Dessa kan ha stor betydelse som klivstenar för generations- och rekombinationsprocesser när laddningsbärare ska ta sig över bandgapet. Men i allmänhet bidrar de inte med några rörliga laddningsbärare. Djupa störställen ger också färg åt många ädelstenar, se figurer 6.24, 6.25 och 6.26.

**Alexandrit**: ligger mellan rubin och smaragd.



Three-level Laser

Figur 6.24: Den första lasern utnyttjade det metastabila tillståndet hos Cr-störställen i rubin.



Figur 6.26: En Cr-jon ersätter en Al-jon i Al2O3. Ljus kan absorberas i övergångar till breddade högre tillstånd, rubin blir röd. Systemet relaxerar sedan snabbt till ett metastabilt tillstånd där det dröjer kvar tills övergången till grundtillståndet åtföljs av röd fluorescens

# 6.13 Kommentar till bundna tillstånd och bandstruktur

Donatorjonen är attraktiv för elektroner, acceptorjonen repulsiv. När vi betraktade bandstrukturen hos en atomär kedja nämnde vi att en attraktiv störatom kunde dra ut ett bundet tillstånd under bandet, en repulsiv störatom trycka upp ett tillstånd över bandet (se figur 6.27). Låter vi bandet representera ledningsbandet så får vi donatortillståndet under bandet. Om figurens band representerar valensbandet så skjuts acceptornivån upp ur bandet.



Figur 6.27: En störatom i en kristall kan leda till elektroniska tillstånd utanför banden. Där periodiciteten är störd är bandgapet inte längre förbjudet. Dessa nya tillstånd är lokaliserade kring störatomen.

#### **Rubin - Smaragd**

# Kapitel 7

# pn-övergången

I en halvledarkristall där det finns ett p-dopat och ett n-dopat område intill varandra uppstår ett utarmningsområde vid gränsytan. Här är koncentrationen av fria laddningsbärare låg jämfört med koncentrationen av dopatomer. Detta rymdladdningsområde leder till ett elektriskt fält som motverkar diffusion av majoritetsladdningsbärare till andra sidan övergången och en potentialbarriär uppstår. I termisk jämvikt är Ferminivån konstant i hela halvledaren. Vid framspänning ökar flödet av laddningsbärare genom övergången exponentiellt med den pålagda spänningen då elektroner och hål injiceras som minoritetsladdningsbärare på den motsatta sidan. I backriktningen har vi endast en liten läckström. Strukturen fungerar därmed som en likriktare. Vid tillräckligt höga backspänningar ökar strömmen kraftigt på grund av tunnling eller lavingenombrott.

## 7.1 pn-övergången i jämvikt

Halvledarnas betydelse beror på att vi kan förändra deras egenskaper, t ex genom dopning, och framförallt på att vi sedan kan bygga användbara elektroniska komponenter genom att modifiera olika områden på olika sätt. Här börjar vi med den enklaste, men mest grundläggande, komponentstrukturen: *pn-övergången*.

#### 7.1.1 Ett tankeexperiment

I ett tankeexperiment låter vi två homogent dopade områden – det ena n-typ, det andra p-typ – gränsa till varandra i en för övrigt homogen halvledare. Vi föreställer oss att dessa områden är åtskilda av en för laddningsbärare ogenomtränglig vägg. Väggen tillåter dock utbyte av värme så att temperaturen är densamma i bägge områden.

Bägge områdena är i termisk, diffusiv jämvikt var för sig och karakteriseras då av var sin Ferminivå (kemisk potential)  $E_F$ . Energiskillnaden mellan de två Ferminivåerna är  $e\psi_0$ , där  $\psi_0$  är diffusionsspänningen (built-in potential, kontaktpotential eller inbyggd spänning).



Figur 7.1: pn-övergången innan jämvikt har etablerats.  $E_{Fi}$  är den intrinsiska Ferminivån, som ligger ungefär mitt i bandgapet.



Figur 7.2: Elektroner och hål diffunderar och rekombinerar.



Figur 7.3: Vid övergången finns ett utarmningsområde utan fria laddningsbärare. Här finns således en nettoladdning med laddningstäthet  $\rho$ . p-sidan är negativt laddad nära gränsytan, n-sidan positivt.  $\oplus$  indikerar positiva donatorjoner (alltså inte hål). Negativa acceptorjoner symboliseras av  $\ominus$ . Vi antar vidare att temperaturen är sådan att att alla störställen är joniserade, dvs  $n = N_D$  i n-typ området,  $p = N_A$  i p-typ området. När vi nu har två olika områden indikerar vi vilket område det gäller med ett index n eller p. Elektronkoncentrationen n på p-sidan skriver vi som  $n_p$ . Vid jämvikt lägger vi till ett index 0, dvs  $n_{p0}$ .

När vi, i vårt tanke<br/>experiment, tar bort väggen i figur 7.1 förväntar vi oss enligt statistisk fysik att partiklar – elektroner – kommer att röra sig från området med högre<br/> $E_F$  till området med lägre  $E_F$ . Detta sker genom att elektroner diffunderar från n-sidan med hög<br/> elektronkoncentration till p-sidan med låg elektronkoncentration. Samtidigt diffunderar då hålen i motsatt riktning från p- till n-<br/>sidan. Nära övergången uppvisar ju både elektroner och hål stora koncentrationsgradienter som driver diffusionen. I n-området ges elektronkoncentrationen av n-dopningen; vi kan exempelvis ha

$$n_n = n_{n0} = N_D = 10^{22} \text{ m}^{-3}$$

medan elektronkoncentrationen i p-området är

$$n_p = n_{p0} = \frac{n_i^2}{N_A} \sim 10^9 \text{ m}^{-3} \text{ om } N_A = 10^{23} \text{ m}^{-3}$$

eftersom (i kisel)  $n_i \sim 10^{16} \text{ m}^{-3}$  vid rumstemperatur. På väg från nområdet till p-området varierar laddningsbärarkoncentrationen alltså typiskt med ca 10 tiopotenser.

Åter till vårt tanke experiment. I p-typ området rekombinerar elektronerna, som nu har blivit många fler än vad termisk jämvikt genom massverkans lag föreskriver, med hål. I n-typ området rekombinerar hålen med elektroner. Rekombination är ett sätt för systemet att gå mot termisk jämvikt. I området kring övergången minskar därmed koncentrationen av majoritetsladdningsbärare, vi får ett *utarmnings-* eller *bristområde*. De joniserade störatomerna (donatorer och acceptorer) finns dock kvar och eftersom de i utarmningsområdet inte neutraliseras av fria laddningsbärare får vi nettoladdningar. Utarmningsområdet kallas därför även *rymdladdningsområde*. Här bestäms laddningstätheten  $\rho$  (enhet C/m<sup>3</sup>) väsentligen av störatomernas koncentrationer,

 $\rho = eN_D^+ = eN_D$  på n-sidan

 $\operatorname{och}$ 

 $\rho = -eN_A^- = -eN_A$  på p-sidan

där vi har antagit fullständig jonisation.

Utanför utarmningsområdet förblir halvledaren neutral. De joniserade störatomerna, positiva donatorjoner såväl som negativa acceptorjoner, neutraliseras av fria laddningsbärare. Vi betraktar nu figur 7.3 och ser att vi har ett positivt laddat område intill ett negativt. Här uppstår ett elektriskt fält med riktning från det positivt laddade området till det negativa. Detta kan också uttryckas med hjälp av Gauss lag; laddningsfördelningen  $\rho$  ger upphov till ett elektriskt fält  $\mathscr{E}$ :

$$\frac{d\mathscr{E}}{dx} = \frac{\rho}{\varepsilon_s}.$$

Detta fält är riktat så att det motverkar diffusionen av laddningsbärare genom övergången, se figur 7.3. En viss fältstyrka bromsar diffusionen helt och i en sådan situation försvinner *netto*flödet av partiklar genom övergången. Övergången är då i termisk jämvikt och vårt tankeexperiment är färdigt. Vi har nu klart för oss pnövergångens struktur och kan analysera strukturen och funktionen mer i detalj.

Med det elektriska fältet  $\mathscr{E}$  förknippar vi en elektrostatisk potential  $\psi(x)$  enligt

$$\mathscr{E} = -\frac{d\psi}{dx}$$

E-fältet pekar från n- mot p-området (se figur 7.3), fältlinjerna startar i positiva donatorjoner och slutar i negativa acceptorjoner. Eftersom då  $\mathscr{E} < 0$  måste potentialens derivata vara positiv,

$$\frac{d\psi}{dx} > 0$$

Potentialen ökar alltså när vi går genom utarmningsområdet från p- till n-sidan (figur 7.4). I de neutrala områdena är det elektriska fältet noll och därmed blir potentialen konstant. Vi känner ännu inte den elektriska potentialens form i utarmningsområdet, men den bör vara kontinuerlig mellan  $\psi = 0$  på p-sidan och  $\psi = \psi_0$  på n-sidan.

En elektrons potentiella energi blir x-beroende när vi placerar den i en elektrostatisk potential  $\psi(x)$ , dvs

$$\Delta E_{pot}(x) = (-e) \cdot \psi(x).$$

Bandkanterna representerar potentiell energi, och vi kan skissa pnövergångens banddiagram om vi adderar  $\Delta E_{pot}(x) = (-e) \cdot \psi(x)$ till den ursprungliga (flata) bandstrukturen för att beskriva elektronernas totala energi (figur 7.5). Hur potentialen i detalj varierar i utarmningsområdet tar vi snart reda på.

I termisk jämvikt, då Ferminivån är konstant genom strukturen, ser vi att  $e\psi_0$  från figur 7.1 återfinns som avståndet mellan ledningsbandskanterna (eller valensbandskanterna) på p- respektive n-sidan Vi väljer gränsytan mellan n- och p-området att ligga i yz-planet. Då har det elektriska fältet bara en x-komponent,  $\mathscr{E} = \mathscr{E}_x$ , och Gauss lag blir  $\nabla \cdot \mathbf{E} = \frac{d\mathscr{E}}{dx} = \frac{\rho}{\varepsilon_s}$  $\varepsilon_s = \varepsilon_r \varepsilon_0$ , där  $\varepsilon_0$  och  $\varepsilon_r$  är permittiviteten i vakuum samt relativa permittiviteten för en specifik halvledare.



Figur 7.4: Den elektriska potentialen kring övergången.  $x_n$  och  $-x_p$  markerar utarmningsområdets gränser.



Figur 7.5: Banddiagrammet för pn-övergången i termisk jämvikt.

i figur 7.5. Detta inses enklast genom geometriskt resonemang. Betrakta figur 7.1. Vi föreställer oss att n-sidan ska flyttas neråt relativt p-sidan tills Ferminivåerna överensstämmer. n-sidan måste då sänkas med  $e\psi_0$ . Bandkanter följer med ner ( $E_F$  måste ligga på samma avstånd från bandkanterna som innan - dessa avstånd ges vid full jonisation av dopningen) och ledningsbandskanten på n-sidan sänks alltså med  $e\psi_0$  relativt ledningsbandskanten på p-sidan. Vi kan då dra slutsatsen att skillnaden i elektrisk potential  $\psi(x)$  mellan p- och n-sidan måste vara  $\psi_0$ , som indikeras i figur 7.4.

**Exempel:** Hur stor är  $\psi_0$ ? Det beror på hur hårt de båda sidorna är dopade, då detta bestämmer Ferminivåernas lägen, men genom att betrakta figur 7.1 eller 7.5 inser vi att  $e\psi_0$  alltid är mindre än bandgapet (iallafall i de strukturer vi beaktar). Vi använder våra exempeldopkoncentrationer;

$$N_A = 10^{23} \text{ m}^{-3} \text{ och } N_D = 10^{22} \text{ m}^{-3},$$

den intrinsiska laddningsbärarkoncentrationen

$$n_i = 1.5 \cdot 10^{16} \text{ m}^{-3} \text{ vid } T = 300 K$$

och permittivitetstalet  $\varepsilon_r = 11,7.$ 

Med hjälp av uttrycket

$$\psi_0 = \frac{kT}{e} \ln\left(\frac{N_A N_D}{n_i^2}\right),$$

som härleds i en övningsuppgift, får vi  $\psi_0$  till 0,75 V, och därmed  $e\psi_0$  till 0,75 eV.

De fria laddningsbärarna har förstås inte helt försvunnit från utarmningsområdet och områdets kanter är inte fullt så skarpa som figur 7.3 antyder. Men vi ser i figur 7.6, som visar resultat vi kommer fram till nedan, att elektron- och hålkoncentrationerna i nästan hela utarmningsområdet är mindre än säg en procent av dopkoncentrationen. För att beräkna potentialen kan vi då med god approximation försumma de fria laddningsbärarna.



Figur 7.6: Banddiagram och laddningsbärarkoncentrationer i en pn-övergång i termisk jämvikt. Majoritetsladdningsbärar-koncentrationerna  $p_{p0}$  och  $n_{n0}$  ges direkt av dopningen, och minoritetsladdningsbärar-koncentrationerna  $n_{p0}$  och  $p_{n0}$  fås sedan genom massverkans lag.
#### 7.1.2 Detaljerad härledning

I förra avsnittet resonerade vi oss fram till en översiktlig insikt om hur det elektriska fältet  $\mathscr{E}$  och elektriska potentialen  $\psi$  bör variera i pn-övergången. Nu ska vi hitta kvantitativa uttryck för dessa storheter.

I termisk jämvikt är Ferminivån konstant och vi har ingen nettotransport av laddningsbärare. Det elektriska fältet från laddningen i rymdladdningsområdet ger upphov till potentialbarriären som hindrar elektroner och hål från att diffundera till den motsatta sidan. Stora koncentrationsgradienter försöker alltså driva diffusionen genom utarmningsområdet men kompenseras av starka elektriska fält som håller partiklarna på plats. I terminsk jämvikt tar de två effekterna exakt ut varandra.

För att bestämma den elektriska potentialen, det elektriska fältet och utarmningsområdets vidd i jämvikt utgår vi ifrån två antaganden som gör det möjligt att lösa problemet för hand.

1) Abrupt övergång. Dopkoncentrationerna  $N_A$  och  $N_D$  på var sin sida om övergången antas vara konstanta och ändrar sig abrupt vid gränsytan mellan n- och p-typ-områdena.

2) Utarmningsapproximationen. När vi bestämmer den elektriska potentialen  $\psi$  antar vi att det inte finns några fria laddningsbärare i utarmningsområdet. Koncentrationerna av fria laddningsbärare försvinner som vi redan noterat naturligtvis inte helt, men antas vara mycket mindre än dopkoncentrationerna  $N_A$  och  $N_D$ , som därmed ensamma bestämmer laddningstätheten och, i förlängningen, den elektriska potentialen i utarmningsområdet.

Vi utgår nu från laddningstätheten  $\rho$  i utarmningsområdet. I de neutrala områdena utanför utarmningsområdet är det elektriska fältet  $\mathscr{E} = 0$ . För att bestämma det elektriska fältet och potentialen startar vi med Gauss lag som relaterar det elektriska fältet till laddningstätheten genom

$$\frac{d\mathscr{E}}{dx} = \frac{\rho}{\varepsilon_s.}$$

För potentialen gäller

$$\frac{d\psi}{dx} = -\mathscr{E}.$$

Dessa två ekvationer kan vi sammanfatta i Poissons ekvation

$$\frac{d^2\psi}{dx^2} = -\frac{\rho}{\varepsilon_s}$$



Figur 7.7: Laddningstätheten i utarmnings- eller rymdladdningsområdet hos pn-övergången i jämvikt.

Vi integrerar nu Poissons ekvation i område I i figur 7.7:

$$\frac{d^2\psi}{dx^2} = -\frac{\rho}{\varepsilon_s} = \frac{eN_A}{\varepsilon_s}; \quad \mathscr{E}(x \le -x_p) = 0$$

Observera att  $x_p$  (och  $x_n$ ) inte har bestämts än.

$$\int_{-x_p}^{x} \frac{d^2 \psi}{dx'^2} dx' = \frac{d\psi}{dx'} \Big|_{x} - \frac{d\psi}{dx'} \Big|_{-x_p}$$
$$= \int_{-x_p}^{x} \frac{eN_A}{\varepsilon_s} dx' = -\mathscr{E}(x) + \mathscr{E}(-x_p) = \frac{eN_A}{\varepsilon_s} (x+x_p)$$

och genom att utnyttja att $\mathscr{E}(-x_p)=0$ får vi $\mathscr{E}(x)$ till

$$\mathscr{E}(x) = -\frac{d\psi}{dx} = -\frac{eN_A}{\varepsilon_s}(x+x_p). \tag{7.1}$$

Vi finner  $\psi(x)$  genom att integrera vårt uttryck för det elektriska fältet:

$$\int_{-x_p}^{x} \frac{d\psi}{dx'} dx' = \psi(x) - \psi(-x_p) = \int_{-x_p}^{x} \frac{eN_A}{\varepsilon_s} (x'+x_p) dx' = \frac{eN_A}{2\varepsilon_s} (x+x_p)^2$$
$$\psi(x) = \psi(-x_p) + \frac{eN_A}{2\varepsilon_s} (x+x_p)^2. \tag{7.2}$$

Nu tittar vi på område II:

$$\frac{d^2\psi}{dx^2} = -\frac{\rho}{\varepsilon_s} = -\frac{eN_D}{\varepsilon_s}; \quad \mathscr{E}(x \ge x_n) = 0$$

$$\begin{aligned} \int_{x}^{x_{n}} \frac{d^{2}\psi}{dx'^{2}} dx' &= \left. \frac{d\psi}{dx'} \right|_{x_{n}} - \left. \frac{d\psi}{dx'} \right|_{x} \\ &= \left. - \int_{x}^{x_{n}} \frac{eN_{D}}{\varepsilon_{s}} dx' = -\mathscr{E}(x_{n}) + \mathscr{E}(x) = -\frac{eN_{D}}{\varepsilon_{s}}(x_{n} - x) \end{aligned}$$



Figur 7.8: a) Det elektriska fältet i pn-övergången.  $\mathscr{E}_m$  är absolutvärdet av den maximala fältstyrkan i övergången. b) Elektriska potentialens variation i pn-övergången.

och genom att utnyttja att  $\mathscr{E}(x_n) = 0$  får vi  $\mathscr{E}(x)$  till

$$\mathscr{E}(x) = -\frac{d\psi}{dx} = \frac{eN_D}{\varepsilon_s}(x - x_n).$$
(7.3)

Vi finner  $\psi(x)$  i område II genom att integrera vårt uttryck för det elektriska fältet:

$$\int_{x}^{x_n} \frac{d\psi}{dx'} dx' = \psi(x_n) - \psi(x) = -\int_{x}^{x_n} \frac{eN_D}{\varepsilon_s} (x' - x_n) dx' = \frac{eN_D}{2\varepsilon_s} (x - x_n)^2.$$
$$\psi(x) = \psi(x_n) - \frac{eN_D}{2\varepsilon_s} (x - x_n)^2 \tag{7.4}$$

Det elektriska fältet måste vara kontinuerligt vid x = 0:

$$\mathscr{E}(0^{-}) = -\frac{eN_A}{\varepsilon_s} x_p = \mathscr{E}(0^{+}) = -\frac{eN_D}{\varepsilon_s} x_n = -\mathscr{E}_m$$
(7.5)

där  $\mathscr{E}(0^-)$  ligger infinitesimalt till vänster om x = 0 och  $\mathscr{E}(0^+)$  infinitesimalt till höger. Indexet 'm' står för 'max' då  $\mathscr{E}$  når sin maximala styrka precis vid övergången, dvs vid x = 0 i våra figurer.

Ur kontinuitetskravet (7.5) för  $\mathscr{E}$  ser vi att

$$N_A x_p = N_D x_n. aga{7.6}$$

Samma samband följer från *laddningsneutralitet*:  $-eN_Ax_p$  är laddningen (negativ) per ytenhet på övergångens p-sida och  $eN_Dx_n$  den positiva laddningen per ytenhet på n-sidan. Finns det positiv nettoladdning på ett ställe måste det finnas negativ nettoladdning på ett annat. Halvledaren i sin helhet förblir neutral.

Det elektriska fältet är negativt i utarmningsområdet (figur 7.8a), dvs pekar mot p-sidan. Eftersom integration av det elektriska fältet i figur 7.8a ska ge  $\psi_0$  i figur 7.8b;

Kom ihåg:

 $\frac{d\psi}{dx}$ 

är arean i figur 7.8a lika med diffusionsspänningen  $\psi_0$ .

