

# Pokroky matematiky, fyziky a astronomie

---

Emanuel Klier

Čtyřicet let tranzistoru

*Pokroky matematiky, fyziky a astronomie*, Vol. 35 (1990), No. 4, 192--212

Persistent URL: <http://dml.cz/dmlcz/137821>

## Terms of use:

© Jednota českých matematiků a fyziků, 1990

Institute of Mathematics of the Academy of Sciences of the Czech Republic provides access to digitized documents strictly for personal use. Each copy of any part of this document must contain these *Terms of use*.



This paper has been digitized, optimized for electronic delivery and stamped with digital signature within the project *DML-CZ: The Czech Digital Mathematics Library* <http://project.dml.cz>

# Čtyřicet let tranzistoru

*Emanuel Klier, Praha*

## 1. Hrotový a slitinový tranzistor

Myšlenka náhrady vakuových a plynových elektronek polovodičovými součástkami má dlouhou historii, kterou zde nebudeme podrobně probírat. Připomeňme si však, že ve třicátých letech se již běžně užívalo usměrňovačů na bázi kysličníku měďného („kuprox“) a selenu a v r. 1938 byla vypracována jejich teorie nezávisle N. Mottem a B. Davydovem a později podstatně zdokonalena W. Schottkym. Kontakt kov-polovodič se svými usměrňovacími vlastnostmi se proto nazývá Schottkyho kontakt. Podstatný pokrok v aplikacích polovodičů byl umožněn přechodem ke germaniu, pokrokem v technologii čistých látek a v pěstování monokrystalů za druhé světové války. Čistota a dokonalost struktury polovodičů umožňují totiž reprodukovatelně ovládat elektrické vlastnosti a porovnávat experimentální výsledky s teorií založenou na idealizaci dokonalých materiálů, popřípadě realizovat jevy předpověděné teorií.

Počátkem čtyřicátých let byl vývoj germaniových hrotových diod pro slaboproudé účely na komerčně zralé úrovni díky zdokonalené čistotě výchozího materiálu a také byly k dispozici dobré monokrystaly germania. Tento vývoj byl podnícen selháním vakuových elektronek ve vysokofrekvenční radarové technice. Zbýval krůček k objevu hrotového tranzistoru, který učinili r. 1947 a uveřejnili v r. 1948 J. Bardeen a W. H. Brattain. Přišli však k tomuto významnému objevu vlastně oklikou přes myšlenku toho, co dnes označujeme termínem „polní tranzistor“ neboli FET (Field Effect Transistor).

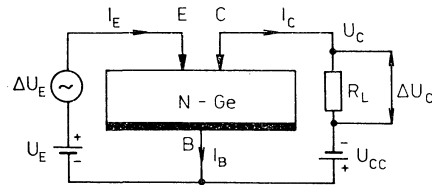
Snaha po náhradě aktivních elektronek typu triody polovodičovými prvky se datuje do r. 1930, kdy J. Lilienfeld v USA a později r. 1935 O. Heil v Anglii ohlásili patenty na zesilovací prvek, jehož izolovaná řídicí elektroda má ovládat proud v polovodiči pomocí příčného pole, tedy v uspořádání, které bychom dnes označili MIS FET (Metal-Insulator-Semiconductor Field-Effect-Transistor). Navrženými polovodiči tehdy byly CuS v prvním a Cu<sub>2</sub>O aj. v druhém případě. K realizaci však nedošlo, zřejmě pro slabé ovlivnění vodivosti polem v nedokonalých a tlustých polykrystalických vzorcích.

V Bellových laboratořích v USA ve čtyřicátých letech sledovala však podobnou myšlenku skupina badatelů v čele s M. Shockleyem a zmíněnými dvěma badateli, Bardeenem a Brattainem. Poslední dva z nich hlavně zkoumali příčinu toho, že příčné pole izolované elektrody neproniká ani do poměrně dokonalého krystalu germania. Dospěli k závěru, že na povrchu existují nepohyblivé náboje, které odstiňují pole, a proto zkoumali povrchové vlastnosti pomocí rozložení potenciálu mezi hrotovými sondami.

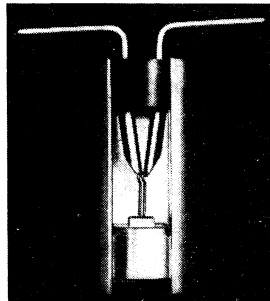
---

Prof. RNDr. EMANUEL KLIER (1918) působí na katedře fyziky polovodičů MFF UK, Ke Karlovu 5, 121 16 Praha 2.