Med uttrycken för  $\mathscr{E}_m$  (ekv. 7.5), dvs

$$\mathscr{E}_m = \frac{eN_D}{\varepsilon_s} x_n = \frac{eN_A}{\varepsilon_s} x_p$$

kan vi ta fram en relation mellan  $x_n$ ,  $x_p$  och  $\psi_0$ :

$$\psi(x_n) - \psi(x_p) = -\int_{-x_p}^{x_n} \mathscr{E} dx = \frac{\mathscr{E}_m x_p}{2} + \frac{\mathscr{E}_m x_n}{2}$$
$$= \frac{eN_A}{\varepsilon_s} \frac{x_p^2}{2} + \frac{eN_D}{\varepsilon_s} \frac{x_n^2}{2} = \psi_0$$

Slutligen vill vi bestämma utarmningsområdets vidd $w=x_p+x_n.$  Ovan fann vi

$$\psi_0 = \frac{eN_A}{\varepsilon_s} \frac{x_p^2}{2} + \frac{eN_D}{\varepsilon_s} \frac{x_n^2}{2} \quad \text{och} \quad \mathscr{E}_m = \frac{eN_A}{\varepsilon_s} x_p = \frac{eN_D}{\varepsilon_s} x_m$$

Vi får

$$\psi_0 = \frac{\mathscr{E}_m}{2}(x_p + x_n) = \frac{\mathscr{E}_m}{2}w = \frac{eN_D}{2\varepsilon_s}\left(\frac{N_D}{N_A} + 1\right)x_n^2,$$

vilket ger

$$\psi_0 = -\int_{-x_p}^{x_n} \mathscr{E} dx$$

Diffusionsspänningen  $\psi_0$  har inte bestämts än. Ett typiskt värde i kisel är  $\psi_0 \sim 0.8$  V.

$$x_n = \sqrt{\frac{2\varepsilon_s\psi_0}{e}}\frac{1}{N_D + N_A}\sqrt{\frac{N_A}{N_D}}.$$

Ur laddningsneutralitet (ekv. 7.6) följer att  $x_p = (N_D/N_A)x_n$ 

$$x_p = \frac{N_D}{N_A} x_n = \sqrt{\frac{2\varepsilon_s \psi_0}{e}} \frac{1}{N_D + N_A} \sqrt{\frac{N_D}{N_A}},$$

och slutligen

$$w = \sqrt{\frac{2\varepsilon_s\psi_0}{e}\frac{1}{N_D + N_A}} \left(\sqrt{\frac{N_A}{N_D}} + \sqrt{\frac{N_D}{N_A}}\right)$$

Ett alternativt uttryck för w är (verifiera!)

$$w = \sqrt{\frac{2\varepsilon_s\psi_0}{e}\left(\frac{N_D + N_A}{N_D N_A}\right)} \tag{7.7}$$

**Exempel:** Vi beräknar w och  $\mathscr{E}_m$  i Si för dopkoncentrationerna

 $N_A = 10^{23} \text{ m}^{-3} \text{ och } N_D = 10^{22} \text{ m}^{-3},$ 

samt  $n_i = 1.5 \cdot 10^{16} \text{ m}^{-3}$  vid T = 300 K och  $\varepsilon_r = 11.7$ .

Vi bestämde tidigare  $\psi_0$  till 0,75 V för de givna  $N_A$  och  $N_D,$  och finner nu

$$w = \sqrt{\frac{2\varepsilon_s \psi_0}{e} \frac{1}{N_D + N_A}} \left( \sqrt{\frac{N_A}{N_D}} + \sqrt{\frac{N_D}{N_A}} \right) = 0.33 \ \mu \text{m}$$

 $\operatorname{och}$ 

$$\mathscr{E}_m = rac{2\psi_0}{w} = 4.5\cdot 10^6 ~\mathrm{V/m}$$

#### Sammanfattning: pn-övergång i termisk jämvikt

I utarmningsområdet finns mycket få fria laddningsbärare och laddningen ges väsentligen av de joniserade dopatomerna. Utanför utarmningsområdet är nettoladdningen noll.

Rymdladdningstätheten  $\rho$  i pn-övergången. Vi gör approximationen att det finns en nettoladdning bara i utarmningsområdet.

Laddningstätheten  $\rho$  ger upphov till ett elektriskt fält  $\mathscr{E}$ .

Det elektriska fältet  $\mathscr{E}$  kan relateras till en elektrisk potential  $\psi$  - fältet är (minus) gradienten av potentialen.

Den elektriska potentialen  $\psi$  ger potentiell energi  $E_{pot}$  för en laddning q via  $E_{pot}(x) = q\psi(x)$ . Bandkanterna visar potentiell energi för elektroner, alltså ges de av  $-e\psi(x)$ . En elektron vid ledningsbandskanten på nsidan måste tillföras energi för att befinna sig vid ledningsbandskanten på p-sidan den har där högre potentiell energi.



#### 7.1.3 Ström vid jämvikt?

Under de förutsättningar vi diskuterar i den här kursen inträffar jämvikt när den yttre pålagda spänningen U = 0 och när vi inte har någon optisk excitation. I jämvikt kan det inte flyta någon ström genom pn-övergången. Men detta gäller för nettoströmmarna. De enskilda laddningsbärarna är i ständig rörelse, vissa rör sig åt vänster, andra åt höger utan att det sker någon nettotransport av varken elektroner eller hål. Elektroner på n-sidan med koncentration  $n_n = N_D$  vid fullständig jonisation har i medel en termisk (kinetisk) energi  $3kT/2 \approx 0.038$  eV vid T = 300 K. De flesta har då energier som är för små för att de ska kunna passera barriären över till p-området. I banddiagrammet på föregående sida ser vi dock att det finns ett fåtal elektroner på n-sidan som har tillräckligt stor termisk energi för att kunna ta sig över barriären, förutsatt att de rör sig i rätt riktning. Samtidigt finns det ett fåtal elektroner på p-sidan som, om de råkar diffundera till utarmningsområdet, sveps över till n-sidan av det starka elektriska fältet. Detta innebär att det hela tiden finns elektroner som rör sig genom utarmningsområdet i bägge riktningarna. Men vi vet att nettoströmmen som uppstår på grund av dessa flöden genom övergången måste vara noll. Motsvarande argument gäller för hålen, deras nettoström är också noll.

Kom ihåg:  $kT \approx 1/40$  eV vid T = 300 K  $e\psi_0 \sim 0.8$  eV

# 7.2 pn-övergången med pålagd spänning

Vi börjar med en kort kvalitativ introduktion för att få en översiktlig förståelse för pn-övergången som komponent.

#### 7.2.1 pn-övergången som likriktare

pn-övergången är en likriktare, dvs en komponent som leder ström för en polaritet på den pålagda spänningen U men inte för motsatt polaritet. Köper man en pn-övergång i en elektronikhandel kallas den ofta diod.

När pn-övergången är i termisk jämvikt, figur 7.10, finns inga nettoflöden av laddningsbärare genom bristområdet. Potentialbarriären i bristområdet hindrar de flesta elektroner från n-området från att ta sig över till p-området. Barriärens höjd är  $e\psi_0$ , där  $\psi_0$  är diffusionsspänningen.

Några få elektroner på n-sidan har dock tillräckligt hög energi för ta sig förbi barriären (de följer en termisk fördelning), och på p-sidan finns också några få elektroner som kan svepas över till n-sidan av det elektriska fältet i utarmningsområdet. Dessa två flöden (grå pilar i fig 7.10) tar precis ut varandra, ingen ström flyter genom pn-övergången. Helt analoga argument gäller för hålen som vi för enkelhets skull utelämnar här.

Nu lägger vi en spänning över pn-övergången, se figur 7.12, där vi sätter spänningens pluspol på p-sidan och minuspol på n-sidan. Då ökar elektronernas potentiella energi på n-sidan med

 $\Delta E_{pot} = -e(-U) = eU \quad (U > 0)$ 

relativt p-sidan. Därmed minskar barriärhöjden på motsvarande sätt. Elektronerna följer en termisk fördelning i ledningsbandet och



Figur 7.10: pn-övergång i termisk jämvikt



Figur 7.9: Krets-symbolen för en likriktare.





189

Figur 7.12: Vänster: då pn-övergången är framspänd (U > 0) minskar barriärhöjden. Höger: då pn-övergången är backspänd (U < 0) ökar barriärhöjden.

betydligt fler elektroner från n-sidan kan nu passera barriären medan elektronflödet från p-sidan förblir detsamma. Övergången är *framspänd* och bristområdets vidd minskar. I dioden flödar nu en nettorström från p- till n-sidan (elektronerna rör sig från n- till p-sidan).

Byter vi polaritet på spänningen, U < 0, minskar elektronernas potentiella energi på n-sidan relativt p-sidan. Barriärhöjden ökar då och endast det lilla elektronflödet från p-sidan (som ger "läckströmmen" genom pn-övergången eller dioden) finns kvar. Barriärhöjden spelar ingen roll för detta flöde - det bestäms av tillgången av elektroner på p-sidan vilken bestäms av massverkans lag. Läckströmmen blir liten och konstant i en första approximation. Övergången är *backspänd* och bristområdets vidd har ökat.

I nästa avsnitt kommer vi att diskutera mekanismerna bakom laddningstransport i pn-övergången mer i detalj, men redan nu kan vi använda våra insikter till att ta fram ett uttryck som beskriver strömmen genom pn-övergången som funktion av spänningen. Strömmen I fås från flödet F genom att multiplicera flödet med elektronladdningen. För elektronerna blir därmed strömmen motriktad flödet. Vi kommer dock i det följande att diskutera flöden och inte strömmar, då figurerna visar flöden. Även fortsatt koncentrerar vi oss på elektronernas rörelse, men samma diskussion kan föras för hålen.

I figur 7.10 är den yttre, pålagda, spänningen U = 0. De grå pilarna i figuren som representerar elektronflöde är lika stora. Vi kallar vart och ett av dessa flöden för  $F_0$  (motsvarande strömmar kan vi kalla  $I_0$ ), och totala strömmen/flödet är noll.

I figur 7.12a är pn-övergången framspänd. Flödet av elektroner från p-sida till n-sida är oförändrat, och ger fortfarande upphov till en



Figur 7.11: Ström-spänningskurvan för en pn-övergång. I en framspänd övergång kan det gå en stor ström.



Figur 7.13: I en backspänd övergång finns en liten läckström.

ström  $I_0$ , medan flödet från höger till vänster har ökat kraftigt. Vi vet att elektroner besätter tillstånd i ledningsbandet enligt en exponentiellt avtagande "Fermi-svans",  $e^{-E/kT}$ . När vi lyfter n-sidan relativt p-sidan kan vi anta att exponentiellt fler elektroner har tillräcklig energi för att diffundera till p-sidan. Flödet dessa elektroner ger borde ges av

$$F_{fram} = F_0 e^{eU/kT},$$

då vi har minskat barriärhöjden med energin eU. Totala flödet borde då bli (flödet framåt) - (flödet bakåt):

$$F(U) = F_0 e^{eU/kT} - F_0 = F_0 (e^{eU/kT} - 1).$$

eller, uttryckt som ström,

$$I(U) = I_0 e^{eU/kT} - I_0 = I_0 (e^{eU/kT} - 1)$$

Ekvationen vi nu har resonerat oss fram till kallas *ideala diodekvationen*. Vi kommer strax att göra en mer utförlig härledning som dessutom leder fram till en insikt om de faktorer som påverkar  $I_0$ .

Vid backspänning, U < 0, blir strömmen  $I = -I_0$  som också ses i figur 7.12b.

Vid tillräckligt hög backspänning kan genombrott (diskuteras mer senare) ske i pn-övergången, vilket resulterar i en stor backström.

Efter denna kvalitativa introduktion ska vi titta närmare på de antaganden vi redan har gjort och de ytterligare som vi behöver göra för att kunna verifiera våra idéer om ström i pn-övergången.

#### 7.2.2 Var hamnar spänningsfallet?

Vi lägger nu en framspänning, exempelvis U = 0.5 V, över övergången, med polaritet (+) på p-sidan och (-) på n-sidan.

De neutrala områdena har hög konduktivitet (många fria laddningsbärare) i förhållande till utarmningsområdet (med mycket färre fria laddningsbärare). Spänningsfallet borde då till allra största delen hamna över utarmningsområdet.

För att försöka övertyga oss om att så är fallet betraktar vi en pnövergång i kisel med några rimliga värden på strömtäthet, spänning och dopnivåer. Vi antar att en viss framspänning, säg U = 0.5 V, ger upphov till en strömtäthet på  $J = 10^4$  A/m<sup>2</sup> i komponenten.

Vidare antar vi följande data: på p-sidan är hålmobiliteten  $\mu_h = 450 \cdot 10^{-4} \text{ m}^2/\text{Vs}$  och  $N_A = 10^{23} \text{m}^{-3}$ . På n-sidan är elektronmobiliteten  $\mu_e = 1350 \cdot 10^{-4} \text{ m}^2/\text{Vs}$  och  $N_D = 10^{22} \text{m}^{-3}$ . Vi antar full jonisering.



Figur 7.14: Om en pn-övergång backspänns mycket kan det orsaka en stor backström.

#### Ideala diodekvationen

Antagande 1: Hela den pålagda spänningen ligger över utarmningsområdet. Strömtätheten är densamma i hela komponenten. Om vi antar att strömmen i de neutrala delarna utgörs av majoritetsladdningsbärarnas driftström, kan vi beräkna det elektriska fält som behövs i dessa områden för att ge den önskade strömtätheten. Vi har strömtätheten

$$J_h = \sigma \mathscr{E} = e\mu_h p_{p0} \mathscr{E}$$

på p-sidan samt

$$J_e = \sigma \mathscr{E} = e\mu_e n_{n0} \mathscr{E}$$

på n-sidan. Vi beräknar det elektriska fältet i de neutrala områdena (utanför bristområdet) på p- respektive n-sidan och får  $\mathscr{E} = 13.9$  V/m på p-sidan och  $\mathscr{E} = 46$  V/m på n-sidan.

Om vi nu vidare antar att de neutrala områdena är  $100\mu$ m långa, leder detta med de ovan beräknade fältstyrkorna till ett spänningsfall över det neutrala p-området på 1,39 mV och över det neutrala n-området på 4,6 mV. Spänningen totalt över pn-övergången satte vi till 0,5 V, och vår räkning indikerar alltså att all pålagd spänning utom några få procent hamnar över utarmningsområdet. Framöver kommer vi därför att förutsätta att spänningsfallet hamnar över bristområdet.

#### 7.2.3 Bidrar minoritetsladdningsbärarna till strömmen i de neutrala områdena?

Vi har nyss bestämt de elektriska fälten i de neutrala områdena som krävdes för att majoritetsladdningsbärarnas driftströmmar skulle ha ett visst givet värde. Hur stora är nu minoritetsladdningsbärarnas driftströmmar för dessa fält? Långt borta från övergången i p-sidans neutrala område blir *elektron*strömtätheten med fältet  $\mathscr{E} = 13.9 \text{ V/m}$ 

$$J_e = 4.87 \cdot 10^{-7} \mathrm{A/m^2}$$

och hålströmtätheten på n-sidan med fältet  $\mathscr{E} = 46 \text{ V/m}$ 

$$J_h = 1.6 \cdot 10^{-10} \mathrm{A/m^2}$$

Dessa strömmar är helt försumbara jämfört med den strömtäthet på  $J = 10^4 \text{ A/m}^2$  som vi har tagit som exempel här, vilket beror på att koncentrationerna av minoritetsladdningsbärare är ungefär 12 tio-potenser mindre än majoritetsladdningsbärarkoncentrationerna. Elektronernas driftström på p-sidan långt borta från övergången är då försumbar i förhållande till deras driftström långt borta på Verifiera  $J_e$  på p-sidan och  $J_h$  på n-sidan. Du behöver först beräkna minoritetsladdningsbärarkoncentrationerna.



Figur 7.16: Fälttriangelns area är bestämd av potentialskillnaden vid övergången. Vid framspänning minskar det elektriska fältet i hela utarmningsområdet.

n-sidan. Detta innebär att av alla de elektroner som kommer in från kontakten på n-sidan är det nästan inga som kommer fram till kontakten på andra sidan av pn-övergången - på p-sidan bärs ju strömmen så gott som uteslutande av hålen (majoritetsladdningsbärarna). Vi kan säga att alla fria laddningsbärare rekombinerar någonstans kring övergången. Detta ger oss en bild av pn-övergången där majoritetsladdningsbärare flyter in från kontakterna mot övergången för att rekombinera någonstans kring övergången.

#### 7.2.4 Ideala diodekvationen

Efter dessa inledande kommentarer ska vi härleda den så kallade ideala diodekvationen på ett sätt som ger mer förståelse för mekanismerna bakom strömtransport i pn-övergången. Ideala diodekvationen baseras på ett antal antaganden eller förenklingar, och man kan sedan ta fram diverse korrektioner till denna första ideala beskrivning (vi ska titta kort på ett par sådana korrektioner, utan detaljer).

Framspänningen över utarmningsområdet reducerar barriärhöjden från  $\psi_0$  till  $\psi_0 - U$ . Övergångens n-sida ökar i energi med  $\Delta E_{pot} = eU$  relativt p-sidan. Detta gör att fler laddningsbärare kan ta sig förbi potentialbarriären.

Ett spänningsfall U som reducerar potentialskillnaden mellan de två neutrala områdena leder till en ändring (minskning) av det elektriska fältet i utarmningsområdet. Fältet blir, vid framspänning, svagare än det ursprungliga fältet vi hade i termisk jämvikt och det blir därmed lättare därmed för majoritetsladdningsbärare att ta sig genom utarmningsområdet. Figur 7.16 visar ändringen av det elektriska fältet  $\mathscr{E}$  i utarmningsområdet.

Att diskutera effekten fältet har på laddningsbärarna eller att studera potentialens effekt är förstås bara två olika sätt att se på samma

Av de majoritetsladdningsbärare som rör sig fram till och genom utarmningsområdet är det inga som överlever resan fram till kontakten på motsatta sidan, alla rekombinerar på vägen.



Figur 7.15: Den pålagda framspänningen U har motriktad polaritet jämfört med diffusionsspänningen  $\psi_0$ , vilket leder till att barriärhöjden reduceras.

fenomen. Vi ser också i figur 7.16 att utarmningsområdets vidd w minskar. Detta följer även ur det uttryck vi härledde ovan;

$$w = \sqrt{\frac{2\varepsilon_s\psi_0}{e} \left(\frac{N_D + N_A}{N_D N_A}\right)}$$

där vi nu får ersätta  $\psi_0 \mod \psi_0 - U$  så att vi får

$$w \sim \sqrt{\psi_0 - U}.$$

När barriärhöjden sänks och det elektriska fältet minskar rubbas balansen vi hade i termisk jämvikt mellan diffusionen i utarmningsområdet som drivs av stora koncentrationsgradienter och det elektriska fältet som försöker hålla laddningsbärarna på avstånd på var sin sida om övergången.

Detta måste då resultera i ett nettoflöde av laddningsbärare genom utarmningsområdet, vilket vi redan har konstaterat inledningsvis, och därmed elektron- och hålströmmar  $J_e$  och  $J_h$ . Figur 7.17 visar att antalet elektroner och hål som kan ta sig över barriären har ökat kraftigt på grund av framspänningen U.

Vi antar nu att vi inte har någon netto-rekombination eller generation i utarmningsområdet utan att strömmar av hål eller elektroner är konstanta i hela utarmningsområdet. Lika många elektroner (hål) som passerar in i utarmningsområdet vid  $x_n(-x_p)$  passerar ut på andra sidan. I rymdladdningsområdet är nettorekombination en klar möjlighet, både elektron och hålkoncentrationer har ökat under framspänningen, men vi bortser alltså från denna möjlighet tills vidare. Vi ska dock återkomma till detta i slutet av kapitlet.

Vi tänker oss att dessa, i utarmningsområdet konstanta, strömmar kommer hela vägen från kontakterna, genom respektive neutrala område, genom utarmningsområdet och därifrån fortsätter in i de neutrala områdena på motsatta sidan (se figur 7.17). I det stationära tillstånd vi betraktar är alla laddningstätheter konstanta med avseende på tid.

Eftersom strömmarna i figur 7.17 leder till ett konstant inflöde av elektroner/hål i de neutrala områdena på övergångens motsatta sida blir det en kraftig ökning av elektronkoncentrationen i  $-x_p$  och av hålkoncentrationen i  $x_n$ . I bägge punkterna ökar alltså minoritetsladdningsbärarkoncentrationen. Vi säger att minoritetsladdningsbärare *injiceras* i de neutrala områdena vid dessa punkter.

Hur stora bli då minoritetsladdningsbärarkoncentrationerna i punkterna  $-x_p$  och  $x_n$  på grund av denna injektion? I p-området är

#### Ideala diodekvationen Antagande 2: Ingen netto-rekombination eller generation i utarmningsområdet.



Figur 7.17: De vita pilarna markerar flöden av laddningsbärare genom utarmningsområdet. De svarta pilarna visar strömmen (inte partikelflödena). eU är energiskillnaden mellan Ferminivåerna i de neutrala områdena långt från övergången.

elektronkoncentrationen  $n_{p0}$  i termisk jämvikt bestämd av massverkans lag

$$n_{p0} = \frac{n_i^2}{p_{p0}} = \frac{n_i^2}{N_A}.$$

Elektronkoncentrationen i ledningsbandet ges som vi fann i tidigare kapitel av

$$n(x) = N_C e^{-(E_C(x) - E_F)/kT}$$

Skillnaden mot tidigare fall, där vi betraktade en homogen halvledare, är att ledningsbandskanten i pn-övergången är lägesberoende,  $E_C(x)$ , och därmed blir även elektronkoncentrationen lägesberoende; n(x).

Betraktar vi först övergången i termisk jämvikt så har vi i punkten  $x = -x_p$ 

$$n_{p0} = n(-x_p) = N_C e^{-(E_C(-x_p) - E_F)/kT}$$

och i punkten  $x_n$ 

$$n_{n0} = n(x_n) = N_C e^{-(E_C(x_n) - E_F)/kT}$$

Ledningsbandskanten på p-sidan ligger högre i energi,

$$E_C(-x_p) = E_C(x_n) + e\psi_0,$$

Du kan utnyttja ekv. 7.8 för att härleda ett uttryck för  $\psi_0$ .

och därmed kan vi skriva



Figur 7.18: pn-övergången i

termisk jämvikt.

#### 7.2. PN-ÖVERGÅNGEN MED PÅLAGD SPÄNNING 195

$$n_{p0} = N_C e^{-(E_C(x_n) - E_F + e\psi_0)/kT} = n_{n0} e^{-e\psi_0/kT}.$$
(7.8)

Detta uttryck ger alltså ett samband mellan elektronkoncentrationerna på ömse sidor om utarmningsområdet vid jämvikt.

När barriärhöjden minskar från  $e\psi_0$  till  $e(\psi_0 - U)$  förväntar vi oss, på grund av injektionen, att elektronkoncentrationen i punkten  $-x_p$ ökar kraftigt relativt sitt jämviktsvärde  $n_{p0}$ . Det ligger nu nära till hands att modifiera jämviktssambandet 7.8 genom att helt enkelt ersätta  $\psi_0$  med  $\psi_0 - U$  vilket ger

$$n_p(-x_p) = n_{n0}e^{-e(\psi_0 - U)/kT}$$

eller

$$n_p(-x_p) = n_{n0}e^{-e\psi_0/kT} \cdot e^{eU/kT} = n_{p0}e^{eU/kT}$$
(7.9)

Med  $n_p$  betecknar vi den ökade elektronkoncentrationen på p-sidans gräns mot utarmningsområdet (punkten  $-x_p$ ) när en framspänning U lagts över övergången.

Denna extrapolation är långt ifrån trivial, när vi överger jämvikt förlorar vi mycket av den trygghet som statistiska fysikens jämviktsvillkor erbjuder. Men om vi antar att vi inte befinner oss alltför långt ifrån jämvikt så kan vi hoppas att samband extrapolerade från jämvikt ändå kan ge rimliga resultat.

Elektronkoncentrationen  $n_p(-x_p)$  vid utarmningsområdets gräns till p-området ökar alltså kraftigt vid framspänning enligt 7.9. Vi får en stor nettoladdning som majoritetsladdningsbärarna (hålen) kommer att neutralisera. Hål strömmar in från kontakten för att skapa en stationär laddningsfördelning som återspeglar elektronfördelningen men har motsatt tecken. I punkten  $-x_p$  gällde vid jämvikt (om vi tar dopnivåerna i fig. 7.6 som exempel)

$$n_{p0}(-x_p) \sim 10^8 \text{ m}^{-3}$$

Injektionen ökar värdet med en faktor  $e^{eU/kT} \sim 10^8$  (enligt ekv. 7.9) när U = 0.5 V till

$$n_p(-x_p) \sim 10^{16} \text{ m}^{-3}$$

Denna ändring är fantastiskt stor, men hålkoncentrationen  $p_{p0} = N_A = 10^{24} \text{ m}^{-3}$  är fortfarande 8 tio-potenser större! Detta betyder att ökningen i hålkoncentration som behövs för att neutralisera den ökade mängden elektroner är försumbar relativt sett,  $p_p$  förblir nästan oförändrad.

När

$$n_p(-x_p) - n_{p0} \ll p_{p0} = N_A$$



Figur 7.19: Koncentrationerna av injicerade laddningsbärare som funktion av avståndet från utarmningsområdets gränser. Långt borta från utarmningsområdet (i punkt A och B) har alla injicerade laddningsbärare rekombinerat och endast jämviktskoncentrationen av minoritetsladdningsbärare återstår.

gäller talar vi om låg-nivå injektion. Detta antagande utnyttjade vi implicit ovan när vi i sambandet

$$n_p(-x_p) = n_n e^{-e(\psi_0 - U)/kT}$$

ersatte  $n_n \mod n_{n0}$ .

För de injicerade hålen vid gränsen till n-sidan  $(x_n)$  gäller – vid låg-nivå injektion – analogt

$$p_n(x_n) = p_{n0}e^{eU/kT}.$$

Dessa injicerade minoritetsladdningsbärare befinner sig i ett fältfritt (vi har ju antagit att spänningsfallet ligger enbart över utarmningsområdet), neutralt område. De kan bara fortsätta på sin väg genom diffusion. Samtidigt är de omgivna av majoritetsladdningsbärare, som de efter hand kommer att rekombinera med. I stort sett inga injicerade elektroner överlever längre än till säg punkten A, inga hål längre än till säg punkten B i figur 7.19. Dessa punkter är något godtyckligt valda och införs endast för att grovt beskriva området inom vilket minoritetsladdningsbärarna rekombinerar. De injicerade laddningsbärarnas koncentration avtar därmed när vi avlägsnar oss från utarmningsområdet och en skiss av koncentrationerna bör då se ut som i figur 7.19.

Man kan visa (se appendix i slutet av kompendiet) att kurvorna i figur 7.19 beskrivs av exponentialfunktioner

$$n_p(x) - n_{p0} = (n_p(-x_p) - n_{p0}) \cdot e^{(x+x_p)/L_e}$$
(7.10)

 $\operatorname{och}$ 

$$p_n(x) - p_{n0} = (p_n(x_n) - p_{n0}) \cdot e^{-(x - x_n)/L_h}$$
(7.11)

#### Ideala diodekvationen

Antagande 3: Vi antar lågnivåinjektion, dvs att ändringar av minoritetsladdningsbärarkoncentrationer är mycket små i förhållande till majoritetsladdningsbärarkoncentrationen. där

$$L_h = \sqrt{D_h \tau_h}$$
 och  $L_e = \sqrt{D_e \tau_e}$ 

är de så kallade diffusionslängderna för hål resp. elektroner. Överskottet av minoritetsladdningsbärare har minskat med en faktor 1/epå en diffusionslängd.  $D_h$  och  $D_e$  är diffusionskonstanterna och  $\tau_h$ och  $\tau_e$  livstiderna för hål resp. elektroner som kort introducerats i kapitel 6.11. Livstiderna ger medellivslängden för minoritetsladdningsbärare, t.ex.  $\tau_h$  för hål på n-sidan, dvs hur länge ett visst hål i medel existerar innan det rekombinerar med en majoritetsladdningsbärare. Hur långt det injicerade överskottet av minoritetsladdningsbärare når innan det har rekombinerat beror på diffusionslängden som i sin tur beror av både diffusionskonstanten och livstiden.

I termisk jämvikt sker hela tiden både rekombination och generation. Vid jämvikt är processerna lika starka och ger konstanta laddningsbärarkoncentrationer. Om vi injicerar extra minoritetsladdningsbärare, får vi lokalt i halvledaren en nettorekombination när systemet strävar mot jämvikt.

**Exempel:** Vi betraktar en pn-övergång i kisel och jämför diffusionslängden  $L_e$  med utarmningsområdets bredd w. Vi tar som indata:

$$\begin{split} &N_A = N_D = 10^{21} \text{ m}^{-3} \text{ (typiska värden)} \\ &\varepsilon_r(\text{Si}) = 11.4 \\ &T = 300 \text{ K} \\ &m_e^* = m_h^* = m \text{ (approximerar lite här)} \\ &E_g = 1.1 \text{ eV} \\ &D_e \sim 20 \cdot 10^4 \text{ m}^2/\text{s samt } \tau_e \sim 5 \cdot 10^{-7} \text{ s} \quad \text{(typiska värden)} \end{split}$$

$$&n_i^2 = N_C N_V e^{-E_g/kT} = 2.07 \cdot 10^{32} \text{ m}^{-6} \\ &\psi_0 = \frac{kT}{e} \ln \left(\frac{n_{n0} p_{p0}}{n_i^2}\right) = 0.58 \text{ V} \quad (e\psi_0 \leq E_g) \\ &w = \sqrt{\frac{2\varepsilon\psi_0}{e} \frac{1}{(N_D + N_A)}} \left(\sqrt{\frac{N_A}{N_D}} + \sqrt{\frac{N_D}{N_A}}\right) = 1.59 \cdot 10^{-6} \text{ m} \\ &\mathcal{E}_m = 2\frac{\psi_0}{w} = 7.30 \cdot 10^5 \text{ V/cm} \\ &\Rightarrow \text{ Diffusionslängd } L_e = \sqrt{D_e \tau_e} = 3.2 \cdot 10^{-5} \text{ m} \end{split}$$

Diffusionslängden kan vara mycket större än utarmningsområdets vidd w. Observera att livstiderna inte är samma sak som spridnings- eller kollisionstiderna vi använder för att beräkna mobilitet.



Figur 7.20: Alla injicerade minoritetsladdningsbärare rekombinerar mellan A och  $-x_p$  (elektronerna) eller mellan  $x_n$  och B (hålen) – gråa pilar.

Från ekvation 7.10 (illustrerad i fig 7.19) ser vi att vi har en ickekonstant fördelning av minoritetselektroner på p-sidan av pn-övergången. Enligt vad vi har diskuterat i kapitel 2 om diffusion, inser vi att koncentrationsgradienten ger upphov till en diffusionsström. I punkten  $-x_p$  blir elektronströmmen

$$J_e(-x_p) = eD_e \frac{dn_p}{dx} = \frac{eD_e(n_p(-x_p) - n_{p0})}{L_e}$$
$$J_e(-x_p) = \frac{eD_e n_{p0}}{L_e} \left(e^{eU/kT} - 1\right)$$

För hålströmmen finner vi analogt (med hjälp av ekv. 7.11)

$$J_h(x_n) = -eD_h \frac{dp_n}{dx} = \frac{eD_h(p_n(x_n) - p_{n0})}{L_h}$$
$$J_h(x_n) = \frac{eD_h p_{n0}}{L_h} \left(e^{eU/kT} - 1\right)$$

Någonstans mellan A och  $-x_p$  rekombinerar i princip alla elektroner som kommer från n-området, någonstans mellan  $x_n$  och B rekombinerar alla hål, se figur 7.20. Eftersom

$$\frac{D}{L} = \sqrt{\frac{D}{\tau}}$$

ser vi att strömmarna ökar för större diffusionskonstanter och minskar när livstiden blir längre.

Ovan har vi betraktat enbart diffusionsströmmarna hos minoritetsladdningsbärarna. Driftströmmen bortser vi från då den för minoritetsladdningsbärare är mycket liten eftersom koncentrationen av laddningsbärare såväl som fältstyrka är små.



Figur 7.21: Majoritetsladdningsbärare möter de injicerade minoritetsladdningsbärarna för att rekombinera.

I ett stationärt tillstånd ersätts minoritetsladdningsbärarna som rekombinerar hela tiden av nya laddningsbärare som diffunderar från utarmningsområdet, dvs levereras av strömmen  $J_e(-x_p)$  eller  $J_h(x_n)$ . Därmed upprätthålls en stationär koncentration. Lika lite kan majoritetsladdningsbärarna bara försvinna, även de måste kontinuerligt levereras. Vi måste öka deras inflöde från kontakterna som visas i figur 7.21, dvs just med samma strömtäthet som  $J_e(-x_p)$  för hålen och med  $J_h(x_n)$  för elektronerna.

Nu är bilden över laddningsbärarnas flöden i pn-övergången äntligen klar. Vi har i punkten  $x = x_n$  en viss diffusionsström av hål – detta är den enda hålström vi har i denna punkt. Vi vet att hålströmmen är konstant genom utarmningsområdet – den har alltså samma värde i punkten  $x = -x_p$  som i punkten  $x = x_n$ . Mekanismen som ger hålströmmen i punkten  $-x_p$  är svårare att säga något om – hålen är här majoritetsladdningsbärare och vi kan anta att drift såväl som diffusion bidrar. Även ett svagt elektriskt fält kan ge en märkbar driftström hos de många majoritetsladdningsbärarna. Oavsett mekanism vet vi att hålströmmen i punkten  $-x_p$  ges av  $J_h(x_n)$  samt att elektronströmmen (från diffusion) i samma punkt ges av  $J_e(-x_p)$  – lägger vi ihop dem får vi hela strömmen genom pn-övergången. Detta ger oss då med

$$J = J_e(-x_p) + J_h(x_n)$$

den ideala diodekvationen

$$J = \left(\frac{eD_e n_{p0}}{L_e} + \frac{eD_h p_{n0}}{L_h}\right) \left(e^{eU/kT} - 1\right) = J_0 \left(e^{eU/kT} - 1\right)$$

I  $J_0$  har vi samlat en hel mängd parametrar. Vi ser att  $J_0$  är spänningsoberoende, men också att alla termer i  $J_0$  (utom e) är tempe-



Figur 7.22: Totala strömmen får vi genom att addera  $J_e(-x_p)$  och  $J_h(x_n)$ .



Figur 7.23: Ideala diodekvationen  $I = I_0(e^{eU/kT} - 1)$ 

199

raturberoende. Starkast – exponentiellt – är temperaturberoendet i minoritetsladdningsbärarkoncentrationerna.

#### 7.2.5 Backspänning

När vi byter polaritet på den pålagda spänningen, dvs lägger på spänningen U < 0, hamnar hela spänningsfallet återigen över bristområdet. Detta innebär att det elektriska fältet ökar, potentialbarriären ökar och även bristområdets vidd ökar. Situationen beskrivs av figur 7.24 och 7.25. Diodekvationen blir

$$J = J_0 \left( e^{eU/kT} - 1 \right) = J_0 \left( e^{-e|U|/kT} - 1 \right)$$

(då U < 0) vilket reducer<br/>as till

 $J \approx -J_0$ 

för backspänningar sådana att |U| > kT/e.  $J = -J_0$  kallas för läckström eller backström.

Vilka laddningsbärare utgör denna lilla läckström? Antalet elektroner som har tillräcklig energi för att ta sig över potentialbarriären har sjunkit drastiskt, elektronflödet mot p-sidan kommer att bli försumbart litet. Den enda kvarstående möjligheten för laddningstransport genom övergången är att minoritetsladdningsbärare råkar diffundera till bristområdets kant där det starka elektriska fältet sveper dem över till motsatta sidan (fig. 7.25). Nu kan dessa minoritetsladdningsbärare inte ersättas av nya laddningsbärare som kommer från kontakterna, utan måste skapas genom termisk generation i områdena mellan A och  $-x_p$  (elektroner) och mellan  $x_n$  och B (hål). Vid framspänning rekombinerade laddningsbärare i dessa områden, nu genereras de i samma områden, se fig. 7.26.



 $\psi_0 - U$ 

Figur 7.24: pn-övergången vid backspänning.



Figur 7.25: Backspänningen hindrar att elektroner rör sig från n- till psidan och att hål tar sig från p- till n-sidan. Därmed återstår endast den svaga ström som uppstår då minoritetsladdningsbärare rullar ner för backen till n-sidan (elektroner) eller bubblar upp till p-sidan (hål).



Figur 7.26: Minoritetsladdningsbärare genereras i angränsande områden kring utarmningsområdet, diffunderar till utarmningsområdets kanter och sveps genom utarmningsområdet av det elektriska fältet. De termiskt genererade majoritetsladdningsbärarna flyter ut till kontakterna.

# 7.3 Spärrskiktskapacitans (backspänning)

I termisk jämvikt gäller för vidden av utarmningsområdet på n<br/>sidan

$$x_n = \sqrt{\frac{2\varepsilon_s\psi_0}{e}\frac{1}{N_D + N_A}}\sqrt{\frac{N_A}{N_D}}$$

När en backspänning U < 0 läggs över dioden ökar  $x_n$  och därmed ändras den positiva laddningen i utarmningsområdet som de joniserade donatorerna ger upphov till.  $x_n$  blir

$$x_n = \sqrt{\frac{2\varepsilon_s(\psi_0 - U)}{e} \frac{1}{N_D + N_A}} \sqrt{\frac{N_A}{N_D}}$$

och den positiva laddningen per ytenhet, Q, i n-området av övergången ges av

$$Q = eN_D x_n = \sqrt{2\varepsilon_s e(\psi_0 - U) \frac{N_D N_A}{N_D + N_A}}$$

En småsignal i form av en växelspänning som är överlagrad den konstanta backspänningen U ser en småsignalskapacitans, eller spärrskiktskapacitans

$$C = \frac{dQ}{dU} = \sqrt{\frac{\varepsilon_s e}{2(\psi_0 - U)} \frac{N_D N_A}{(N_D + N_A)}}.$$

Utarmningsområdets vidd w gavs av

$$w = \sqrt{\frac{2\varepsilon_s(\psi_0 - U)}{e} \left(\frac{N_D + N_A}{N_D N_A}\right)}$$

och vi ser att kapacitansen då kan tecknas

$$C = \frac{\varepsilon_s}{w}$$

Vi framspänning tillkommer ytterligare kapacitans relaterad till de injicerade minoritetsladdningsbärarna. Den totala kapacitansen kan då avvika från vad vi har härlett ovan. Vid backspänning, och ibland även för svaga framspänningar (hur svaga beror på den specifika pn-övergångens parametrar), kan man ofta approximera den totala kapacitansen med spärrskiktskapacitansen.



Figur 7.27: Laddning på n-sidan som funktion av spänning.



Figur 7.28: Vänster: I(U) vid låga spänningar. Höger: I(U) vid höga spänningar.

#### 7.4 Ström-spänningskarakteristik

Ser vi nu tillbaka på denna långa diskussion så noterar vi att strömmarna bestämdes av injektionen vid utarmningsområdets gränser,

$$n_p(-x_p) = n_{p0}e^{eU/kT}$$

 $\operatorname{och}$ 

$$p_n(x_n) = p_{n0}e^{eU/kT}$$

och den därpå följande diffusionen in i de neutrala områdena. Dessa är då de faktorer som bestämmer och begränsar diodens ström vid en given framspänning U. Denna ideala diodström kallas därför även diffusionsström.

Ofta tänker vi på diodens karakteristik som i den vänstra delen av figur 7.28. Inom elektronik där man har diodströmmar kring t ex 1 mA betraktar man en annan skala som i den högra figuren. Man förenklar sedan bilden och säger att dioden leder när spänningen överstiger tröskelspänningen  $U_T \sim 0.7$  eV (kiseldiod). I germanium är  $U_T \sim 0.3$  V.

För att komma fram till den ideala diodekvationen gjorde vi ett antal antaganden; t. ex lågnivå-injektion, som i en verklig komponent under olika förhållanden kan vara mer eller mindre uppfyllda, vilket medför att diodekvationen beskriver ström-spänningskarakteristiken för en verklig pn-övergång mer eller mindre bra.

I nästa avsnitt beskriver vi kortfattat vad som händer då generation eller rekombination förekommer i utarmningsområdet – något som vi helt bortsåg från i härledningen av den idela diodekvationen.

Innan dess, ska vi kort fundera över vårt antagande att hela spänningsfallet hamnar över utarmningsområdet. Vi motiverade detta med att utarmningsområdet har få fria laddningsbärare, och att Kan du komma på en förklaring varför tröskelspänningen är mindre i germanium än i kisel? Kom ihåg att bandgapen är olika!