Při tom zpozorovali, že při velmi malé vzdálenosti sond (50–250  $\mu\text{m}$ ), které fungovaly jako hrotové usměrňovače polarizované v opačných směrech, závěrný proud jednoho z nich je silně ovlivněn přímým proudem druhého z nich. Dokonce se někdy pozorovalo, že změna závěrného proudu kontaktu, který nazvali *kolektor*, byla větší než změna proudového proudu kontaktu, který nazvali *emitor*. Dostali tak aktivní prvek schopný zesílení, ať už jen výkonového, nebo dokonce i proudového. Tento objev uveřejnili v krátkém sdělení v časopise *Physical Review* v r. 1948 [1]. Nový prvek jimi nazvaný *tranzistor* (TRANSfer-ResISTOR) byl na světě. Jeho schematický princip a fotografie s otevřeným pouzdem jsou na obr. 1 a 2.



Obr. 1. Schéma hrotového tranzistoru v normálním zapojení se společnouází.



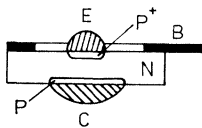
Obr. 2. Fotografie hrotového tranzistoru s otevřeným pouzdem (1948).

Výklad funkce hrotového tranzistoru podali Bardeen a Brattain [2]. Hlavním rysem je účast obou druhů nosičů proudu, tj. elektronů i děr. Na obr. 1 se předpokládá, že základní destička Ge – báze – je typu N, takže v jejím objemu převládá za normálních okolností koncentrace elektronů o několik řádů nad koncentrací děr. Na povrchu však existují pevné záporné náboje, které odpuzují elektrony zpět do objemu a přitahují volné menšinové nosiče, díry, do prostoru těsně pod povrchem. Proto se proud kladně pólovaného emitoru skládá z velké části z děr proudících do objemu a ke kolektoru a z menší části z elektronů proudících z objemu k emitoru. Jestliže jsou kontakty dostatečně blízko, dostane se většina děr až ke hrotu kolektoru. Poněvadž prostor pod záporně pólovaným kolektorem je zbaven velké části elektronů, vzniká pod kolektorem *vrstva ochuzená* o elektrony. Ta tvoří bariéru pro tok elektronů z kolektoru, ale nijak nebrání toku děr do kolektoru. Za existence emitorového proudu je tedy proud v kolektoru větší než bez ní. Je to též *vrstva prostorového náboje* ionizovaných donorů. Tento obraz

by však vysvětloval jen situaci, kdy proud kolektorem je nejvýše roven proudu emitoru, tj. když proudové zesílení  $\alpha = I_C/I_E \leq 1$ . Ve skutečnosti se však pozorovalo, že *činitel proudového zesílení*  $\alpha$  byl často větší než 1 a dosahoval hodnoty i větší než 3. To vysvětlili autoři změnou hustoty kladného prostorového náboje pod kolektorem díky přitékajícím díram. Tato změna snižuje bariéru pro elektrony, a tím vzrůstá proud kolektoru o elektronovou složku. Ve skutečnosti jsou děje v hrotovém tranzistoru složitější.

Poněvadž hrotové tranzistory brzy ztratily význam pro svou nestabilitu, nebudeme se zabývat podrobněji jejich vlastnostmi.

Další vývoj tranzistoru pokračoval rychle. Pokus, který provedl Shive (1949) s velmi tenkou bází a s hrotovými kontakty na protilehlých stěnách, dokázal, že jde o proud nosičů proudu z emitoru samého do objemu báze a dále do kolektoru, takže povrchová vrstva nehraje primární roli [3]. To vedlo další badatele fy Bell pod vedením W. Shockleye k vytvoření teorie trojvrstvého tranzistoru, dnes nazývaného *bipolární* vzhledem k účasti obou typů nosičů, elektronů i děr. V r. 1950 popsal R. N. Hall slitinovou technologii bipolárních tranzistorů, která se dlouho osvědčovala při zavádění tranzistorů do slaboproudé praxe.



Obr. 3. Slitinový tranzistor; schematická struktura PNP.

Schéma slitinového tranzistoru je na obr. 3. Při výrobě se na stěnu monokrystalické germaniové destičky typu N položí kousek india a vše se zahřívá v peci v inertní atmosféře při takové teplotě, kdy se indium roztaví a germanium se v něm částečně rozpustí. Při chladnutí tavenina utuhne jako monokrytal tuhého roztoku In v Ge a vytvoří silně dopovanou vrstvu typu P mezi bází a ztuhlou indiovou kapkou. Indium totiž jako trojmocný prvek tvoří v germaniu akceptory, jejichž koncentrace  $N_A^E$  převyší koncentraci donorů v bazi  $N_D$  ( $N_A^E \gg N_D$ ). Tím vznikne přechod  $P^+N$  ( $P^+$  znamená silnou dotaci akceptory), který slouží jako emitorový. Podobně se vyrobí i kolektorový přechod, avšak vhodným režimem se docílí nižší dotace akceptory, pokud možno tak, aby výsledná koncentrace děr v kolektoru ( $p = N_A^C - N_D$ ) byla menší než koncentrace elektronů v bazi  $n = N_D$ . Důležité pro všechny technologie tranzistoru je, že všechny jeho části na sebe navazují jako koherentní části monokrystalu, tj. tak, že nepřerušeně pokračuje pravidelná struktura monokrystalu celým trojvrstvím. Jinak by totiž na rozhraní vznikaly nežádoucí bariéry a nepohyblivé náboje, které by funkci tranzistoru znemožnily. Skutečná geometrie slitinového tranzistoru je složitá. Pro teoretické úvahy se nahrazuje jednorozměrným modelem planparalelních vrstev s nekonečnými bočními rozměry a počítá se s proudovými hustotami. Tomuto modelu velmi dobře vyhovují tranzistory pokročilejších struktur, o nichž se dále zmíníme. Teorii plošného tranzistoru založenou na analýze dějů na přechodu P-N podal v r. 1949 W. Shockley [4]. Její hlavní myšlenky v názorné formě podáme v následujících dvou částech. Podrobnější teorii tranzistorů najde čtenář např. v knize H. Franka a V. Šnejdara [5].