Figur 7.29: Diodström utan (heldragen linje) och med (streckad linje) hänsyn tagen till serieresistans. Notera att strömmen är logaritmerad!

det därmed är den mest resistiva delen av pn-övergången (mer resistiv än de neutrala n- och p-områdena). Allt eftersom vi lägger på en större framspänning, kan vi se i fig 7.28 att resistansen minskar (I/U blir mindre och mindre). Vid höga strömmar kan vi nå en situation där resistansen associerad med utarmningsområdet är mycket liten och andra resistanser associerade med pn-övergången dominerar. Strömspänningskurvan blir då linjär och karakteristisk för pn-övergångens totala resistans. Detta kallas att diodströmmen domineras/karakteriseras av serieresistans. De resistiva element som typiskt ingår i serieresistans är resistansen hos de neutrala n- och p-områdena samt kontaktresistanser.

Vårt antagande om att allt spänningsfall hamnar över utarmningsområdet är alltså sämre ju större strömmen är. Vi får då tänka oss att bara en del av den pålagda spänningen hamnar över utarmningsområdet och därmed blir strömmen lägre än vad diodekvationen ger vid den aktuella spänningen (se figur 7.29).

Vid höga framspänningar faller också vårt antagande om lågnivåinjektion, vilket även det kommer att leda till en minskning av strömmen relativt den ideala diodekvationen.

205

Sammanfattning: Framspänd pn-övergång, U > 0

Diffusionslängderna  $L_e$  och  $L_h$  är oftast mycket längre än w.

Injektion av minoritetsladdningsbärare. Minoritetsladdningsbärarna rör sig endast genom diffusion i de neutrala områdena där de rekombinerar med majoritetsladdningsbärarna.

*Diffusionslängden* avgör hur långt minoritetsladdningsbärarna diffunderar innan de försvinner genom rekombination. Diffusionslängden beror på både diffusionskonstanten och livstiden.

a - Hålström som injiceras i n-sidan därhålen rekombinerar med elektroner.b - Hålström som möter injicerade elektroner på p-sidan för rekombination.

c - Elektronflöde som injiceras i p-sidan där elektronerna rekombinerar med hål.
d - Elektronflöde som möter injicerade hål på n-sidan för rekombination.





# Sammanfattning: Backspänd pn-övergång, U < 0

Diffusionslängderna  $L_e$  och  $L_h$  är oftast mycket längre än w.

Extraktion av minoritetsladdningsbärare. Minoritetsladdningsbärare som diffunderar in i utarmningsområdet sveps över till andra sidan - extraktion. Motsvarande mängd minoritetsladdningsbärare genereras i de neutrala områdena vid steady state.

Diffusionslängden avgör hur långt ifrån utarmningsområdet denna extraktion med påföljande generation sker. Diffusionslängden beror på både diffusionskonstanten och livstiden.

a - Hålström som genereras på n-sidan, passerar utarmningsområdet och fortsätter till kontakten på p-sidan.

b - Hålström som genereras på p-sidan och fortsätter till kontakten på p-sidan.

c - Elektronflöde som genereras på p-sidan, passerar utarmningsområdet och fortsätter till kontakten på n-sidan.

d - Elektronflöde som genereras på n-sidan och fortsätter till kontakten på n-sidan.

De totala hål- och elektronströmmarna i övergången vid backspänning.





Figur 7.31: Rekombination i utarmningsområdet ger ett bidrag till strömmen genom dioden, på grund av de svarta flödena.

# 7.5 Generation och rekombination i utarmningsområdet

Vi bortsåg ovan från rekombination och generation i utarmningsområdet och vi skall nu till sist även betrakta denna möjlighet. Dessa processer kommer att ge upphov till ytterligare strömmar, förutom den ideala diffusionsström vi redan diskuterat.

#### 7.5.1 Framspänning

Vid framspänning ökar koncentrationerna av elektroner och hål i bristområdet, dvs de blir större än i termisk jämvikt. Då blir produkten

$$np > n_i^2$$
.

Systemet strävar mot jämvikt och försöker reducera koncentrationerna i utarmningsområdet. Vi förväntar oss då en netto-rekombination i utarmningsområdet. Vi ser i figur 7.31 att denna rekombination kräver ett ytterligare inflöde av elektroner och hål (svarta pilar) från kontakterna vilket åstadkoms genom en liten ökning av spänningsfallet över de neutrala områdena.

Detta bidrag till diodströmmen som vi kallar rekombinationsström brukar vara speciellt märkbart vid låga framspänningar när injektionen av minoritetsladdningsbärare på p- och n-sidan är liten. För att kunna ta sig över potentialbarriären behöver en laddningsbärare ju energin  $e(\psi_0 - U)$ , för att två laddningsbärare ska kunna mötas i utarmningsområdet och rekombinera räcker det att de tillsammans har denna energi.



Figur 7.30: Rekombination i utarmningsområdet.



Figur 7.32: Strömmen I som funktion av framspänningen U.  $\eta$  är idealitetsfaktorn.

Rekombinationsströmmens spänningsberoende kan visas ges av uttrycket

$$I_{rec} \sim e^{eU/2kT}$$

och därmed får vi den totala strömmen genom att addera rekombinationsströmmen till den ideala diodströmmen. Man brukar skriva denna generaliserade diodekvation på formen

$$I = I_0 \left( e^{eU/\eta kT} - 1 \right), \ (1 \le \eta \le 2)$$

där  $\eta$  är den så kallade *idealitetsfaktorn*. Figur 7.32 visar att strömspännings-karakteristiken vid låga framspänningar domineras av rekombinationsströmmen ( $\eta = 2$ ) och vid högre framspänningar av den ideala diffusionsströmmen ( $\eta = 1$ ).

I ett material med indirekt bandgap som kisel sker rekombinationsoch generationsprocesser normalt via djupa störnivåer i bandgapet.

#### 7.5.2 Backspänning

Vid backspänning (U < 0) minskar koncentrationen fria laddningsbärare i utarmningsområdet,

 $np < n_i^2$ .

Därmed uppstår en netto-generation av laddningsbärare i utarmningsområdet på grund av systemets inneboende strävan mot jämvikt. De i utarmningsområdet genererade laddningsbärarna sveps bort till de neutrala områdena av det starka fältet.

Denna generationsström kommer att vara temperaturberoende eftersom den beror på termiska generations-processer. Den kommer också att leda till att läckströmmen, dvs  $I = -I_0$  som vi har vid backspänning, uppvisar ett tydligt spänningsberoende som i figur 7.36. Sannolikheten för termisk generation är, förutom att vara temperaturberoende, även proportionell mot storleken av det område där generationen kan ske. Utarmningsområdet ökar i vidd enligt

$$w = \sqrt{\frac{2\varepsilon(\psi_0 - U)}{e} \frac{1}{N_D + N_A}} \left( \sqrt{\frac{N_A}{N_D}} + \sqrt{\frac{N_D}{N_A}} \right)$$

För större backspänningar  $(|U| \gg \psi_0)$  gäller då

$$w \sim \sqrt{|U|}$$

 $\operatorname{och}$ 

$$\ln I \sim \ln w \sim rac{1}{2} \ln |U|$$
 jfr. figur 7.36



Figur 7.33: Generationsprocess i utarmningsområdet.



Figur 7.35: Upp-och-ner-vänd IVkurva - jämför med data i figur 7.36. Vid genombrott ökar strömmen kraftigt.



Figur 7.36: Backströmmen ökar med ökande backspänning särskilt vid låga temperaturer på grund av generation.



Figur 7.34: Generation i utarmningsområdet ger ett bidrag till diodströmmen, generationsströmmen (svarta pilar).



Figur 7.37: a) En pn-övergång fungerar som likriktare, b) en småsignal-växelspänning uöverlagras en konstant förspänning  $U_1$  vilket resulterar i en växelström i bestämd av småsignal-resistansen r.

### 7.6 Ekvivalent småsignalschema

När en växelspänning läggs över pn-övergången fungerar övergången som likriktare enl. figur 7.37 a). Strömmen genom dioden blir en pulserande likström.

När en småsignal-växelspänning u överlagras en konstant förspänning  $U_1$ , ser växelspänningen dioden som en resistans r, vars storlek bestäms av spänningen  $U_1$ ,

$$\frac{1}{r} = \left. \frac{dI}{dU} \right|_U$$

och den resulterande växelströmmen bestäms av

 $u = r \cdot i.$ 

Småsignal-växelspänningens amplitud förutsätts vara liten i förhållande till  $U_1$ . I ett småsignalschema är signalen det centrala, konstanta förspänningar exempelvis träder i bakgrunden.

# 7.7 Genombrott vid backspänning

Tunnel- eller Zener-genombrott Tunnelgenombrott uppstår vanligtvis i mycket kraftigt dopade dioder där det elektriska fältet i utarmningsområdet är högt, storleksordningen  $10^8$  V/m. Vi hade ovan

$$\mathscr{E}_m = \frac{eN_D}{\varepsilon_s} x_n \quad \text{och}$$



Figur 7.40: Vid genombrottsspänningen ökar backströmmen dramatiskt.

$$x_n = \sqrt{\frac{2\varepsilon_s\psi_0}{e}\frac{1}{N_D + N_A}}\sqrt{\frac{N_A}{N_D}}$$

Därmed följer

$$\mathscr{E}_m \sim \sqrt{\frac{N_D N_A}{N_D + N_A}} = \sqrt{\frac{N_D}{2}}$$

om exempelvis dopkoncentrationerna är lika,  $N_A = N_D$ . Banden är då mycket branta, eftersom  $x_n$  är litet, och avståndet mellan valensoch ledningsbandet – "horisontellt" vid samma energi – blir litet. Då kan en elektron i p-sidans valensband tunnla genom bandgapet över till n-sidan (se figur 7.38). Vi kommer ihåg att möjligheten för tunnling beror kritiskt på bredden av det förbjudna området partiklarna måste passera.

Lavingenombrott När laddningsbärare som accelereras i utarmningsområdet får tillräckligt hög energi kan de i en kollision med en elektron i valensbandet excitera denna till ledningsbandet och därmed skapa elektron-hål par som i sin tur kan excitera ytterligare par. Detta kan då leda till en lavin av laddningsbärare i utarmningsområdet. Strömmen rusar iväg. Genombrott modifierar den ideala diodkarakteristiken enligt figur 7.40.



Figur 7.38: Är det horisontella avståndet mellan lednings- och valensband tillräckligt litet, så kan elektroner tunnla genom övergången och backströmmen ökar.



Figur 7.39: På sin väg genom utarmningsområdet accelereras elektroner och hål och kan i kollisionsprocesser skapa nya elektron-hål par, backströmmen ökar.

#### Exempel:

I en pn-övergång i kisel är dopkoncentrationerna

 $N_A = 10^{23} \text{ m}^{-3} \text{ och } N_D = 10^{22} \text{ m}^{-3},$ 

den intrinsiska laddningbärarkoncentrationen

 $n_i = 1.5 \cdot 10^{16} \text{ m}^{-3} \text{ vid } T = 300 \text{ K}$ 

och den relativa dielektricitetskonstanten (permittivitet<br/>stalet)  $\varepsilon_r=11,7.$  Diffusionskonstanterna är

$$D_e = 25 \cdot 10^{-4} \text{ m}^2/\text{s och } D_h = 10 \cdot 10^{-4} \text{ m}^2/\text{s}$$

och laddningsbärarnas livstider  $\tau_e = \tau_h = 5 \cdot 10^{-7}$  s.

a) Bestäm läckströmtätheten  $J_0$  vid T = 300 K.

$$J_0 = \frac{eD_e n_{p0}}{L_e} + \frac{eD_h p_{n0}}{L_h}$$

Med  $L_e=\sqrt{D_e\tau_e}$  och  $L_h=\sqrt{D_h\tau_h}$  och  $n\cdot p=n_i^2$ kan strömtätheten skrivas

$$J_0 = en_i^2 \left( \sqrt{\frac{D_e}{\tau_e}} \cdot \frac{1}{N_A} + \sqrt{\frac{D_h}{\tau_h}} \cdot \frac{1}{N_D} \right) = 1.87 \cdot 10^{-7} \text{ A/m}^2$$

b) Bestäm pn-övergångens strömtä<br/>thet vid en framspänning  $U=0,5~\mathrm{V}.$ 

$$J = J_0(e^{eU/kT} - 1) = 47 \text{ A/m}^2$$

# Kapitel 8

# Den bipolära transistorn

Den bipolära transistorn består av två pn-övergångar, pnp eller npn, där det mellersta området, basen, är mycket tunt. Den framspända emitter-basövergången injicerar laddningsbärare in i basen som till största delen överlever passagen genom basen och sveps iväg till kollektorn av det starka elektriska fältet i den backspända bas-kollektorövergången. Transistorn kan fungera som förstärkare, och den kan även fungera som switch i logiska kretsar. Är bägge övergångarna framspända, kan en stor ström flyta mellan emitter och kollektor samtidigt som spänningsfallet mellan emitter och kollektor är litet. Switchen är sluten. Är bägge övergångarna backspända är switchen öppen.

# 8.1 Vi bygger en strömkälla

Strömmen genom en backspänd ideal diod, diodens läckström, är konstant över ett brett spänningsintervall, dvs oberoende av den pålagda backspänningen - se figur 8.1

Dioden fungerar som en strömkälla. Detta innebär att den – inom intervallet – genererar en konstant om än mycket liten ström som är oberoende av den spänning som ligger över dioden. Dioden som strömkälla är en aktiv komponent. (I en resistans däremot bestäms strömmen av den pålagda spänningen, I = U/R, och resistorn själv är passiv.)

Med tanke på det som följer räknar vi nu strömmen i backriktning,  $I_C$ , som positiv och räknar även spänningen över dioden i backriktningen,  $U_{BC}$ , som positiv. Beteckningarna kan verka omotiverade men de är anpassade till den följande diskussionen.

En *ideal strömkälla* ger en konstant ström för alla spänningar (streckad linje i figur 8.2) medan dioden (heldragen linje och punkter) endast i ett begränsat intervall motsvarar en ideal strömkälla. Detta område begränsas av diodens genombrott vid stora och det fram-



Figur 8.1: pn-övergångens ström-spänningskarakteristik och kretssymbol.



Figur 8.4: Den framspända emitterbas dioden levererar hål som minoritetsladdningsbärare till basen. Hålen diffunderar genom basen och samlas upp av den backspända bas-kollektor dioden.



Figur 8.2: Jämförelse mellan karakteristiken hos en ideal strömkälla och hos en pn-övergång.



Figur 8.3: pn-övergången med backspänning  $U_{BC}$ 

spända området med  $U_{BC} < 0$ . Men framför allt är strömmen som dioden genererar alltför liten för att vara av intresse i praktiken.

Hur skulle vi kunna öka strömmen i vår strömkälla? Backströmmen beror ju på koncentrationen av minoritetsladdningsbärare i de neutrala områdena nära övergången i både p- och n-området (dvs vid A och B i figur8.3). Tillför vi energi t ex i form av ljus eller värme så kan dessa koncentrationer ökas kraftigt och därmed även strömmen. Denna ökning i strömmen kan t ex utnyttjas för att bestämma intensiteten på det infallande ljuset. Siktar vi mot tillämpningar inom elektroniken måste vi dock hitta ett sätt att åstadkomma samma resultat på elektrisk väg.

Nära utarmningsområdet vill vi alltså öka koncentrationen minoritetsladdningsbärare som kan diffundera till utarmningsområdet och bidra till strömmen. Vi kommer ihåg att det i en pn-övergång just sker en sådan injektion av minoritetsladdningsbärare när vi framspänner övergången. Lägger vi då t ex en framspänd pn-övergång bredvid vår backspända np-övergång som i figur 8.4, så kan vi öka hålkoncentrationen på n-sidan av vår ursprungliga övergång avsevärt. Därmed ökar också strömmen  $I_C$  kraftigt, de injicerade hålen diffunderar genom basen och sveps över till p-området i den backspända övergången av det elektriska fältet. Inte nog med det! Vi ser också att framspänningen i den ena övergången *styr* hur många hål som injiceras och därmed hur stor strömmen blir genom den *andra* övergången.

När vi satt samman denna struktur har vi byggt en transistor. Är



Figur 8.5: Strömmarna i en p<br/>np transistor. De mörkare områdena är utarmningsområdena mellan emitter och bas och mellan bas och kollektor. Neutrala områden är vita. De dominerande strömmarna i strukturen, hålströmmen från emitter till bas,  $I_{Ep}$ , och hålströmmen från bas till kollektor,  $I_{Cp}$ .

spänningarna valda som ovan, dvs  $U_{EB} > 0$  och  $U_{BC} > 0$ , säger vi att transistorn är i sitt *aktiva* område.

Den vänstra pn-övergångens p-område som levererar hålen kallas emitter, den högra pn-övergångens p-område som fångar upp dem kallas kollektor. Det centrala n-området är basen. Emitterströmmen  $I_E$  går in i emittern och kollektorströmmen  $I_C$  lämnar kollektorn. Hålens bidrag till dessa strömmar betecknas med ytterligare ett index p,  $I_{Ep}$  och  $I_{Cp}$ , och elektronbidrag med index n. Basströmmen  $I_B$  flyter ut ur basen, i detta första steg är den noll, och vad den beror på diskuterar vi snart.

Genom att variera emitter-bas övergångens framspänning  $U_{EB}$  kan vi nu styra  $I_{Ep}$  (hålflödet från emitter till bas) och därmed också hålflödet från bas till kollektor, dvs strömmen  $I_{Cp}$  som ger det dominerande bidraget till  $I_C$ . Vi kan därmed öka strömmen som vår strömkälla levererar kraftigt så som figur 8.6 visar.

### 8.2 Hålens väg genom basen

När hålen injicerats in i basen genom emitter-bas övergångens utarmningsområde befinner de sig i basens neutrala område. Här neutraliseras de av majoritetsladdningsbärarna, elektronerna. Området är därmed fältfritt, E = 0, och hålen rör sig enbart genom diffusion. Om nu injektionen från emittern till basen faktiskt skall öka kollektorströmmen måste hålen överleva passagen genom basen utan att rekombinera med elektroner. Ett uppenbart sätt att minska risken för rekombination är att göra basen så tunn som möjligt. Då blir tiden hålen spenderar i den neutrala basen kort. Hur bra vi lyckas anger **bastransportfaktorn**  $\alpha_T$ , dvs förhållandet mellan







Figur 8.8: Transistorn i sitt aktiva område. Elektroner injiceras från basen till emittern vilket ger en oönskad  $I_{En}$ .

hålbidraget till kollektorströmmen,  $I_{Cp}$ , och hålbidraget till emitterströmmen,  $I_{Ep}$ ,

$$\alpha_T = \frac{I_{Cp}}{I_{Ep}}.$$

Parametern  $\alpha_T$  talar alltså om för oss hur många hål som överlevt resan genom basen i förhållande till antalet som skickats iväg av emittern. Om alla hål överlever blir  $\alpha_T = 1$ . Några hål kommer dock att hinna hitta en elektron att rekombinera med i basen och om vi förutsätter att vi har ett stationärt tillstånd med tidsoberoende strömmar och laddningsfördelningar, då måste nya elektroner flyta från baskontakten in i basen. Denna ström kallas basrekombinationsström,  $I_{BB}$ , den ger ett bidrag till basströmmen  $I_B$  men är typiskt tusen gånger mindre än den stora hålströmmen  $I_{Ep} \approx I_{Cp}$ .

#### 8.3 Elektroner injiceras från bas till emitter

Den önskvärda strömmen är hålströmmen från emittern till kollektorn, det var ju kollektorströmmen vi ville kunna styra med emitterbas övergångens framspänning. Men när vi framspänner emitterbas-övergången kommer vi inte bara att injicera hål från emittern till basen utan även injicera elektroner från basen till emittern. Denna elektronström har ingen direkt inverkan på kollektorströmmen utan ökar bara emitter- och basströmmen. Den bidrar då inte till transistorns egentliga funktion och bör minimeras. Vi kan inte bli av med elektroninjektionen men vi kan minska dess relativa storlek i förhållande till hålinjektionen genom att välja en  $p^+n$  övergång, där



Figur 8.7: Basrekombinationsströmmen  $I_{BB} = I_{Ep} - I_{Cp}.$ 



Figur 8.9:  $I_{En}$  bidrar till  $I_B$
$p^+$ - emittern är mycket kraftigare dopad än n-typ basen. Emitterns *injektionseffektivitet* (*emitter efficiency*) anger förhållandet mellan emitterströmmens hålbidrag och den totala emitterströmmen

$$\gamma = \frac{I_{Ep}}{I_{Ep} + I_{En}}$$

Injektionseffektiviteten bör ligga så nära ett som möjligt.

### 8.4 Två pn-övergångar – en transistor

Vi har talat om transistorns strömmar och spänningar lite abstrakt utan att placera transistorn i sin naturliga miljö, en elektrisk krets. I figur 8.11 visas transistorn nu i en enkel krets, som vi kallar gemensam-bas koppling. Här är baskontakten gemensam för både den styrande ingångskretsen med emitter-bas övergången och utgångskretsen med bas-kollektor övergången. Batteriet  $U_1$  ser till att emitter-bas övergången är framspänd och batteriet  $U_2$  att baskollektor övergången är backspänd så att transistorn är i sitt aktiva område.

Åven om vi i vår diskussion tänker oss att vi för samman två pnövergångar är det viktigt att inse att kretsen i figur 8.10 med två pn-övergångar inte alls beskriver den slutliga situationen. Väsentligt är att basen är så smal att emitterströmmen (de injicerade hålen) fortsätter direkt in i den backspända kollektorövergången utan att rekombinera med elektroner. I figur 8.10 skulle alla hål ha rekombinerat när strömmen lämnar emitter-bas dioden (i den metalliska ledningen mellan dioderna flyter endast elektroner) och kollektorströmmen  $I_C$  skulle faktiskt försvinna så när som på en liten läckström i den backspända högra dioden.

Vi sammanfattar i stället de två pn-övergångarna i en **transistor** i figur 8.11. Spänningen i ingångskretsen,  $U_{EB}$ , bestämmer  $I_E$  som dyker upp som den nästan lika stora kollektorströmmen  $I_C$  i utgångskretsen, och denna är då oberoende av spänningen mellan bas och kollektor i utgångskretsen,  $U_{BC}$ . Därmed kan vi se transistorn som en **spännings-styrd strömkälla**, utkretsens ström  $I_C$  (till höger i figuren) styrs av en spänning  $U_{EB}$  i en annan del av kretsen, dvs vid ingången till vänster. Produkten av parametrarna  $\alpha_T$  och  $\gamma$  blir

$$\gamma \cdot \alpha_T = \frac{I_{Cp}}{I_E} = \alpha$$

där vi definierat

$$\alpha = \frac{I_{Cp}}{I_E} \approx \frac{I_C}{I_E}$$



Figur 8.10:  $Tv \dot{a}$  pn-övergångar



Figur 8.11: En transistor



Figur 8.12: Utgångskarakterisktik i gemensam-bas koppling  $I_C = \alpha I_E$  $\alpha$ -gemensam-bas strömförstärkning  $(\alpha \leq 1)$ 

Aktivt område:  $U_{EB} > 0$ , EB-övergången framspänd,  $U_{BC} > 0$ , BC-övergången backspänd Vi antar här att  $I_{Cp} \approx I_C$  eftersom elektronernas bidrag till kollektorströmmen,  $I_{En}$ , (en ytterst liten läckström som vi hittills knappast ens nämnt) kan försummas. Förhållandet  $\alpha$  mellan utgångens och ingångens strömmar är då *strömförstärkningen* i gemensam-bas koppling. Både  $\alpha_T$  och  $\gamma$  bör ha värden mycket nära 1 och därmed blir även  $\alpha$  nästan 1. En strömförstärkning som är en aning mindre än ett låter i första ögonblicket inte särskilt imponerande men vi skall strax se att vi med denna koppling kan uppnå en rejäl spännings- eller effektförstärkning.

Transistorns *utgångskarakteristik* visas i figur 8.12. För varje värde på  $I_E$  som bestäms och hålls konstant av ingångskretsen får vi en kurva  $I_C$  som ger som funktion av  $U_{BC}$  i utgångskretsen. I det aktiva området gäller

$$I_C = \alpha \cdot I_E$$

218

här fungerar transistorn som strömkälla.

Om bas-kollektor övergången däremot framspänns,  $U_{BC} < 0$ , samlar kollektorn inte längre effektivt in hålen som kommer från basen utan börjar själv försöka injicera hål in i basen och kollektorströmmen sjunker. Transistorn sägs bottna (den når saturation). $U_{EB}$  har ju ett värde kring tröskelspänningen (~ 0,7 V) hos emitter-bas övergången. När UBC är ca. 0,7 V injicerar kollektorn ungefär lika många hål som emittern till basen och hålströmmen blir mindre. Samtidigt injicerar basen elektroner in i kollektorn och därmed minskar också  $I_C$  och blir till slut negativ för större framspänningar hos bas-kollektor övergången. Emitter och kollektor har bytt plats och transistorn har hamnat i sitt inverterade område, se även figur 8.18.

### 8.5 Tillverkning

Figur 8.13 visar ett schematiskt tvärsnitt genom en pnp transistor. Man utgår här från ett p-dopat skikt (betecknat *epitaxial layer* i figuren) som n-dopas i ett begränsat område för att skapa basen. Därefter skapas emittern genom kraftig p-dopning i basområdet. Med detta tillvägagångssätt blir dopkoncentrationerna ordnade enligt

 $N_E > N_B > N_C,$ 

dvs acceptorkoncentrationen  $N_E$  i emittern är större än donatorkoncentrationen  $N_B$  i basen, som i sin tur är större än acceptorkoncentrationen  $N_C$  i kollektorn.



Figur 8.13: Schematiskt tvärsnitt genom en pnp transistor.

Figur 8.14 visar en förstoring av transistorns centrala område som är den egentliga pnp transistorn. Vi ser i figur 8.13 att kollektorn är hårt dopad under metallkontakten. Detta är ett vanligt sätt att få en kontakt som är Ohmsk, dvs inte likriktande som metall-halvledarövergångar annars oftast är (se kapitlet om metallhalvledarövergången). De två  $n^+$ -områdena i transistorns ytterkanter avser att isolera transistorn från omgivningen genom att bilda backspända övergångar med den p-dopade kollektorn.

### 8.6 Sammanfattning av transistorns strömmar och arbetssätt

Innan vi tittar på transistorns funktion som förstärkare skall vi kort sammanfatta vår bild av pnp transistorn. Vi börjar med att betrakta strömmarna i det **aktiva området** där emitter-bas övergången är framspänd och bas-kollektor övergången backspänd. Hål injiceras från emittern in i basen  $(I_{Ep})$  och fortsätter vidare till kollektorn  $(I_{Cp})$ . Några hål rekombinerar i basen, de elektroner som då försvinner levereras av bas-rekombinationsströmmen

$$I_{BB} = I_{Ep} - I_{Cp}.$$

Elektroner injiceras från basen till emittern  $(I_{En})$  och elektroner i kollektorn "läcker" över till basen  $(I_{Cn})$ . Dessa två strömmar är ej lika stora, fler elektroner lämnar basen än som strömmar till, skillnaden mellan dessa strömmar måste basströmmen ta hand om, men i förhållande till  $I_{Ep}$  och  $I_{Cp}$  är basströmmen fortfarande mycket liten. Vi skriver både emitterströmmen

$$I_E = I_{Ep} + I_{En}$$

och kollektorströmmen

$$I_C = I_{Cp} + I_{Cn}$$



Figur 8.14: En detalj ur figur 8.13 som visar transistorns emitter, bas och kollektor.



Figur 8.15: pnp-transistorns strömmar.



Figur 8.16: pnp transistorns schematiska uppbyggnad och dess kretssymbol till höger

som en summa av ett hål- och ett elektronbidrag. Basströmmen ges (med Kirchhoffs första lag) av

$$I_B = I_E - I_C = I_{Ep} + I_{En} - I_{Cp} - I_{Cn} = I_{En} - I_{Cn} + I_{BB}$$

där

$$I_{Ep} - I_{Cp} = I_{BB}$$

är basens rekombinationsström.

Transistorns schematiska struktur och spänningarna i gemensambas koppling visas i figur 8.16 tillsammans med kretssymbolen.

I kretssymbolen anger den lilla pilen vid emittern att emitterströmmen flyter in i emittern, dvs att det är en pnp transistor.

I termisk jämvikt är Ferminivån konstant och för dopkoncentrationerna gäller som vi såg ovan vanligtvis

$$N_E \gg N_B \gg N_C.$$

Beroende på spänningarna  $U_{EB}$  och  $U_{CB}$  har pnp transistorn fyra arbetssätt, se figur 8.18.

#### Kirchhoffs första lag:

Summan av strömmarna som går in mot en punkt i en krets måste vara lika med summan av strömmarna som går ut från samma punkt.

#### Kirchhoffs andra lag:

Summan av delspänningarna i en sluten krets är noll.



Figur 8.17: p<br/>np transistorn i termisk jämvikt Hur kan du här se att<br/>  $N_E \gg N_B \gg N_C$  ?



Figur 8.18: pnp transistorns fyra arbetssätt

## 8.7 Spänningsförstärkning - gemensam-bas koppling

Betrakta en pnp transistor i gemensam-bas koppling, figur 8.19. Transistorn är i sitt **aktiva område**, dvs emitterbas övergången är framspänd, kolletor-bas övergången backspänd. Emitterströmmen bestäms av inkretsen, dvs genom val av  $U_1$  och  $R_1$  ( $u_{in} = 0$ ). Har inkretsens ström t ex värdet  $I_{E1}$ , så väljs därmed en av kurvorna i figur 8.20 ut som aktuell utgångskarakteristik (den tjockare kurvan).

I det aktiva området är  $I_C = \alpha I_{E1} \approx I_{E1}$ , men vi vet inte om transistorn där den sitter i kretsen verkligen hamnar i det aktiva området eller vilket värde  $U_{BC}$  har, dvs var vi ligger på karakteristikkurvan. För att bestämma  $I_C$  och  $U_{BC}$  måste vi ta hänsyn till den övriga utkretsen, dvs spänningskällan  $U_2$  och resistansen  $R_2$ . Med Kirchhoffs andra lag följer

$$U_2 = U_{BC} + I_C \cdot R_2$$

eller

$$I_C = \frac{U_2 - U_{BC}}{R_2}$$

 $I_C$  som funktion av  $U_{BC}$  är här en rät linje med lutning  $-1/R_2$ .



Figur 8.19: Gemensam-bas koppling

Be last ning slinje

$$I_C = \frac{U_2 - U_{BC}}{R_2}$$



Figur 8.20: Utgångskarakteristiken med belastningslinje

Denna linje kallas belastningslinjen och skärningspunkten med den aktuella (tjockare) utgångskarakteristiken bestämmer *arbetspunk-ten* eller vilopunkten.

Med versaler (stora bokstäver) betecknar vi konstanta förspänningar och strömmar och med gemener (små bokstäver) faktiska, i allmänhet tidsberoende, spänningar och strömmar.

En liten signalspänning  $u_i n$  överlagras nu på den konstanta förspänningen  $U_1$ . Signalspänningen ändrar emitterströmmen med

$$\Delta i_E = \frac{u_{in}}{R_1 + R_{EB}}$$

där  $R_{EB}$  är det effektiva motståndet hos emitter-bas övergången, se figur 8.21. Kollektorströmmen ändras då med

$$\Delta i_C = \alpha \Delta i_E = \alpha \frac{u_{in}}{R_1 + R_{EB}}$$

vilket resulterar i en ändring av spänningsfallet över  $R_2$ ,

$$u_{ut} = R_2 \Delta i_C = \alpha R_2 \frac{u_{in}}{R_1 + R_{EB}}.$$

Här är gemensam-bas strömförstärkningen  $\alpha \approx 1$ . Den effektiva resistansen  $R_{EB}$  är liten, typiskt 10  $\Omega$ .

Antag nu att vi väljer  $R_1 = 100 \ \Omega$  och  $R_2 = 10^4 \ \Omega$ , då får vi

$$u_{ut} = 91u_{in}$$

Detta innebär en rejäl förstärking av spänningen.

Resultatet ser ut att vara oberoende av batterispänningen  $U_2$ . Transistorn verkar kunna generera höga utspänningar och därmed – i princip obegränsad – effekt ur tomma intet. Detta är förstås omöjligt.



Kom ihåg (figur 8.31):

Figur 8.21: pn-övergångens effektiva resistans som växelspänningen u ser.



Figur 8.22: Utgångskarakteristik med belastningslinjer, fallet  $U_2 = 0$  streckat.

Vi tänker oss att vi tar bort spänningskällan  $U_2$ . Kirchhoffs andra lag ger

$$i_C R_2 = U_{CB} = -U_{BC}$$

och om  $i_C$  fortfarande flyter blir kollektor-bas övergången framspänd. Då har transistorn lämnat det aktiva området och bottnat (nått saturation). Det innebär att kollektorn inte längre sköter sitt jobb att samla upp hålen som emittern injicerat in i (och genom) basen utan själv försöker injicera h ål tillbaka till basen. Belastningslinjen förskjuts, nu är ju  $U_2 = 0$ , och ges för detta fall av den streckade linjen. Vi ser att kollektorströmmen  $i_C$  näst intill försvinner och med den den stora utgångsspänningen. Batteriet  $U_2$  är därmed nödvändigt för att hålla transistorn kvar i det aktiva området där den fungerar som strömkälla. Med andra ord behöver vi en likspänning vilken transistorn kan omvandla till signalspänning.

## 8.8 Gemensam-emitter koppling (npn transistor)

I de flesta fallen används transistorn i gemensam-emitter koppling som visas i figur 8.24. Emittern är jordad, ingångskretsens ström är nu den lilla basströmmen  $I_B$  och utgångskretsens ström är  $I_C$ . Därmed förväntar vi oss att med den lilla strömmen  $I_B$  kunna styra den stora strömmen  $I_C$ .

I det följande avsnittet går vi också över till den i praktiken vanligare npn-transistorn. Detta ändrar ingenting väsentligt, vi får ombytta roller mellan elektroner och hål. Vi måste dock komma ihåg att om elektronerna rör sig åt ett håll så går strömmen åt det andra hållet.



Figur 8.23: Kretssymbol - npn-transistor



Figur 8.25: Banddiagram för transistorn i sitt aktiva område



Figur 8.26: Strömförstärkning i gemensam-emitter koppling



Figur 8.24: Gemensam-emitter koppling

Spänningar och strömmar byter tecken. I kretssymbolen i figur 8.23 har den lilla pilen vid emittern bytt riktning, elektroner går från emittern till basen men strömmen går i motsatt riktning.

I analoga kretsar används transistorn som strömförstärkare. Transistorns aktiva området definierades av

$$U_{BE} > 0$$
 (emitter-bas framspänd) och  $U_{CB} = U_{CE} - U_{BE} \ge 0$  (bas-kollektor backspänd).

Därmed ges gränsen mellan aktiv och bottnad transistor av den streckade kurvan i figur 8.26 där  $U_{CE} = U_{BE}$ . Utgångskarakteristiken visar nu  $I_C$  som funktion av  $U_{CE}$  för olika värden på basström-

men  $I_B$  som bestäms av inkretsen. Kirchhoffs andra lag ger

$$U_{CC} = I_C R_L + U_{CE}$$

och härur följer belastningslinjen

$$I_C = \frac{U_{CC}}{R_L} - \frac{U_{CE}}{R_L}$$

Låt ingångskretsens signal vara en signalspänning  $u_{in}$  som ger upphov till en växelström  $i_{in}$  som överlagras den konstanta strömmen  $I_B$  vilken bestämde arbetspunkten. På grund av signalströmmen  $i_{in}$  varierar nu  $i_B = I_B + i_n$ , den aktuella arbetspunkten flyttas upp och ner längs belastningslinjen, utgångskretsens ström  $i_C = I_C + I_{ut}$  varierar i takt med  $i_B$ . Men  $i_C$  är en stor ström i förhållande till  $i_B$ , vi förväntar oss en avsevärd strömförstärkning, denna ska vi nu bestämma.

I det aktiva området är

$$i_C = \alpha i_E$$

med  $\alpha$  som gemensam-bas strömförstärkning och ur

$$i_B = i_E - i_C$$

följer

$$i_B = i_E - \alpha i_E = (1 - \alpha)i_E$$
 eller  $i_E = \frac{i_B}{1 - \alpha}$ 

Därmed gäller för kollektorströmmen

$$i_C = \frac{\alpha}{1 - \alpha} i_B = h_{fe} i_B$$

med **gemensam-emitter strömförstärkningen** som är av storleksordningen 100 om  $\alpha$  nästan är 1. Detta överensstämmer med vår tidigare diskussion som gav att basströmmen är mycket mindre än emitter- och kollektorströmmarna.

I digitala kretsar kan transistorn användas som en switch, figur 8.27. Allt vi behöver göra för att få en switch är att låta basströmmen variera kraftigare än i förstärkarfallet ovan. Med en stor basström  $I_B$ kan transistorn lämna det aktiva området och hamna i läge ON där spänningsfallet över transistorn,  $U_{CE}$ , är litet och strömmen  $I_C$  stor. Är basströmmen liten (eller negativ) hamnar transistorn i läge OFF där strömmen  $I_C$  är liten och spänningsfallet stort,  $U_{CE} \approx U_{CC}$ , switchen är öppen.

En sådan switch kan vi använda i en logisk grind där en insignal bearbetas och resulterar i en viss utsignal. Som ett exempel diskuterar vi inverteraren i följande avsnitt.







Figur 8.27: Utgångskarakteristiken med belastningslinje. Banddiagram för ON och OFF lägen.



Figur 8.28: Inverteraren

### 8.9 Inverteraren

En av de fundamentala byggstenarna i digitala logiska kretsar är inverteraren. Den har en ingång och en utgång, är inspänningen låg blir utspänningen hög och tvärtom. Vi studerar en npn transistor i gemensam-emitter koppling, både matningsspänningen  $U_{CC} = +5$  V, och spänningarna  $u_i n$  och  $u_u t$  mäts relativt den jordade emittern i figur 8.28

En låg ingång,  $u_{in} = 0$  V, ger  $U_{BE} = 0$  V. Inga elektroner injiceras in i basen för att sedan fortsätta till kollektorn, dvs ingen ström flyter in i kollektorn, transistorn är strypt, sedd som switch är den öppen, "OFF" (se figur 8.29). Matningsspänningen ligger över transistorn,  $U_{CE} = U_{CC} = 5$  V, och därmed är  $u_{ut} \approx U_{CC} = 5$  V.

Ändrar vi  $u_{in}$  till 5 V blir bas-emitter övergången framspänd. För

Inverterare

U <sub>in</sub>	U <sub>ut</sub>
0 V	5 V
5 V	0 V



Figur 8.29: Transistorn är strypt i OFF läge

inkretsen gäller med Kirchhoffs andra lag

$$u_{in} = i_B R_B + u_{EB}$$

Eftersom emitter-bas övergångens tröskelspänning är ungefär 0,7 V kommer den största delen av spänningsfallet (lite drygt 4 V) att hamna över resistorn. En stor basström flyter som begränsas av resistorn

$$i_B = \frac{u_{in} - u_{EB}}{R_B} \approx \frac{u_{in}}{R_B}$$

Om resistansen  $R_B$  inte är vald alltför stor kommer basströmmen att bli stor och transistorn bottnar. Kollektorströmmen  $I_C$  är stor och spänningsfallet  $u_{CE}$  övertransistorn litet, eftersom  $u_{CE} \leq u_{BE} \sim$ 0,7 V. Transistorn som switch är nu sluten, den är ON. Och eftersom  $u_{CE} = u_{ut}$  blir utgången "låg".

Vi ser förresten i utgångskarakteristiken att vi i detta bottnade läge inte längre kan öka kollektorströmmen genom att öka  $U_{BE}$  eller  $I_B$ . Strömförstärkningen

$$\frac{\Delta i_C}{\Delta i_B}$$

minskar och går mot noll, transistorn "bottnar".

Strömmen genom transistorn leder till betydande effektförbrukning i ON tillståndet. npn transistorns fyra arbetssätt visas i figur 8.33.



Figur 8.30: Utgångsläge OFF,  $U_{in} = 5$  V Transistorn är strypt, OFF, när  $U_{in}$  plötsligt blir hög.  $U_{BE}$  ökar mot sin tröskelspänning,  $I_C$  börjar öka.



Figur 8.31: Slutläge ON,  $U_{in}=5$ V Transistorn har gått längs belastningslinjen genom det aktiva området för att slutligen bottna, transistorn hamnar i läge ON. $U_{ut}$  blir låg.



Figur 8.32: Transistorn switchar från OFF till ON. Pilen markerar det transienta förloppet.



Figur 8.33: npn transistorns arbetssätt

• Resistor-transistor logik, RTL: NOR



Figur 8.34:  $U_{ut} = 0$  V om endera eller bägge transistorer är ON. Detta är fallet om endera eller bägge input spänningar är höga.