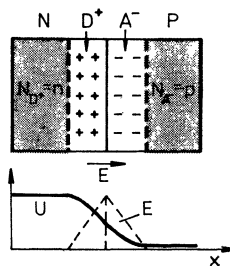
## 2. Děje na přechodu P-N

Jestliže chceme pochopit tranzistorový jev, je nutno porozumět především dějům na přechodu P-N. Pro jednoduchost předpokládejme, že na straně P rovinného přechodu je koncentrace akceptorů  $N_A$  ( $N_D = 0$ ) a na druhé straně koncentrace donorů  $N_D$  ( $N_A = 0$ ). Předpokládá se úplná ionizace obou druhů příměsí, takže v objemu polovodičů P a N máme koncentrace volných nosičů  $p_p = N_A$ ,  $n_n = N_D$ . Kromě většinových nosičů o koncentracích  $p_p$  a  $n_n$  existují v obou typech polovodičů menšinové nosiče, tj. elektrony v typu P a díry v typu N, pro jejichž koncentrace  $n_p$  a  $p_n$  bylo při termické rovnováze (tj. bez předpětí na přechodu) odvozeno

$$(1) \quad n_p \cdot p_p = n_n \cdot p_n = n_i^2,$$

kde  $n_i$  je tzv. intrinzičná koncentrace. Její velikost je nezávislá na koncentracích příměsí a lze ji vypočítat jako funkci teploty, známe-li základní parametry elektronové struktury polovodiče.

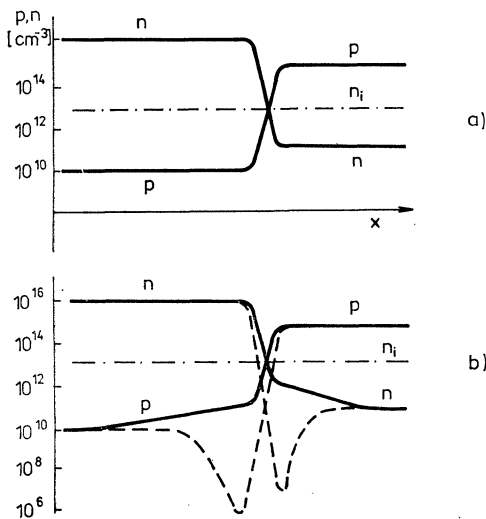
Obr. 4. Prostorové nepohyblivé náboje donorů  $D^+$  a akceptorů;  $A^-$ , elektrické pole a potenciál na rozhraní N-P. Přechodová vrstva prakticky bez volných nosičů.



Přechod je strmý, tj. v ideálním případě se na rozhraní skokem mění koncentrace dopantů z  $N_A$  (při  $N_D = 0$ ) na  $N_D$  (při  $N_A = 0$ ). Volné nosiče ovšem takovou skokovou změnu nemohou kopírovat. Protože jde o volné částice, snaží se difundovat z místa vyšší koncentrace do míst s nižší koncentrací. To znamená, že díry difundují rozhraním do polovodiče N a elektrony do polovodiče P. Zanechávají za sebou nekompenzované náboje ionizovaných příměsí, tj. záporně nabitě akceptory  $A^-$  a kladně nabitě donory  $D^+$  (obr. 4); po obou stranách rozhraní vznikají *nepohyblivé prostorové náboje*, v typu P záporné, v typu N kladné. Tak vzniká *vnitřní elektrické pole*  $E$  směřující z N do P, které způsobuje polní proud. Tento polní proud je realizován pohybem elektronů z typu P do N a děr z typu N do P, tedy proti proudu difúznímu. Ustaví se rovnováha, v níž *proud polní* je právě roven *proudu difúznímu* a této rovnováze odpovídají: 1. rovnovážné rozdělení koncentrací nosičů v přechodu; 2. průběh intenzity vnitřního pole  $E$ ; 3. rovnovážný rozdíl potenciálu  $U_D$  mezi částmi P a N, kterému říkáme *difúzní napětí*. Toto napětí je bariérou pro pohyb elektronů z části N do P a pro pohyb děr z části P do N. Přes tuto bariéru se nosiče dostávají difúzí, což je tepelně aktivovaný děj: se vzrůstající teplotou roste počet nosičů, které mají energii dostatečnou k překonání potenciální bariéry. Tak např. v germaniu, kde je přibližně za pokojové teploty  $n_i = 10^{13} \text{ cm}^{-3}$  a při koncentracích majoritních nositelů v částech N a P  $n_n = N_D =$