Utgångens spänning		
U <sub>A</sub>	$\mathrm{U}_B$	$\mathrm{U}_{ut}$
0 V	0 V	5 V
5 V	0 V	0 V
0 V	5 V	0 V
5 V	5 V	0 V

Truth table: NOR

А	В	Т
0	0	1
1	0	0
0	1	0
1	1	0







Relä OR



# Kapitel 9

# Metall-halvledarövergången

Ingen halvledarkomponent är användbar om vi inte kan komma i kontakt med den. En pn-övergång kräver två kontakter till de två olika dopade områdena och den bipolära transistorn tre. En kontakt består ofta av metall som lagts på halvledarytan. Vi ska här kortfattat beskriva egenskaperna hos en sådan metall-halvledarövergång.

Figur 9.1 visar en tänkt metall-halvledarövergång innan termisk jämvikt har etablerats, dvs inga elektroner har ännu tillåtits passera gränsytan. Vi betraktar en n-typ halvledare där Ferminivån ligger i övre halvan av bandgapet. Vakuumnivån representerar energin hos en fri elektron i vila som inte känner av vare sig metallens eller halvledarens inflytande. Metallens och halvledarens utträdesarbeten är olika. Vi antar att gränsytan är ideal; dvs det finns inte något oxidskikt på halvledaren och inga yttillstånd i halvledaren. Där periodiciteten som ledde till bandstruktur är bruten t ex av en yta eller en kristalldefekt är bandgapet ju inte längre förbjudet. Det betyder att det kan förekomma *lokala* elektroniska tillstånd, med energier i bandgapet, på halvledarens yta.



Utträdesarbete metall:  $\phi_m$  oxid:  $\phi_s$ 

Elektronaffinitet:  $\chi$ Ferminivå, metall:  $E_{Fm}$ halvledare:  $E_{Fs}$ 

Figur 9.1: Metall-halvledarövergång innan jämvikt har etablerats.

När vi släpper loss elektronerna kommer de att röra sig från halvledaren med högre Ferminivå,  $E_{Fs}$ , till metallen med lägre Ferminivå,  $E_{Fm}$ . Detta talar statistisk fysik om för oss. Jämvikt uppnås när Ferminivåerna blivit lika, se figure 9.3. Detta innebär att övergångens kontaktpotential är

$$\psi_0 = (\phi_m - \phi_s)/\epsilon$$

Tittar vi i detalj på vad som händer så diffunderar elektroner i halvledarens ledningsband över till metallen med avsikt att sänka sin energi. Elektronerna lämnar då ett utarmningsområde med positiva donatorjoner efter sig. Motsvarande negativa laddning i metallen kan endast förekomma på metallens yta. I metallen och i halvledarens neutrala område är det elektriska fältet  $\mathscr{E} = 0$ . Enligt Gauss lag gäller

$$\frac{d\mathscr{E}}{dx} = \frac{\rho}{\varepsilon_s}$$

där  $\varepsilon_s$ är halvledarens dielektricitetskonstant. I utarmningsområdet är

$$\rho = eN_D$$

och en första integration ger det elektriska fältet som

$$\mathscr{E}(x) = \frac{eN_D}{\varepsilon_s}(x-w)$$

där vi tagit hänsyn till att  $\mathscr{E}(w) = 0$ . Eftersom

$$\mathscr{E} = -\frac{d\psi}{dx}$$

leder en andra integration till elektriska potentialen

$$\psi(x) - \psi(0) = -\frac{eN_D}{2\varepsilon_s}(x-w)^2 + \frac{eN_D}{2\varepsilon_s}w^2.$$

Vi väljer randvillkoren  $\psi(0)=0$  och  $\psi(w)=\psi_0$  med kontakt<br/>potentialen

$$\psi_0 = \frac{\phi_m - \phi_s}{e}.$$

Utarmningsområdets vidd w bestäms då ur

$$\psi(w) = \frac{eN_D}{2\varepsilon_s}w^2 = \frac{\phi_m - \phi_s}{e}$$

eller

$$w^2 = \frac{2\varepsilon_s}{eN_D}\psi_0 = \frac{2\varepsilon_s}{eN_D}\frac{\phi_m - \phi_s}{e}$$



Figur 9.2: Laddningstäthet  $\rho$ , elektriskt fält  $\mathscr{E}$  och elektrisk potential  $\psi$  i metallhalvledarövergången



Figur 9.3: Banddiagram för en metall-halvledarövergång i termisk jämvikt.

Den negativa ytladdningen i metallen och den positiva laddningen på grund av de joniserade donatorerna i utarmningsområdet leder till en elektrisk potential och ett elektriskt fält som förhindrar ytterligare elektrondiffusion. Adderar vi den potentiella energin

$$E_{pot}(x) = -e \cdot \psi(x)$$

till banden ser vi att banden böjs och en barriär uppstår vid gränsytan.

Vi ser förresten att sambanden som ovan gav det elektriska fältet, elektriska potentialen och utarmningsområdets vidd är precis desamma som för en  $p^+n$ -övergång.

# 9.1 Metall-halvledarövergången med pålagd spänning

I halvledarens neutrala område gäller:

$$n = N_D$$

I termisk jämvikt är flödet av elektroner som kan ta sig över potentialbarriären från halvledaren proportionellt mot elektronkoncentrationen i halvledaren vid gränsytan,  $n_s$ . I det neutrala området, dvs för x > w, gäller

$$n = N_D = N_C e^{-(E_C - E_F)/kT}$$

och eftersom ledningsbandet vid ytan ligger högre med  $\phi_m-\phi_s$ 

$$n_s = N_C e^{-(E_C + \phi_m - \phi_s - E_F)/kT} = N_D e^{-(\phi_m - \phi_s)/kT}$$

Därmed blir flödet från halvledaren till metallen

$$F_{s \to m} \sim n_s = N_D e^{-(\phi_m - \phi_s)/kT}$$



Figur 9.4: Metall-halvledarövergången i termisk jämvikt.



$$eU$$
 I termisk jämvikt måste flödet från metallen till halvledaren vara  
till beloppet lika stort, dvs nettoflödet genom gränsytan måste vara  
-- noll:

 $F_{m \to s} = F_{s \to m} = F_0$ 

Vid framspänning U (metallen positiv relativt halvledaren) ökar flödet

Figur 9.5: Metall-halvledarövergången vid framspänning U



Figur 9.6: Metall-halvledarövergången vid backspänning U < 0

$$F_{s \to m} \sim N_D e^{-(\phi_m - \phi_s - eU)/kT} = F_0 e^{eU/kT}$$

och strömmen är proportionell mot nettoflödet

$$I \sim F_{s \to m} - F_{m \to s}$$
 eller  
 $I = I_0(e^{eU/kT} - 1)$ 

Vid backspänning U < 0 blir, för e|U| > kT

$$I \approx -I_0$$

Metall-halvledarövergången fungerar liksom pn-övergången som en likriktare och används i likriktande Schottky-dioder som är snabba och har låg tröskelspänning. Snabbheten beror på att det inte förekommer någon injektion av minoritetsladdningsbärare och därmed ingen upplagring av laddning som kan fördröja omslaget mellan fram- och backspänt läge. Strömmen i framriktningen domineras av majoritetsladdningsbärarna, likriktaren är unipolär. Den låga tröskelspänningen (kring 0,2 V) medför att små framspänningar ger stora strömmar.

Även om man kan tänka sig situationer där metall-halvledarövergången enligt teorin inte borde vara likriktande, t ex när  $\phi_m < \phi_s$ , så verkar de flesta metall-halvledarövergångarna vara likriktande. Detta kan bero på oönskade tunna gränsskikt – t ${\rm ex}$ ett oxidlager – eller på alla de elektroniska yttillstånd som kraftigt kan påverka laddningsfördelningen och därmed banddiagrammet vid övergången.

Det visar sig också experimentellt att barriärhöjden för en metallhalvledarövergång (n-typ)

$$\phi_{Bn} = \phi_m - \chi$$

endast svagt beror på metallens utträdesarbete  $\phi_m$ .

Ohmska kontakter görs i praktiken ofta genom att dopa halvledaren hårt under kontakten så att barriären blir mycket tunn och elektronerna lätt kan tunnla igenom. Utarmningsområdets vidd wgavs ju av

$$w = \sqrt{\frac{2\varepsilon_s}{eN_D}\psi_0}$$

och därmed minskar w när vi ökar  $N_D$ .

## 9.1. METALL-HALVLEDARÖVERGÅNGEN MED PÅLAGD SPÄNNING235



Figur 9.7: Barriärhöjden visar ett svagare beroende på metallens utträdesarbete än vad vår enkla modell ger.

### 236 KAPITEL 9. METALL-HALVLEDARÖVERGÅNGEN

# Kapitel 10

# Metall-Oxid-Halvledarövergången

Metall-oxid-halvledarövergången (MOS-övergången) utgör den centrala delen i en MOSFET (Metal-Oxide-Semiconductor Field Effect Transistor) och är därmed en av de viktigaste strukturerna i modern elektronik. En MOS-struktur utgörs av en metallkontakt ovanpå en halvledare belagd med en oxid, så att ingen ström kan passera mellan halvledaren och metallen. Genom att variera spänningen på metallkontakten kan man styra ledningsförmågan i halvledaren.

Figur 10.1 visar bandstrukturen för en ideal MOS-övergång vid termisk jämvikt. Som exempel betraktar vi här en MOS med en pdopad halvledare. *Ideal* innebär bl a att utträdesarbetet hos metallen,  $\phi_m$ , är lika med utträdesarbetet i halvledaren,  $\phi_s$ . I termisk jämvikt är Ferminivåerna i metallen,  $E_{Fm}$ , och i halvledaren,  $E_{Fs}$ , lika. I en ideal MOS-struktur blir då bandkanterna i halvledaren flata i termisk jämvikt – vi har ingen bandböjning.

För en ideal MOS antar vi vidare att det inte finns några laddningar i eller på oxiden, samt att oxiden inte tillåter någon passage av laddningsbärare.



- Utträdesarbete metall:  $\phi_m$ halvledare:  $\phi_s$
- Elektronaffinitet:  $\chi$

Ideal MOS struktur:
1) φ<sub>m</sub> = φ<sub>s</sub>
2) Inga laddningar i oxiden
3) Inga likströmmar genom oxiden

Oxidtjocklek:  $d \sim 0.01 - 1.0 \mu \text{m}$ 

Figur 10.1: En ideal MOS struktur i termisk jämvikt.  $E_{Fi}$  är halvledarens intrinsiska Ferminivå.

Oxiden (elektriskt isolerande) är omgiven av två ledande material och strukturen påminner därmed om en plattkondensator. Om vi lägger en spänning över en kondensator kommer laddningar att samlas på plattorna. Sambandet mellan laddning per ytenhet Q och spänning U ges av

För en plattkondensator gäller:

 $Q = C \cdot U.$ 

Vi använder Q (laddning per ytenhet) och C (kapacitans per ytenhet). Alltså:

Q har enheten C·m<sup>-2</sup>

C har enheten  ${\rm F}{\cdot}{\rm m}^{-2}$ 



Figur 10.2: MOS-övergången vid ackumulation. Banddiagram, elektrisk potential samt laddning i metallen,  $Q_m$ , och i halvledaren,  $Q_s$ .

$$QA = CA \cdot U$$

där kondensatorns kapacitans per ytenhet, C är relaterad till avstånden mellan plattorna d genom

$$C = \frac{\varepsilon_r \varepsilon_0}{d}.$$

A är plattornas area och  $\varepsilon_r$  permittivitetstalet (relativa dielektricitetskonstanten) hos mediet mellan plattorna.

Lägger vi en spänning över en MOS-struktur (som ofta kallas MOSkondensator) kommer laddningar att samlas nära oxidens ytor på liknande sätt. I både metallen och halvledaren finns ju rörliga laddningsbärare som påverkas av ett elektriskt fält. Inga laddningar kan dock passera oxiden, dvs inga strömmar flyter, och därmed förblir både metallen och halvledaren i jämvikt var för sig. Vi har då konstanta Ferminivåer i halvledaren respektive i metallen.

Vi ska nu fundera lite mer i detalj på vad som händer då vi lägger en spänning över strukturen. I det följande tänker vi oss att halvledaren är jordad, och att vi lägger en positiv eller negativ spänning på metallen.

### 10.1 MOS-övergången med pålagd spänning

#### 10.1.1 Ackumulation

Ackumulation innebär att vi samlar majoritetsladdningsbärare intill oxiden. För att åstadkomma det för en p-typ halvledare lägger vi en spänning U < 0 på metallen (figur 10.2).

I en elektrisk potential ändras elektronernas potentiella energi med

$$\Delta E_{pot} = -eU$$

Alla energier för elektrontillstånd i metallen – inklusive Ferminivåns position – förskjuts med e|U| mot högre energi då U < 0 (jämför med hur Ferminivåerna förskjuts i en pn-övergång). Potentialen i metallen måste vara konstant vid jämvikt. Om potentialen är varierande finns det också elektriska fält, som orsakar strömmar som omfördelar laddningar. Dessa skulle flyta tills metallens inre blir fältfritt. Hela metallen får alltså potentialen  $\psi = U$ .

Fria elektroner i metallen attraheras av den (relativt sett) positiva potentialen i halvledaren och samlas i ett tunt skikt på metallens yta (laddning per ytenhet  $Q_m$ ). I halvledaren drar sig majoritetshålen något mot den negativa spänningen på metallen så att hålkoncentrationen ökar och ger en positiv laddning vid ytan mellan oxid och halvledare (laddning per ytenhet  $Q_s$ ).  $Q_m + Q_s = 0$  då vi inte, netto, tillfört eller tagit bort laddningar. Från

$$p = N_V e^{-(E_{Fs} - E_V)/kT}$$

följer att  $E_{Fs} - E_V$  måste minska precis intill ytan och banden böjs upp något, se figur 10.2. Redan en liten ändring av  $E_V$  har på grund av det exponentiella beroendet stor effekt på p.

Vi ritar  $Q_m$  och  $Q_s$  som höga smala staplar, dvs hög koncentration i ett tunt skikt intill oxiden. I metallen är elektronkoncentrationen mycket hög, och  $Q_m$  kan betraktas som en nästan ideal ytladdning (sitter bara på ytan). I halvledaren styrs hålkoncentrationen av avståndet mellan  $E_{Fs}$  och  $E_V$ , och hålkoncentrationen ökar exponentiellt då vi närmar oss oxiden. Hålkoncentrationen blir mycket hög intill oxiden, och det räcker det med att bandböjningen sker i ett smalt område för att få laddningen  $Q_s$ .

#### 10.1.2 Utarmning

Vi byter nu polaritet på spänningen på metallen, dvs U > 0. Hålen drar sig bort från metallens positiva potential och lämnar efter sig joniserade, negativt laddade acceptorer så att ett utarmningsområde med bredd w uppstår. Här, precis som i pn-övergången, bestäms laddningstätheten av koncentrationen av joniserade dopatomer. Koncentrationen blir inte så hög som i fallet med de ackumulerade hålen tidigare, och bandböjningen sker då i ett bredare område.

Avståndet mellan  $E_{Fs}$  och  $E_V$  ökar för att reflektera den minskade hålkoncentrationen. På metallens yta uppstår en positiv nettoladdning – avsaknad av elektroner – då metallens fria elektroner drar sig undan den negativa laddningen i halvledaren.

Det elektriska fältet som uppstår mellan laddningen  $Q_m$  i metallen och laddningen  $Q_{sc}$  (där sc står för *space charge* eller rymdladdning) i halvledarens utarmningsområde pekar i positiv *x*-riktning. På samma sätt som i vår diskussion av pn-övergången kan vi då



Figur 10.3: MOS-övergången vid utarmning. Banddiagram, elektrisk potential och laddning i metallen,  $Q_m$ , och i halvledarens utarmningsområde,  $Q_{sc}$ .

dra slutsatsen att potentialen avtar när vi går från metallen in i halvledaren. Potentialen i halvledaren ges av

$$\psi(x) = \frac{eN_A}{2\varepsilon}(w-x)^2,$$

där vi satt  $\psi(w) = 0$  som randvillkor eftersom halvledaren är jordad.

På halvledarens yta är potentialen  $\psi_{yt} = \psi(0)$ . Som vi ser i diagrammet för  $\psi$  kan den totala spänningen U skrivas  $U = V_{ox} + \psi_{yt}$ ; dvs summan av spänningsfallet  $\psi_{yt}$  över halvledarens utarmningsområde och spänningsfallet  $V_{ox}$  över oxiden.

Vi tänker oss nu att vi ökar den positiva spänningen på metallen. Fler hål drar sig då undan så att utarmningsområdets vidd w ökar. Då ökar laddningen  $Q_{sc}$  och därmed även  $Q_m$ . Potentialen  $\psi_{yt}$  vid halvledarytan ökar och banden böjs ner ytterligare.

Vid ett visst värde på spänningen är banden, och därmed den intrinsiska Ferminivån  $E_{Fi}$ , så kraftigt böjda att  $E_{Fi}$  ligger på samma energi som Ferminivån  $E_{Fs}$  på halvledarens yta. Halvledarens yta är då intrinsisk ( $E_{Fs} = E_{Fi}$ ) och de lokala elektron- och hålkoncentrationerna vid ytan är båda lika med  $n_i$ .

#### 10.1.3 Inversion

I inversion har vi lagt så stor spänning på metallen att vi i halvledaren fått fler elektroner än hål intill oxiden. Halvledaren av p-typ har på ytan blivit n-typ. Detta kallas inversion.

I halvledarens neutrala område (x > w) kan vi skriva

$$N_A = p = n_i e^{(E_{Fi} - E_{Fs})/kT} = n_i e^{e\Delta/kT}$$

där vi antagit fullständig jonisation  $(p = N_A)$ . Här betecknar  $e\Delta$ energiskillnaden mellan den intrinsiska Ferminivån  $E_{Fi}$  – eller mitten av bandgapet – och den faktiska Ferminivån  $E_{Fs}$  i halvledaren (se figur 10.4). Då spänningen över MOS-strukturen ökar kommer vi så småningom till en situation då ytpotentialen har ökat till  $2\Delta$ . Då ligger  $E_{Fi}$  vid ytan lika långt under  $E_{Fs}$  som den låg över  $E_{Fs}$ inne i halvledaren.

På halvledarens yta har vi då

$$n = n_i e^{(E_{Fs} - E_{Fi})/kT} = n_i e^{e\Delta/kT}$$

dvs n vid ytan är lika stor som p i halvledarens neutrala område. Den spänning som krävs för att åstadkomma detta kallar vi

Här är det smidigt att använda uttrycken:

$$n = n_i e^{(E_{Fs} - E_{Fi})/kT}$$
$$p = n_i e^{(E_{Fi} - E_{Fs})/kT}$$

som (kapitel 6) är ekvivalenta med

 $n = N_C e^{-(E_C - E_{Fs})/kT}$  $p = N_V e^{-(E_{Fs} - E_V)/kT}$ 

 $U_T$ , tröskelspänningen, för inversion. Vi förväntar oss alltså att en icke-försumbar koncentration elektroner dyker upp i den triangulära potentialgropen nära ytan vid termisk jämvikt. I figur 10.4 syns dessa elektroner som laddningen  $Q_n$ . Det tunna skikt vid ytan där de ansamlade laddningsbärarna finns kallas för inversionskanal.

När U ökar ytterligare (så att  $U > U_T$ ) kommer utarmningsområdets bredd inte att öka nämnvärt. En liten bandböjning vid ytan, dvs en liten ökning av  $E_{Fs} - E_{Fi}$ , är allt som behövs för att öka  $Q_n$  kraftigt. Ökningen av den negativa laddningen i halvledaren sker då i kanalen och inte längre genom att utarmningsområdet växer, utarmningsområdet har nått sin maximala bredd  $w_{max}$ .

Varifrån kommer elektronerna som samlats intill oxiden i inversion? Några har diffunderat från det neutrala området, men här är elektroner i minoritet så det finns inte så många att tillgå. De flesta genereras genom termisk excitation över bandgapet i utarmningsområdet. Hålen som då skapas sveps bort till det neutrala p-området och elektronerna dras ner i gropen. Dessa processer tar dock viss tid. Den tid det tar att nå termisk jämvikt, den så kallade *termiska relaxationstiden*, är ungefär av storleksordningen en sekund.

Om det tar någon sekund för elektronerna att anpassa sig efter en förändrad spänning, hur ser då strukturen ut precis efter det att vi lagt på en spänning  $U > U_T$ ? Det som sker först – och snabbt – är att hålen dras bort från ytan av det ökade elektriska fältet och vi får ett brett utarmningsområde, bredare än  $w_{max}$ . Detta initala tillstånd kallas djup utarmning (deep depletion), figur 10.5. Elektronerna tar en stund på sig att samlas vid ytan. Allt eftersom elektroner dyker upp intill oxiden, huvudsakligen genom termisk generation, ökar laddningen i inversionskanalen och utarmningsområdet krymper tills vi slutligen når termisk jämvikt. Djup utarmning används i CCD-detektorer.