$= 10^{16} \text{ cm}^{-3}$  a  $p_p = N_A = 10^{15} \text{ cm}^{-3}$  by byly koncentrace minoritních nositelů podle (1)  $n_p = 10^{10} \text{ cm}^{-3}$  a  $p_n = 10^{11} \text{ cm}^{-3}$ . Při tom by se vytvořila bariéra vysoká přibližně 0,3 V. Oblast prostorového náboje by byla široká asi 0,5  $\mu\text{m}$ . Tato vrstva se chová jako kondenzátor, neboť je bez volných nábojů. Tato situace je znázorněna na obr. 5a.

Jestliže připojíme vnější napětí  $U$  v *propustném směru*, tj. tak, že část P je připojena ke kladnému a část N k zápornému pólu zdroje, potenciálová bariéra se sníží na  $U_D - U$  a poruší se rovnováha *ve prospěch difúzního proudu*. Z obou stran pronikají většinové nosiče do přechodu a dostávají se hluboko do protilehlých částí, tj. elektrony do části P a díry do části N a stávají se tak menšinovými. To je naznačeno pomocí koncentrací na obr. 5b. Tomuto jevu říkáme *injekce menšinových nosičů*. Je třeba zdůraznit, že jde převážně o difúzní proces, nikoli však o nějaké urychlení nábojů v elektrickém poli nebo o prosté snížení odporu přechodové vrstvy. Ukazuje se, že přímý proud způsobený injekcí velmi rychle narůstá s napětím  $U$  a dá se přibližně vyjádřit exponenciální funkcí  $U$ . Tento obraz platí pro nepřilíši vysokou injekci.



Obr. 5. Rozložení koncentrací děr ( $p$ ) a elektronů ( $n$ ) na přechodu PN v logaritmickém měřítku:

- a) v rovnovážném stavu,
- b) s předpětím v propustném směru  
 (plná čára),  
 s předpětím v závěrném směru  
 (čárkovaná čára).

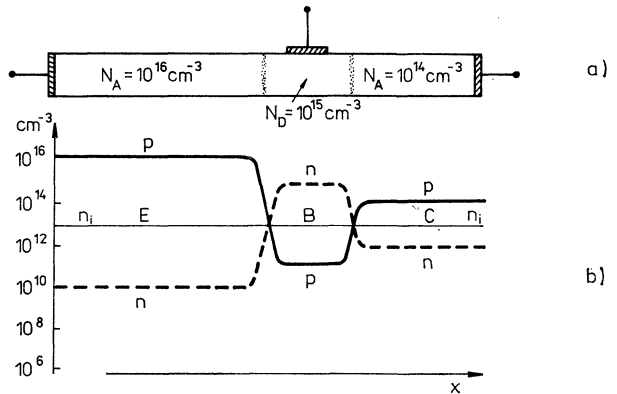
Jestliže je přechod pólován *v závěrném směru*, (tj. N-vrstva kladná a P-vrstva záporná) s napětím  $U$ , zvýší se potenciální bariéra na  $U_D + U$  (obr. 5b) a *difúze majoritních nosičů přes bariéru se potlačí*. Proud polní, který tvoří díry proudící z části N a elektrony z části P, je omezen, neboť je jich na hranici přechodné vrstvy velmi málo k dispozici. Nadto se proti rovnovážnému stavu koncentrace všech nosičů v přechodové vrstvě sníží, oblast prostorového náboje se rozšíří, jak naznačuje obr. 5b (čárkovaná křivka). Teorie ukazuje, že proud minoritních nosičů z obou stran je omezen difúzním proudem ve směru koncentrací menšinových nosičů z objemu směrem k přechodové vrstvě. Tato difúze probíhá při neobyčejně nízkých úrovních koncentrací (značně menších než  $n_p$  a  $p_n$ ), takže závěrný proud je o mnoho řádů menší než propustný a stoupajícím závěrným napětím se nasýtí při napětích blízkých jednomu voltu.

Je třeba si uvědomit, že přiložením napětí k přechodu je porušena rovnováha, takže při propustném napětí je všude v přechodu  $p \cdot n > n_i^2$  a při závěrném napětí je  $p \cdot n < n_i^2$ . Porušení rovnováhy vede k dějům směřujícím k obnovení rovnováhy: nastává *rekombinace nosičů* při  $n \cdot p > n_i^2$  a *generace nosičů* při  $n \cdot p < n_i^2$ . Zde nás zvláště zajímá stav injekce, kdy převládá rekombinace menšinových nosičů. Tento děj je charakterizován *dobou života a pohyblivostí menšinových nosičů*, které jsou tím větší, čím je krystal čistší a dokonalejší. Doba života a pohyblivost nosičů určují vzdálenost, do které se mohou průměrně dostat při injekci, dříve než zmizí rekombinací. Pro dobrou injekční účinnost je žádoucí co nejdelší doba života. Při nesymetrickém přechodu (např.  $N_D \gg N_A$ ) převládá ten injekční proud, který přechází ze silněji dotovaného polovodiče. V našem případě (obr. 5b) převládá injekce elektronů do polovodiče P.

### 3. Princip tranzistorového jevu

Budeme předpokládat, že náš tranzistor má strukturu P N P, kde levá část P tvoří emitor (E), část N bázi (B) a pravá část P kolektor (C) (obr. 6a). Pro koncentrace příměsí ve vrstvách E, B, C platí  $N_A^E > N_D^B > N_A^C$ . Analogicky by vypadala i struktura N P N.

Obr. 6. Model plošného tranzistoru (podle Spenkeho [7]):  
 a) základní struktura,  
 b) rozložení koncentrací elektronů a děr bez předpětí,

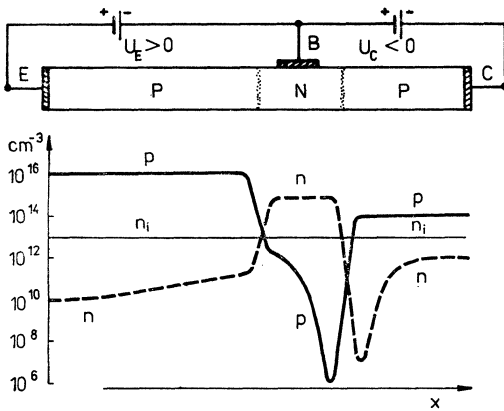


Při přechodu k této komplementární struktuře je třeba ve výkladu zaměnit všechny pojmy za komplementární:  $N_A \leftrightarrow N_D$ ,  $n \leftrightarrow p$ , elektron  $\leftrightarrow$  díra,  $+U \leftrightarrow -U$ .

Uvažujeme tranzistor v normálním režimu s uzemněnou („společnou“) bází, takže napětí jednotlivých částí jsou  $U_B = 0$ ,  $U_E > 0$ ,  $U_C < 0$ , tj. přechod emitorový je ve stavu propustném, kolektorový ve stavu závěrném. Pravidelně je  $|U_E| < |U_C|$ . Předpokládáme, že všechny příměsi jsou ionizovány, takže rovnovážné koncentrace většinových nosičů jsou rovny koncentracím příměsí v příslušných vrstvách. Dále budeme zanedbávat pole v neutrálních částech vrstev E a C. Průřez přechodů předpokládáme jednotkový.

Pokud je tranzistor bez předpětí, tj. v rovnováze podle rovnice (1), je průběh koncentrací podle obr. 6b.

Po vložení napětí  $U_E$  a  $U_C$  (obr. 6c) se podle teorie koncentrace nosičů změní tak, jak je znázorněno na obr. 6d. Všimněme si nejdříve emitorového přechodu (obr. 6d). Je pólován v propustném směru, takže nastává injekce děr do báze a elektronů do emitoru. Tato injekce tvoří složky emitorového proudu  $I_E = I_{E_p} + I_{E_n}$ . Na rozdíl od § 2 zde máme silně nesymetrický přechod, takže difúze děr do báze se odehrává na úrovních koncentrací o několik řádů vyšších než difúze elektronů, takže je  $I_{E_n} \ll I_{E_p}$ . Jde tedy prakticky o jednostrannou injekci děr do báze.



c) Obr. 6. c) zapojení tranzistoru PNP v normálním režimu se společnou bází,  
d) rozložení koncentrací elektronů a děr při předpětí podle obr. 6c; situace odpovídá přibližně germaniu při velmi slabé injekci a široké bázi.

V bázi vidíme monotónní pokles koncentrace děr až ke kolektorovému přechodu (obr. 6d). Injektované díry se pohybují bází převážně difúzním mechanismem. Jakmile dorazí do kolektorové přechodové vrstvy, jsou strženy silným vnitřním polem a rychle proletí až do objemu neutrální kolektorové vrstvy. Je-li jejich životní doba  $\tau_B$  delší než doba transportu báze  $T_B$ ,  $\tau_B > T_B$ , pak se jich v bázi jen malá část ztratí rekombinací a téměř stejný jejich proud, který pramení v emitoru, ústí do kolektoru  $I_{C_p} \doteq I_{E_p}$ .

Podle obr. 6d by se na první pohled zdálo, že vzhledem k silnému poklesu koncentrace děr v bázi, se jich jen málo dostane do kolektoru. Tomu tak není. Teorie ukazuje, že tam, kde koncentrace děr klesá, jejich difúzní rychlost  $v_D$  stoupá, takže proudová hustota ( $e \cdot p \cdot v_D$ ) zůstává téměř konstantní od emitoru až do kolektoru.

Přenos děrového proudu emitoru do kolektoru tvoří podstatu tranzistorového jevu, tedy kontrolu kolektorového proudu proudem emitorovým.