### 10.2 Charge-coupled device (CCD)

En CCD-detektor registrerar den ljusintensitet som träffar olika delar av detektorns yta. Detektorn består av en matris av MOSkondensatorer. Ljus exciterar elektron-hål-par över bandgapet i MOSkondensatorerna. Om ljuset som träffar en del av detektorn har högre intensitet (fler fotoner per sekund) skapas lokalt fler elektron-hål par.

En MOS-kondensator (var och en med en separat metallkontakt) med en halvledare av p-typ visas i figur 10.6. En positiv gatespän-



Figur 10.4: MOS övergången vid inversion, elektrisk potential  $\psi$ och laddning i metallen,  $Q_m$ , i utarmningsområdet,  $Q_{sc}$ , och i kanalen,  $Q_n$ .



Figur 10.5: MOS-struktur vid djup utarmning, innan termisk jämvikt har etablerats.



Figur 10.6: MOS-kondensator på ett p-typ substrat.

ning  $U_G > U_T$  läggs på en av kontakterna. I första ögonblicket skapas ingen inversionskanal utan bara ett brett utarmningsområde. Vi har djup utarmning.

Om MOS-strukturen i djup utarmning belyses skapar fotonerna elektron-hålpar genom att excitera elektroner från valensbandet till ledningsbandet. Varje inkommande foton (med energi större än bandgapet) har en viss chans att skapa ett elektronhål-par. Ju fler fotoner som träffar MOS-strukturen, desto fler elektronhål-par bildas. Om denna excitation sker inom den utarmade delen av halvledaren, där det finns ett elektriskt fält, kommer de skapade hålen att dras iväg till halvledarens neutrala område. De skapade elektronerna knuffas på andra hållet, mot oxiden, och samlas i gropen som det böjda ledningsbandet bildar (se figur 10.5). Antalet elektroner borde vara proportionellt mot ljusets intensitet.

Genom att använda en matris av MOS-kondensatorer kan vi mäta ljusintensitet i olika punkter och då lagra en bild som man t ex gör i en digital kamera. Genom att räkna elektronerna på en viss kondensator får vi ett mått på ljusintensiteten. Räkningen måste ske snabbt, innan systemet når termisk jämvikt, för då kommer det ju för samma styrspänning  $U_G$  över MOS-kapacitanserna att finnas lika många elektroner i varje grop, och vi har förlorat informationen om ljusintensiteten.

Figurerna 10.7 och 10.8 visar principen för hur man för ut information (dvs antalet elektroner på en kondensator), från enskilda MOS-kondensatorer för en viss typ av CCD-detektor. För att förstå principen betraktar vi figur 10.7. Vi lägger en spänning  $U_1 = +10$ V på kontakt 1 (så att vi har djup inversion) och låter ljus träffa strukturen. Då skapas elektron-hål par och elektronkoncentrationen under metallkontakten är karakteristisk för ljusets intensitet. Genom att sedan snabbt lägga en hög spänning  $U_2 = +15$  V på nästa kontakt flödar elektroner över till denna kontakt. Vi har alltså samlat in information om ljusintensiteten vid kontakt 1 och flyttat informationen ett steg genom strukturen.



Figur 10.7: Principen för hur elektroner transporteras inom en CCD-detektor.

I trefas-CCD behövs tre MOS-strukturer för varje bildpunkt (en detekterar ljus och ytterligare två används för förflyttningen) och genom lämpliga spänningspulser (se figur 10.8) kan vi flytta elektronerna från grop till grop, ut till en läskrets vid radens ände, där en förstärkare konverterar laddningsmängden till en spänningssignal. MOS-matrisens rader fungerar som horisontella shiftregister.



Figur 10.8: Spänningspulser flyttar laddningen från MOS till MOS. Den undre delen av figuren visar hur spänningarna  $U_1$ ,  $U_2$  och  $U_3$  varierar, och den övre delen visar hur elektronerna förflyttas mha spänningspulserna.



Figur 10.9: a) Strukturen hos en MOSFET. Metallelektroden (*gate*, inte utsatt i figuren) ligger på oxiden. b) Dopning och utarmningsområde nära halvledarens yta. c) Banddiagram som visar ledningsbandskanten.

### 10.3 MOSFET

Med en spänning  $U_G > U_T$  på metallen kan vi – som vi har sett – skapa en inversionskanal vid halvledarytan under oxiden. I en p-typ halvledare blir då elektronkoncentrationen vid ytan hög. Elektronkoncentrationen ökar sedan med ökande gatespänning, dvs kanalens konduktivitet styrs av gatespänningen.

Figur 10.9 visar en metalloxid-halvledar-fält-effekt-transistor (MOS-FET). Central i transistorn är en MOS-övergång som utgörs av en metallkontakt (gate; G), ett oxidskikt och den underliggande halvledaren, här av p-typ. Genom dopning skapas två  $n^+$  områden; source (S) och drain (D).

I en MOSFET kan man styra laddningstransporten (här av elektroner) mellan source- och drainkontakterna genom att variera spänningen på gate-kontakten. Utan inversionskanalen kan det inte gå ström mellan source och drain. Elektronerna i n<sup>+</sup>-områdena (S och D) ser p-området som en potentialbarriär (ledningsbandskanten ligger högre i energi här; se fig 10.9). Med en positiv spänning på gaten kan vi böja ner bandkanten och skapa en inversionskanal med elektroner som förbinder source och drain.

Det är alltså möjligt att styra strömmen mellan source och drain genom inversionskanalen med hjälp av den gatespänningen. Den $U_T$  är tröskelspänningen för inversion.



Figur 10.10: MOSFET med en styrspänning större än tröskelspänningen. En inversionskanal har bildats vid halvledarens yta under oxiden. Utarmningsområdet utgörs av den streckade arean.

na styrning, som kallas fälteffekt, diskuterades redan på 1920-talet medan MOSFET-strukturer realiserades tillförlitligt först på 1960-talet.

Source och drain  $(n^+)$  bildar pn-övergångar med p-typ halvledaren, där utarmningsområdena (streckade i figur 10.9) väsentligen ligger på den lågdopade sidan av övergången, dvs i p-området.

För  $U_G > U_T$  bildas en inversionkanal under oxiden; se figur 10.10. Som vi har sett i diskussionen om utarmning och inversion finns det också ett utarmningsområde mellan oxiden och den neutrala p-typ halvledaren. Här har de fria hålen trängts undan av den positiva gatespänningen och kvar finns de negativa acceptorjonerna.

Potentialen vid gränsytan till oxiden är

$$\psi_{yt} = \psi_{yt}^0 = \frac{eN_A}{2\varepsilon_s}w^2 \quad \text{när} \quad U_S = U_D = 0.$$

där vi återigen antagit att  $\psi = 0$  i den neutrala p-havledaren pga kontakten till jord. Här är w utarmningsområdets vidd i p-typ halvledaren,  $N_A$  acceptorkoncentrationen och  $\varepsilon_s$  halvledarens dielektricitetskonstant.

Med en positiv drainspänning  $U_D$ , se figur 10.12, rör sig elektroner från source till drain förutsatt att vi mha lämplig gatespänning har skapat en inversionskanal. För låga värden på  $U_D$  ökar strömmen linjärt med drainspänningen. Kurvans lutning bestäms av konduktiviteten i inversionskanalen som i sin tur beror av gatespänningen.

För den följande diskussionen av strömmen mellan source och drain, som ofta bara kallas drainström och betecknas  $I_D$ , behöver vi härleda ett samband som relaterar laddningen på metallen,  $Q_m$ , och därmed även laddningen i halvledaren,  $Q_s$ , till ytpotentialen  $\psi_{yt}$ . För härledningen utnyttjar vi MOS-övergångens likhet med en plattkondensator.

Som vi nämnt tidigare kan vi se den totala spänningen på gaten,  $U_G$ , som summan över spänningsfallet över oxiden,  $V_{ox}$ , och över halvle-

Kom ihåg:  $x_p N_A = x_n N_D$ 

 $\psi$  kan vi ta fram på samma sätt som i avsnittet om utarmning, eller som i kapitlet om pn-övergång. Använd relationerna för  $\rho$ ,  $\mathscr{E}$  och  $\psi$ 



Figur 10.12: MOSFET med liten drainspänning  $U_D = U_1 > 0$ . Kanalen smalnar av när vi närmar oss drain och utarmningsområdets vidd ökar.  $I_D$  ökar nästan linjärt med  $U_D$ .  $\psi(y)$  är elektriska potentialen längs kanalen orsakad av drainspänningen.

darens utarmningsområde,  $\psi_{yt}$ . Över oxiden faller alltså spänningen  $V_{ox} = U_G - \psi_{yt}$  och vi kan skriva laddningen  $Q_m$  (positiv i det här fallet) på metallen som

$$Q_m = (U_G - \psi_{yt}) \cdot C_{ox} \quad \left[\frac{C}{m^2}\right].$$

där gatekapacitansen  $C_{ox}$  (per ytenhet) är

$$C_{ox} = \frac{\varepsilon_{ox}}{d} \qquad \left[\frac{\mathbf{F}}{\mathbf{m}^2}\right]$$

Totalt sett är MOS-kondensatorn oladdad; totala laddningen per ytenhet i halvledaren  $Q_s$  är till beloppet lika stor som  $Q_m$ , men med motsatt tecken

$$Q_s = -Q_m = -(U_G - \psi_{yt}) \cdot C_{ox}$$

där  $Q_s$  är totala laddningen i halvledaren; alltså både laddningen associerad med laddningsbärare i inversionskanalen och med laddningen hos joniserade dopatomer i utarmningsområdet.

Spänningsfallet från source till drain längs med kanalen medför nu att potentialen vid halvledarytan varierar med läget y, den ökar från  $\psi_{yt}^0$  vid source till  $\psi_{yt}^0 + \psi(y)$  vid drain. Vi skriver den som

$$\psi_{yt}(y) = \psi_{ut}^0 + \psi(y)$$

där  $\psi(y)$  varierar från 0 till  $U_D$ , ungefär som skissat i figur 10.12. Laddningen på metallytan blir då också beroende av läget och kan skrivas som Q = CV



Figur 10.11: Potentialen i MOS-strukturen.





Figur 10.13: Med ökad drainspänning minskar strömökningen.

$$Q_m = |Q_s| = (U_G - \psi_{yt}^0 - \psi(y)) \cdot C_{ox}$$

Kanalen påverkas lokalt som funktion av y på två sätt:

1) Den totala negativa laddningen i halvledaren,  $Q_s$ , minskar nära drain eftersom  $\psi(y)$  ökar. Detta följer ur relationen ovan.

2) För positiv  $U_D$  blir pn-övergången vid drain backspänd och utarmningsområdet ökar i vidd. Då ökar den negativa laddningen  $Q_{sc}$ i utarmningsområdet nära drain.

Tillsammans leder dessa två effekter till att laddningen i kanalen,  $Q_n$ , måste minska med ökande  $\psi(y)$ , dvs när vi går från source till drain, eftersom

$$Q_n = Q_s - Q_{so}$$

Kom ihåg:  $J = \sigma \mathscr{E}$ 

skissades potentialen  $\psi(y)$ ökar med ökad drainspän strömmen ökar nu endast Utan att härleda något de kanalens smalaste ställe är  $Q_s$  minskat så mycket, oc kanalen  $Q_n = Q_s - Q_{sc}$ vara tom, "pinched-off" e maximala värde. Ökar vi drainspänningen y ningen väsentligen att han är strypt, snarare än över ningsområdet är ju resist

Figur 10.14: Drainspänningen har ökat från  $U_1$  (figur 10.12) till  $U_2$ . MOSFET når pinch-off läget när inversionskanalen är strypt vid drain.  $I_D$  mättas.

Eftersom kanalen smalnar av när vi närmar oss drain så kommer konduktiviteten att minska nära drain. Strömmen är förstås oberoende av y. Då måste det elektriska fältet längs kanalen öka med yför att kompensera för minskningen i konduktivitet. Eftersom

$$\mathscr{E} = -\frac{d\psi}{dx}$$

skissades potentialen  $\psi(y)$  som i figur 10.12. Resistansen i kanalen ökar med ökad drainspänning (pga att kanalen smalnar av), och strömmen ökar nu endast långsamt med drainspänning.

Utan att härleda något detaljerat uttryck för  $\psi(y)$  kan vi inse att kanalens smalaste ställe är vid drain. Vid en viss drainspänning har  $Q_s$  minskat så mycket, och  $Q_{sc}$  ökat så mycket, att laddningen i kanalen  $Q_n = Q_s - Q_{sc} = 0$ . Då kommer kanalen vid drain att vara tom, "*pinched-off*" eller strypt. Strömmen har här nått sitt maximala värde.

Ökar vi drainspänningen ytterligare, så kommer ökningen av spänningen väsentligen att hamna över utarmningsområdet där kanalen är strypt, snarare än över den ledande biten av kanalen. I utarmningsområdet är ju resistiviteten mycket högre än i den ledande kanalen. Vi får då bilden som i figur 10.14. En ledande om än avsmalnande kanal leder fram elektroner till stupet vid drain. Stupets höjd påverkar inte strömmen, avgörande är den ledande kanalbitens förmåga att leverera elektroner. Drainströmmen mättas, vi når "saturation" - ett ord som här har en helt annan betydelse än i diskussionen av den bipolära transistorn!

Vi finner slutligen transistorns utgångskarakteristik, se figur 10.15.



Figur 10.15: Utgångskarakteristik hos en MOSFET.  $I_D$  är plottad som funktion av  $U_D$  med gatespänningen  $U_G$  som parameter.  $U_T$  är tröskelspänningen för inversion. Karakteristiken är mycket lik den bipolära transistorns karakteristik.

### 10.4 CMOS

Vi diskuterade tidigare inverteraren baserad på en bipolär transistor. Vi fann att det i dess ena stationära tillstånd där transistorn var ON flöt en stor kollektorström. Detta leder till en stor effektförbrukning och dessutom värmeutveckling i transistorn. I integrerade kretsar måste vi minimera effektförbrukningen och här har komplementär MOS teknik - CMOS - haft en oerhörd betydelse. Vi avslutar diskussionen av MOS-strukturer med en kort introduktion till CMOS.

En CMOS-inverterare består av en MOSFET av vardera slaget; en PMOS och en NMOS. Beteckningarna NMOS och PMOS anger vilken typ av laddningsbärare som finns i kanalen vid inversion. En NMOS är alltså en transistor gjord med p-substrat, som vid inversion blir n-typ.

I det följande exemplet är båda transistorerna av *enhancement-typ*, dvs en gatespänning behövs för att skapa en inversionskanal - se högra delen av figur 10.16.

En inverterare har en ingångsspänning,  $v_{in}$ , och en utgångsspänning,  $v_{ut}$ . Inverteraren ska, precis som det låter, ge en hög utgångsspänning för en låg ingångsspänning och tvärtom.

Vi börjar med att betrakta låg ingångsspänning,  $v_{in} \approx 0$  V. PMOStransistorn är i inversion. Gatespänningen här sätts direkt av  $v_{in}$ , medan halvledardelen är satt till potentialen +5 V av spänningen  $U_{DD}$ . Relativt halvledaren är gate-kontakten negativ och därmed



Figur 10.16: CMOS inverterare som består av två enhancement transistorer, en p-kanal och en n-kanal. Drainströmmen  $I_D$  är ritad som funktion av gatespänningen  $U_G$  till höger med tröskelspänningar  $U_{Tp}$  och  $U_{Tn}$  för p- resp. n-kanal transistorn. Source och substrat är sammankopplade.

PMOS - p-kanal i n-typ bulk halvledare.

Gatespänningen måste vara negativ relativt substratet och till absolutvärdet större än den negativa tröskelspänningen  $U_{Tp}$ för att skapa en hål-kanal.

NMOS - n-kanal i p-typ bulk halvledare.

Gatespänningen måste vara positiv och större än tröskelspänningen  $U_{Tn}$  för att skapa en elektron-kanal.

Typiskt:  $U_{Tp} \sim -1.0 \text{ V}$  $U_{Tn} \sim 1.0 \text{ V}$  samlas hål under oxiden så att transistorn är ledande.

NMOS-transistorn är däremot stängd. Spänningen på gaten, även denna definierad av  $v_{in}$ , är densamma som potentialen i halvledardelen (jordad; har potentialen noll) och därmed bildas ingen inversionkanal. Kontakten där vi läser av  $v_{ut}$  är isolerad från jord (på grund av stängd NMOS), men hamnar på samma potential som  $U_{DD}$  på grund av den ledande PMOS-transistorn;  $v_{ut} \approx 5$  V.

För hög ingångsspänning,  $v_{in} \approx 5$  V, är PMOS-transistorn stängd och utgångskontakten  $(v_{ut})$  är därmed isolerad från  $U_{DD}$ . NMOStransistorn är nu ledande ty gaten har en potential på +5 V, medan halvledardelen i NMOS-transistorn fortfarande är jordad. Utgångsspänningen  $v_{ut}$  sätts nu till potentialen på jord,  $v_{ut} \approx 0$  V.

Oavsett värde på  $v_{in}$ , högt eller lågt, är en av transistorerna alltid stängd (icke-ledande). Det finns alltså aldrig möjlighet för en stor ström att gå t ex mellan  $U_{DD}$  och jord och den statiska strömförbrukningen är mycket låg. I själva omslaget, när vi byter värde på  $v_{in}$ , omfördelas laddningar vilket innebär att strömmar flödar i transistorerna.

Åven i statiskt läge kan det dock finnas en låg effektutveckling. En stängd transistorer har en viss om än liten ledningsförmåga - ett läckage - mellan source och drain även då vi inte har inversion - så kallad *subthreshold current*. Det är angeläget att den är så liten som möjligt för att minimera effektutvecklingen i transistorn.

# Appendix A

# Värmeledningsförmåga

Varierar temperaturen längs med ett prov kommer ett flöde av termisk energi

$$J^Q = \frac{Q}{A \cdot \Delta t}$$

från den heta änden till den kalla att försöka utjämna temperaturen. Vi betraktar nu flödet i punkten x = 0 och antar att elektroner svarar för transporten.

Elektroner som passerar x = 0 från vänster hade sin senaste kollision för en tid  $\tau$  sedan, dvs vid  $x = -\ell$  där deras energi i medel var

$$E = -\frac{dE}{dx}\ell.$$

Denna energi bär de nu med sig förbi x = 0. Elektronerna som passerar från höger hade sin senaste kollision vid  $x = \ell$  där deras medelenergi var

$$E = +\frac{dE}{dx}\ell.$$

Vi antar – lite förenklat – att 1/3 av elektronerna rör sig i var och en av koordinataxlarnas riktning. Hälften av dem som rör sig längs med x-axeln rör sig åt vänster, hälften åt höger. Flödet av termisk energi (i positiv x-riktning) kan då skrivas:

$$J^{Q} = \frac{1}{6}nv\left(E - \frac{dE}{dx}\ell\right) - \frac{1}{6}nv\left(E + \frac{dE}{dx}\ell\right)$$
$$= -\frac{1}{3}nv \cdot \frac{dE}{dx}\ell.$$

Q - termisk energi ("heat")



E: elektroners medelenergi n: elektronkoncentration v: elektroners medelfart  $\ell$ : fria medelväglängden  $C_V$ : värmekapacitet per volymsenhet. Med värmekapaciteten per volymsenhet

$$C_V = n \frac{dE}{dT}$$

finner vi slutligen

$$J^{Q} = -\frac{1}{3}v\ell n \frac{dE}{dT} \frac{dT}{dx} = -\frac{1}{3}v\ell C_{V} \frac{dT}{dx} = -\kappa \frac{dT}{dx}$$
värmeledningsförmåga:  $\kappa = \frac{1}{3}C_{V}v\ell$ 

Speciellt för elektroner gäller

$$E = \frac{m}{2}v_{th}^2 = \frac{3}{2}kT$$

$$C_V = n \frac{dE}{dT} = \frac{3}{2}nk, \qquad \ell = v_{th}\tau$$

vilket ger

,

$$\kappa = \frac{1}{2}nkv_{th}^2\tau = \frac{3}{2}\frac{nk^2\tau T}{m} \quad \text{och}$$

Kom ihåg:  $\sigma = \frac{n e^2 \tau}{m}$ 

$$\frac{\kappa}{\sigma T} = \frac{3}{2} \left(\frac{k}{e}\right)^2 = 1.11 \cdot 10^{-8} \mathrm{W}\Omega/\mathrm{K}^2 = \mathrm{konst}$$

Det experimentella Lorenz talet är

$$\frac{\kappa}{\sigma T} \sim 2 \cdot 10^{-8} \mathrm{W}\Omega/\mathrm{K}^2.$$

Den empiriska relationen

$$\frac{\kappa}{\sigma T} = \text{konst.}$$

kallas Wiedemann-Franz lag (1853) och Drudes förklaring upplevdes som ett starkt argument för hans modell – även om det sägs att överensstämmelsen ursprungligen varit bättre tack vare ett räknefel.
### Appendix B

## Lådpotential i ett elektriskt fält

En oändlig lådpotential har lösningar som visas i figur B.1a. Lägger vi ett elektriskt fält över lådan,

$$\mathscr{E}=-\frac{dU}{dx}$$

kommer lådans botten att slutta som den streckade linjen i figur B.1b eftersom elektronens potentiella energi ändrar sig med

$$\Delta E_{pot} = (-e) \cdot U(x).$$

Elektronen kommer då att i sitt grundtillstånd att vilja halka ner till det nedre vänstra hörnet. Det betyder att vågfunktionen måste öka något i lådans vänstra halva och minska i den högra halvan. Vi kan approximativt beskriva vågfunktionen i det störda grundtillståndet som en linjärkombination av de två lägsta ostörda vågfunktionerna

 $\Psi(x) = c_1\psi_1(x) + c_2\psi_2(x)$ 

Det resulterande grundtillståndet kommer att se ut ungefär som i figur B.1c. Vi har kommit en liten bit på väg mot en exakt lösning, men det hade säkert gett vågfunktionen större flexibilitet om vi hade tagit med fler termer i vår utveckling, dvs fler ostörda vågfunktioner. Om vi t ex hade blandat in den ostörda vågfunktionen 4 så skulle elektronen kunna komma ännu lite närmare kanten vilket nog ytterligare hade minskat dess energi en aning.