Kromě děrové složky  $I_{C_p}$  obsahuje kolektorový proud též elektronovou složku  $I_{C_n}$ , která je dána nasyceným závěrným proudem kolektorového přechodu. Tato složka prakticky nezávisí na kolektorovém napětí ani na emitorovém proudu a bývá značně menší než  $I_E$ . Je tedy konečně celkový proud kolektoru  $I_C \doteq I_{C_p}$ . Při účinné injekci ( $I_E \doteq I_{E_p}$ ) a dobré účinnosti přenosu ( $\tau_B \gg T_B$ ) je konečně i  $I_C \doteq I_E$  v dosti širokých mezích  $I_E$  a odtud plyne pro zesilovací proudový činitel

$$\alpha = \frac{I_C}{I_E} \doteq 1.$$

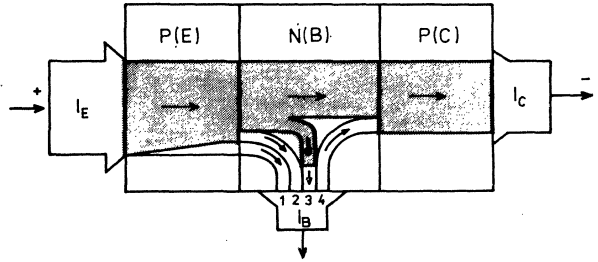
Praktické hodnoty  $\alpha$  jsou mezi  $0,98 \div 0,995$ .



Proud bázi  $I_B$  je v podstatě elektronový a má tyto složky (obr. 7): injekční proud elektronový z báze do emitoru, rekombinační proud způsobený rekombinací jednak v emitorovém přechodu, jednak v objemu báze a konečně závěrný proud kolektorovým přechodem N-P,  $I_C$ . Rekombinační proudy představují ztráty injektovaných děr z emitoru. Celkově lze proud  $I_B$  vyjádřit jako

$$I_B = I_E - I_C = (1 - \alpha) I_E.$$

Obr. 7. Složky proudů v tranzistoru; bílé plochy: elektronový, šedé plochy: děrový proud. Složky bázevého proudu: 1 injekce  $B \rightarrow E$ ; 2 rekombinace v přechodu E; 3 rekombinace v bázi; 4 závěrný proud kolektoru. Šipky značí konvenční směr proudu.

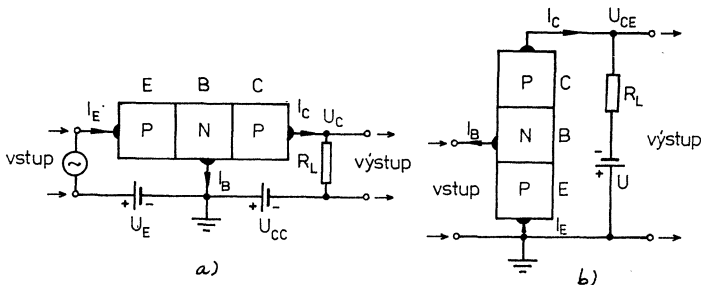


$I_B$  je tím menší, čím delší je doba života nosičů v bázi. V zapojení „se společným emitorem“ podle obr. 8 je vstupním proudem proud bázi  $I_B$ , emitor je na potenciálu  $U_E = 0$ , výstupním napětím je pak napětí  $U_{CE}$  mezi kolektorem a emitorem. Pro činitel proudového zesílení nyní vychází

$$\beta = \frac{I_C}{I_B} = \frac{I_C}{I_E - I_C} = \frac{\alpha}{1 - \alpha}.$$

Je vidět, že  $\beta$  je běžně  $\gg 1$  a že urychleně stoupá, čím více se  $\alpha$  blíží k jedničce.

Pro rychlost odezvy, popř. pro mezní použitelnou frekvenci v.f. signálu, je rozhodujícím faktorem tranzistoru doba průletu děr bázi  $T_B$ . Ta je určena u tranzistorů s homogenní bázi tloušťkou báze a rychlostí difúze, která je poměrně pomalá.  $T_B$  je možno zkrátit zmenšováním tloušťky báze, které má svoje technologické omezení u hodnot řádu  $1 \mu\text{m}$ . Dále je možno snížit  $T_B$  tím, že se vytvoří struktura se „zabudovaným“ vnitřním polem v bázi, které urychluje injektované nosiče. Lze toho docílit tím, že koncentrace příměsí v bázi není rovnoměrná, nýbrž silně klesá od emitoru ke kolektoru.



Obr. 8. Schematické zapojení plošného tranzistoru PNP. a) se společnou bází, b) se společným emitorem.

## 4. Vývoj některých technologií

### 4.1. Výchozí materiály a příprava krystalů Si

Snaha po zdokonalování vlastností tranzistorů a dalších elektronických prvků podnítila vývoj různých technologických postupů a hledání dalších vhodných materiálů. Protože se velmi brzy ukázalo, že tranzistor přináší do elektroniky revoluci významnější než vakuové elektronky, bylo jeho vývoji a technologii výchozích materiálů věnováno úsilí narůstající geometrickou řadou s časem. Jeho podrobné historické zhodnocení by ukázalo příkladnou interakci technologií, experimentátorů a teoretiků z oborů fyziky, elektroniky, chemie i matematiky.

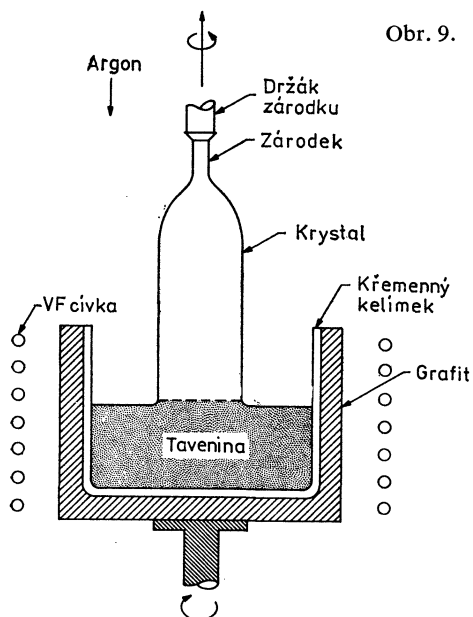
Pokud se týká materiálů, převládalo v padesátých letech *germanium*. Kolem r. 1960 však nastoupil *křemík* a rychle zatlačil germanium do pozadí. Křemík má proti germaniu některé nevýhody, například nižší *pohyblivost nosičů* a větší prahové napětí pro nasazení propustného proudu ( $\sim 0,7$  V). Zato však má řadu pozoruhodných vlastností, které umožňují technologické postupy a konstrukce nových struktur nedosažitelných na jiných materiálech. Jeho elektronová struktura vede k nižším hodnotám nasyceného závěrného proudu, k vyšším hodnotám přípustného závěrného napětí a k možnosti práce při vyšších teplotách než u germania.

Daleko nejdůležitější jeho vlastností je, že se jednoduchým procesem oxidace při vysokých teplotách ( $800-1200$  °C) vytváří na povrchu chemicky odolný *oxid křemičitý* s vysoce kvalitními dielektrickými vlastnostmi a s malou hustotou lokalizovaných elektronových stavů na rozhraní Si – SiO<sub>2</sub>. Tyto vlastnosti umožnily jednak rozvinutí struktur kov-oxid-polovodič (čili MOS – Metal-Oxide-Semiconductor) pro tranzistory a jiné součástky založené na polním efektu, jednak zavedení maskovací techniky, a tím i výrobu integrovaných obvodů. Další výhodou přirozených oxidových vrstev je možnost jejich užití jako pasivační vrstvy na hotových součástkách. Tím, že zabraňují různým chemickým a adsorpčním procesům na povrchu, zvyšují podstatně časovou stabilitu součástek.

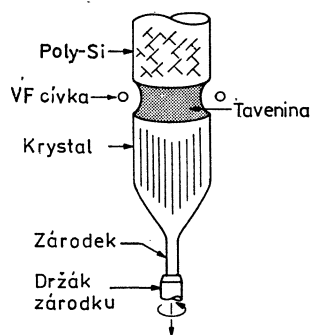
Pozice křemíku v součástkové základně elektroniky se zdá pro nejbližších nejméně 20 let neotřesitelná. Na druhém místě dnes stojí sloučeniny typu A<sup>3</sup>B<sup>5</sup> tj. prvků 3. grupy (Al, Ga, In) s prvky 5. grupy (P, As, Sb), z nich především arzenid galitů (GaAs). Různé kombinace prvků 3. a 5. grupy ve dvojných, trojných a čtverných sloučeninách poskytují nepřehledné možnosti volby požadovaných vlastností a realizaci prvků, jejichž funkce není uskutečnitelná s elementárními polovodiči. Popis technologií, prvků a aplikací značně přesahuje rámec tohoto článku, a proto se obrátíme téměř výhradně ke křemíku.

Nejrozšířenější metodou přípravy čistého křemíku je pyrolytický rozklad plyných sloučenin křemíku, jako je silan (SiH<sub>4</sub>) a jeho chlorové deriváty. Tyto sloučeniny se dají čistit opakovanou destilací. Vznikající křemík je polykrystalický a připravují se z něho monokrystaly převážně *metodou tažení podle Czochralského* (původně pro iontové krystaly 1917). Její princip je na obr. 9. Tavenina v kelímku z křemene je zahřívána indukčním v.f. ohřevem vnějšího grafitového kelímku. Do taveniny se ponoří zárodek a ten se pak vytahuje nad hladinu za stálé protiběžné rotace zárodku