Figur B.1: a) De fyra lägsta vågfunktionerna i en o<br/>ändlig lådpotential. b) Lådpotentialen i ett elektriskt fält. c) Grund<br/>tillståndets vågfunktion som linjärkombination av de två lägsta ostörda vågfunktionerna.

### Appendix C

### Diffusionsekvationen

Vi ska här mycket kortfattat antyda var diffusionsekvationen för minoritetsladdningsbärarna i en halvledare kommer ifrån. *Kontinuitetsekvationen* håller reda på att laddningar inte bara kan försvinna. Minskar laddningstätheten i något område måste den ha strömmat in i eller ut ur området.

$$\nabla \cdot \mathbf{J} = -\frac{\partial \rho}{\partial t}$$

eller i en dimension (i x-riktningen) med  $J = \mathbf{J}_x$ 

$$\frac{\partial J_x}{\partial x} = -\frac{\partial \rho}{\partial t}$$

När det gäller elektroner och hål i en halvledare måste vi ta hänsyn till att både elektron- och hålkoncentrationer även kan ändras genom generations- och rekombinationsprocesser. För elektronkoncentrationen gäller då

$$\frac{\partial n}{\partial t} = -\frac{1}{e} \frac{\partial J_e}{\partial x} + G_e - R_e$$

och för hålkoncentrationen

$$\frac{\partial p}{\partial t} = -\frac{1}{e} \frac{\partial J_h}{\partial x} + G_h - R_h$$

där G och R är generations- och rekombinationstakter för respektive partikel.

Strömtätheterna för elektroner och hål, som innehåller ett drift- och ett diffusionsbidrag ges av

$$J_e = e\mu_e n\mathscr{E} + eD_e \frac{dn}{dx}$$

 $\operatorname{och}$ 

$$J_h = e\mu_h p\mathscr{E} - eD_h \frac{dp}{dx}$$

Vi använder följande notation:  $p_0$  och  $n_0$  betecknar hål- och elektronkoncentrationer i jämvikt.

p och n betecknar motsvarande koncentrationer när jämvikt är rubbad som t ex genom injektion. Detta är alltså de aktuella koncentrationerna. Vi skriver för skillnaderna  $\Delta p = p - p_0$  och  $\Delta n = n - n_0$ .  $G_h$  och  $G_e$  – generationstakter,  $R_h$  och  $R_e$ –

rekombinationstakter,  $J_h$  och  $J_e$  – hål- resp. elektronströmtätheter,  $D_h$  och  $D_e$  – diffusionskonstanter,

 $\mu_h \operatorname{och} \mu_e - \operatorname{mobiliteter.}$ 

Vi sätter in strömtätheterna i de modifierade kontinuitetsekvationerna och får

$$\frac{\partial n}{\partial t} = \mu_e n \frac{\partial \mathscr{E}}{\partial x} + \mu_e \mathscr{E} \frac{\partial n}{\partial x} + D_e \frac{\partial^2 n}{\partial x^2} + G_e - R_e$$

 $\operatorname{och}$ 

$$\frac{\partial p}{\partial t} = -\mu_h p \frac{\partial \mathscr{E}}{\partial x} - \mu_h \mathscr{E} \frac{\partial p}{\partial x} + D_h \frac{\partial^2 p}{\partial x^2} + G_h - R_h$$

I jämvikt har vi laddningsneutralitet

$$p_0 + N_D^+ - n_0 - N_A^- = 0$$

I icke-jämvikt kan vi mycket väl ha nettoladdningar, som då enligt Coulombs lag leder till ett inre – alltså ej pålagt – elektriskt fält enligt

$$\frac{\partial \mathscr{E}}{\partial x} = \frac{e(p + N_D^+ - n - N_A^-)}{\varepsilon_s} = \frac{e(\Delta p - \Delta n)}{\varepsilon_s}$$

Vi antar nu att vi har en n-typ halvledare som är homogen, dvs p och n är oberoende av rum och tid. Vad händer nu om en avvikelse från laddningsneutralitet har gett upphov till ett sådant inre fält i vår halvledare? Fältet kommer att generera strömmar och dessa kommer säkert att domineras av bidraget från elektronerna, dvs majoritetsladdningsbärarna. Hålbidraget är ju storleksordningen 10 tio-potenser mindre. Elektronerna kommer då snabbt att neutralisera hålen och laddningsneutralitet återställs – åtminstone nästan,

$$\Delta p \cong \Delta n.$$

Detta innebär nu att vad hålen än hittar på, så följer elektronerna efter för att upprätthålla neutralitet. Det spelar ingen roll om hålen diffunderar eller drivs av ett (pålagt) elektriskt fält, elektronerna hänger med. Vet vi då vad hålen har för sig, så vet vi allt som händer. Därför är det så viktigt att hålla reda på minoritetsladdningsbärarna!

Nu kan vi i princip översätta resonemanget till formalism. Eftersom vi inte exakt vet hur stor laddningsobalansen är och därmed inte heller vet exakt hur stort det inre elektriska fältet är, eliminerar vi lämpligen först  $\partial \mathscr{E} / \partial x$  ur de två ekvationerna vi kommit fram till ovan. Vi antar vidare att laddningsneutralitet gäller, dvs

$$\Delta p = \Delta n.$$

Efter lite algebra, och med Einsteins relation mellan D och  $\mu$ , kommer vi till slut fram till *diffusionsekvationen* för hålen i en *n-typ* halvledare

$$\frac{\partial p}{\partial t} = D_h \frac{\partial^2 p}{\partial x^2} - \mu_h \mathscr{E} \frac{\partial p}{\partial x} + G_h^{\text{ext}} - \frac{p - p_0}{\tau_h}$$

 $G_h^{\text{ext}}$  och  $G_e^{\text{ext}}$  – generationstakter på grund av yttre inverkan,  $\tau_h$  och  $\tau_e$  – hålens respektive elektronernas livstid.



Figur C.1: I en belyst halvledare skapas elektron-hålpar vid ytan. Överskotts-minoritetsladdningsbärarna diffunderar in i halvledaren och rekombinerar.

och analogt för elektronerna i ett p-typ material

$$\frac{\partial n}{\partial t} = D_e \frac{\partial^2 n}{\partial x^2} + \mu_e \mathscr{E} \frac{\partial n}{\partial x} + G_e^{\text{ext}} - \frac{n - n_0}{\tau_e}.$$

Låt oss titta på ett exempel där vi använder diffusionsekvationen. Situationen är mycket lik injektionen i pn-övergången. Vi betraktar en *n-typ* halvledare under belysning. Belysningen genererar ett laddningsbäraröverskott vid ytan, x = 0, och vi antar att systemet befinner sig i ett stationärt tillstånd.

Belysningen skapar i det stationära tillståndet ett konstant överskott av hål bara på ytan,

/ **T** 

$$p(x=0) = p(0) = \text{konstant}$$

Långt inne i halvledaren har överskottet dött ut, dvs

$$p(x \to \infty) = p_0.$$

Elektronerna neutraliserar hålen, halvledaren är fältfri.

Vi skriver då diffusionsekvationen

$$\frac{\partial p}{\partial t} = 0 = D_h \frac{\partial^2 p}{\partial x^2} - \frac{p - p_0}{\tau_h}.$$

Lösningen blir en avklingande exponentialfunktion

$$p(x) = p_0 + \{p(0) - p_0\} \cdot e^{-x/L_h}$$

där diffusionlängden definierades genom sambandet

$$L_h = \sqrt{D_h \tau_h}.$$

#### 258 APPENDIX C. DIFFUSIONSEKVATIONEN

### Appendix D

# Generation och rekombination – livstider

Låt oss rubba jämvikten i en n-typ halvledare genom belysning, som vi antar genererar elektron-hål par homogent i hela halvledarprovet. Detta ändrar elektronkoncentrationen från sitt jämviktsvärde  $n_{n0}$  till  $n_n$ , och  $p_{n0}$  till  $p_n$ ,

$$n_n = n_{n0} + \Delta n, \qquad p_n = p_{p0} + \Delta p.$$

Belysning skapar ingen nettoladdning vi har fortfarande laddningsneutralitet,

$$\Delta n = \Delta p.$$

Det pågår hela tiden generations- och rekombinationsprocesser i halvledaren. Elektroner i valensbandet utnyttjar den termiska energi som flyter omkring och hoppar upp till ledningsbandet, andra elektroner som redan är i ledningsbandet hittar ett hål i valensbandet och rekombinerar. Vid rekombinationen frigörs energi, som kan leda till att fotoner utsänds, men som oftast bara värmer upp kristallen. Med  $G_{th}$  betecknar vi den termiska generationstakten och med R rekombinationstakten. Takterna anger antalet processer per volyms- och tidsenhet. Vad kan vi säga om  $G_{th}$  och R? Det är uppenbart att rekombinationstakten förutsätter att det finns elektroner i ledningsbandet och att det samtidigt finns hål i valensbandet. R borde då ges av produkten av koncentrationen besatta tillstånd i ledningsbandet,  $n_n$ , och koncentrationen tomma tillstånd i valensbandet,  $p_n$ . Vi skriver

$$R = \beta n_n p_n$$

med en proportionalitetskonstant  $\beta$ , som beskriver detaljerna i övergångsprocessen, men den kan vi inte säga något om här.



Figur D.1: Belysning skapar elektron-hålpar med generationstakten  $G_L$ . Denna balanseras i jämvikt av nettorekombinationen med takten  $R - G_{th}$ .

$$G_{th}, R - [m^{-3}s^{-1}]$$
  
 $G_L - [m^{-3}s^{-1}]$ 

Motsvarande argument ger att den termiska generationstakten  $G_{th}$ beror både på koncentrationen elektroner i valensbandet som kan exciteras och på koncentrationen tomma tillstånd i ledningsbandet dit elektronerna kan exciteras. Men i en halvledare är ju valensbandet normalt alltid (nästan) helt fyllt och ledningsbandet alltid (nästan)helt tomt. Generationstakten är då konstant även om vi rubbar systemets jämvikt.  $G_{th}$  har då även samma värde vare sig halvledaren är intrinsisk eller extrinsisk.

I termisk jämvikt – ingen nettorekombination eller nettogeneration – måste för en intrinsisk halvledare gälla

$$G_{th} = R = \beta n_i^2,$$

och för en extrinsisk n-typ halvledare

$$G_{th} = R = \beta n_n p_n.$$

Om  $G_{th}$  är den samma så har våra funderingar lett oss fram till massverkans lag på en annan väg än tidigare eftersom

$$G_{th} = \beta n_n p_n = \beta n_i^2.$$

Detta säger oss att produktansatsen vi gjort är rimlig.

Låt belysningen nu skapa  $G_L$  elektron-hål par per tidsenhet och volymsenhet i vår n-dopade halvledare. Hålkoncentrationen  $p_n$  ändrar sig då med tiden enligt

$$\frac{dp_n}{dt} = G_L + G_{th} - R$$

För ett stationärt tillstånd gäller

$$0 = \frac{dp_n}{dt} = G_L + G_{th} - R$$

där R, när vi nu rubbat jämvikten, ges av

 $R = \beta n_n p_n$ 

 $n_n = n_{n0} + \Delta n$  $p_n = p_{n0} + \Delta p$ 

Antagande 1.

 $n_{n0} \gg p_{n0}$ 

Halvledaren är n-dopad,

$$R = \beta(n_{n0} + \Delta n)(p_{n0} + \Delta p)$$
  
=  $\beta n_{n0}p_{n0} + \beta(n_{n0} + p_{n0} + \Delta p) \cdot \Delta p$ 

där vi utnyttjat att  $\Delta n = \Delta p$ .

Om halvledaren är n-dopad gäller

 $n_{n0} \gg p_{n0}.$ 

Dessutom begränsar vi analysen till situationer där överskottet av laddningsbärare som belysningen genererar visserligen ökar hålkoncentrationen kraftigt

$$p_{n0} \ll \Delta p$$

men ändå är mycket mindre än majoritetsladdningsbärarkoncentrationen  $n_{n0}$ ,

$$\Delta p \ll n_{n0}.$$

Detta innebär att vi antar så kallad låg-nivå injektion (se även pnövergången). Därmed förenklas uttrycket för R till

$$R = \beta n_{n0} p_{n0} + \beta n_{n0} \cdot \Delta p = G_{th} + \beta n_{n0} \cdot \Delta p$$

 $och\ nettore kombination stakten\ blir$ 

$$R - G_{th} = \beta n_{n0} \cdot \Delta p.$$

Vi har då

$$0 = \frac{dp_n}{dt} = G_L - \beta n_{n0} \cdot (p_n - p_{n0})$$
$$= G_L - \frac{p_n - p_{n0}}{\tau_h}$$

där vi definierat en parameter, som har dimensionen tid,

$$\tau_h = \frac{1}{\beta n_{n0}},$$

och som är konstant vid låg-nivå injektion,  $n_n \cong n_{n0}$ .

I det stationära tillståndet gäller då

$$p_n - p_{n0} = \tau_h \cdot G_L.$$

Detta hålöverskott leder till ökad konduktivitet, som eftersom överskottet beror på belysningen även kallas *fotokonduktvitet*.

Vad händer nu om vi släcker lampan vid T = 0? Ekvationen som beskriver ändringen av hålkoncentrationen med tiden lyder nu

$$\frac{dp_n}{dt} = -\frac{p_n - p_{n0}}{\tau_h}$$

och vi får lösningen

$$p_n(t) - p_{n0} = (p_n(0) - p_{n0}) \cdot e^{-t/\tau_h}$$
  
=  $G_L \tau_h \cdot e^{-t/\tau_h}$ 

Exempel:

 $p_{n0} \ll \Delta p \ll n_{n0}.$ 

Antagande 2.

$$\begin{split} n_{n0} &= N_D = 10^{22} \text{ m}^{-3} \\ n_i &= 10^{16} \text{ m}^{-3} \\ p_{n0} &= \frac{n_i^2}{n_{n0}} = 10^{10} \text{ m}^{-3} \\ \text{Därmed måste gälla vid lågnivå} \\ \text{injektion:} \\ 10^{10} \text{ m}^{-3} \ll p_n \ll 10^{22} \text{ m}^{-3} \\ \text{vilket normalt inte brukar} \\ \text{innebära någon större} \\ \text{begränsning!} \end{split}$$

Vi förutsätter låg-nivå injektion,

 $\begin{aligned} \frac{dy}{dt} &= -\frac{y}{t} \\ \text{har ju som du säkert kommer} \\ \text{ihåg lösningen} \\ y(t) &= y(0) \cdot e^{-\frac{t}{\tau}} \end{aligned}$ 



Figur D.2: En halvledare återgår till termisk jämvikt när vi slutar störa den.

Överskottet dör ut exponentiellt, halvledaren återgår till termisk jämvikt med en relaxationstid  $\tau_h$ . (Vi kallar  $\tau_h$  även minoritetsladdningsbärarnas *livstid*, dvs  $\tau_h$  är ett mått på den tid det tar för minoritetsladdningsbärarna att försvinna när ljuset har släckts, se nedan.)

Vi har här underförstått diskuterat en halvledare där generationsoch rekombinationsprocesser sker genom direkta övergångar mellan lednings- och valensband. I material med indirekt bandgap som t ex kisel dominerar övergångar via djupa störnivåer. Analysen blir mera komplicerad, men även i detta fall kan vi på liknande sätt definiera en relaxationstid som blir beroende av t ex störställenas koncentration. Därmed får vi också möjlighet att styra hur snabbt halvledaren återgår till jämvikt efter en störning genom att medvetet introducera djupa störnivåer.

Vi kan slutligen också ge relaxationstiden  $\tau_h$  ett annat namn. Vi tänker oss som ovan att vi har låg-nivå injektion i en n-typ halvledare, dvs ett överskott av minoritetsladdningsbärare  $p_n - p_{n0}$ . Om vi tillskriver överskottsladdningsbärarna en livstid  $t_p$  så måste överskottet försvinna inom tiden  $t_p$  på grund av nettorekombination  $R - G_{th}$ , dvs vi kan skriva

$$R - G_{th} = \frac{p_n - p_{n0}}{t_p}$$

Jämför vi detta uttryck med tidigare uttryck ovan,

$$R - G_{th} = \beta n_{n0} \cdot \Delta p = \frac{p_n - p_{n0}}{\tau_p}$$

kan vi därmed identifiera  $\tau_h$  med överskottsminoritetsladdningsbärarnas livstid (excess-minority-carrier lifetime). Denna livstid blir en viktig parameter när det t ex gäller att bestämma hur ett lokalt överskott av minoritetsladdningsbärare diffunderar och samtidigt rekombinerar i t ex pn-övergången eller den bipolära transistorn.



# Appendix E

# Konstanter

c	= 299792458  m/s	${f L} {f jushastigheten}$
ħ	$= 1.054573 \cdot 10^{-34} \text{ Js}$	Plancks konstant
	$= 0.6582122 \cdot 10^{-15} \text{ eVs}$	
h	$= 0.662608 \cdot 10^{-33} \text{ Js}$	
	$= 4.135669 \cdot 10^{-15} \text{ eVs}$	
e	$= 1.602177 \cdot 10^{-19} \text{ C}$	${ m Elementarladdningen}$
$a_0$	$= 0.529177 \cdot 10^{-10} \text{ m}$	Bohrradien
$m_e$	$= 0.910939 \cdot 10^{-30} \text{ kg}$	Elektronmassan
$m_p$	$= 1.672623 \cdot 10^{-27} \text{ kg}$	Protonmassan
	$= 1836.15270m_e$	
$m_n$	$= 1.674929 \cdot 10^{-27} \text{ kg}$	Neutronmassan
k	$= 1.380658 \cdot 10^{-23} ~{\rm J/K}$	Boltzmanns konstant
	$= 0.861739 \cdot 10^{-4} \ {\rm eV/K}$	
$N_A$	$= 0.6022137 \cdot 10^{27} \ \rm kmol^{-1}$	Avogadros konstant
Ry	= 13.6057  eV	Rydbergs konstant Ry = $\frac{h^2}{2ma_0^2}$
	$= 109737.32 \text{ cm}^{-1}$	0
R	$= 0.83145 \cdot 10^4 \ {\rm J}/({\rm kmol} \ {\rm K})$	allmänna gaskonstanten
$\alpha$	$=\frac{e^2}{4\pi\varepsilon_0 hc}=\frac{1}{137,036}$	${\it finstrukturkonstanten}$
$\varepsilon_0$	$= 0.885419 \cdot 10^{-11} \text{ As/Vm}$	dielektriskakonstanen för vakuum
		numera: permittiviteten i vakuum
$\mu_0$	$= 1.256637 \cdot 10^{-6} \ \mathrm{Vs/Am}$	permeabiliteten för vakuum
	$=4\pi\cdot 10^{-7}~{ m Vs/Am}$	
$\mu_B$	$= 9.27402 \cdot 10^{-24} \text{ Am}^2$	Bohr magnetonen $\mu_B = \frac{eh}{2m}$