a kelímku. Tavenina se udržuje na teplotě těsně nad bodem tání Si (1420 °C). Rychlosti tažení bývají až několik mm/min a proces probíhá v argonové atmosféře. Dnes se běžně takto připravují krystaly o průměru 10–15 cm, dlouhé kolem 80–150 cm, v dohledu jsou krystaly o průměru 20 cm. Běžně se dosahuje čistoty označované ve zlomcích  $(1 \div 25) \cdot 10^{-9}$  atomů aktivních nečistot (tj. donorů a akceptorů) na jeden atom Si. Hlavními neaktivními nečistotami jsou kyslík a uhlík, které nemají vliv na vodivost i při koncentracích do  $10^{18}/\text{cm}^{-3}$ . Je to nepředstavitelná čistota: na jediný atom aktivní nečistoty připadá 40 miliónů až jedna miliarda atomů křemíku. Dopování příměsí se provádí přidáváním vhodných sloučenin do taveniny nebo tak, že se při tažení nechá probíhat difúze z plynné fáze par ( $\text{B}_2\text{H}_6$ ,  $\text{PH}_3$ ,  $\text{AsH}_3$ ).



Obr. 9. Czochralského tažička krystalů (podle Sze [8]).



Obr. 10. Metoda pěstování monokrystalů pomocí visuté zóny.

Kromě Czochralského metody se užívá též *metody visuté zóny*. Při ní svisle postaveným polykrystalickým ingotem putuje vzhůru indukčně roztavená zóna, která drží pohromadě kapilárními silami (obr. 10). Na začátku ingotu je monokrystalický zárodek, takže zóna krystalizuje ve formě monokrystalu. Výhoda této metody záleží v tom, že tavenina není v kontaktu s žádným kelímkem, takže relativní znečištění kyslíkem a uhlíkem jsou nižší (pod  $10^{-9}$ ).

Metody růstu se označují běžně jako Cz pro Czochralského metodu a FZ (Floating Zone) pro metodu visuté zóny.

#### 4.2. Epitaxní vrstvy

Epitaxní růst znamená souvislé pokračování krystalové mřížky monokrystalické

podložky krystalizací z kapalně nebo plynné fáze. Obvykle se užívá podložky ve tvaru desky s přesně orientovanou vrchní stěnou podél význačné atomové roviny, např. (100), (111) apod. Nejužívanější metoda pro křemík, vhodná pro hromadnou výrobu, je *epitaxe z plynné fáze*, označovaná jako *VPE* (Vapour-Phase-Epitaxy). Je to opět chemický rozklad některého chlór-silanu v proudící vodíkové atmosféře za vývoje křemíku (na podložce) a chlorovodíku při teplotách kolem 1200 °C. Vznikají vrstvy velmi dokonalé struktury, rovnoměrné tloušťky, kterou lze regulovat volbou teploty, složením atmosféry a dobou růstu od tloušťek setin až jednotek  $\mu\text{m}$ . Dopanty (příměsi) se vpravují též ve formě plynů a par do transportního plynu, např. jako páry  $\text{GaCl}_3$ ,  $\text{BBr}_3$  pro typ P nebo jako arsin  $\text{AsH}_3$  či fosfin  $\text{PH}_3$  pro typ N. Vysoká toxicita těchto plynných sloučenin vyžaduje ovšem zvláštní bezpečnostní opatření.

Epitaxe z kapalně fáze (LPE) se hodí pro sloučeniny  $\text{A}^3\text{B}^5$ .

Nejpokročilejší metoda výroby epitaxních vrstev je užití *molekulárních svazků* (MBE, Molecular-Beam-Epitaxy). Při této metodě se v recipientu vyčerpáném na ultravysoké vakuum ( $p < 10^{-8}$  Pa) termicky generované svazky molekul nebo atomů nasměrují na podložku, na niž dochází k růstu monokrystalické vrstvy žádaného složení a žádané koncentrace dopantů. Pomocí počítače se kontroluje složení a rychlost růstu s takovou přesností, že je možno reprodukovatelně pěstovat vrstvy o tloušťce deseti atomových rovin i méně. Tato metoda je velmi nákladná (aparatura stojí řádově až několik miliónů dolarů) a užívá se pro speciální účely v aplikacích sloučenin, u křemíku jen zcela výjimečně.

### 4.3. Difúze a implantace

Difúze a implantace jsou dvě hlavní metody, kterými se zavádějí příměsi do polovodičů.

Difúze do křemíku se v převážné většině provádí z plynné fáze do připravené desky vhodně orientované a oleptané. Nejobvyklejší dopanty jsou pak diboran  $\text{B}_2\text{H}_6$  pro typ P a arsin či fosfin pro typ N. Používaná teplota je 800–1200 °C a jako transportního inertního plynu se užívá čistého dusíku. Důležitý je zde tzv. difúzní profil neboli závislost koncentrace dopantu na hloubce pod povrchem  $x$ . Tento profil se dá teoreticky vypočítat pro různé podmínky procesu. Měněním podmínek lze dosáhnout různých hodnot koncentrací, různých průměrných hloubek proniknutí a různého spádu koncentrací. Profil lze dále upravit žíháním bez dalšího vnášení dopantu.

Jestliže chceme vnést příměs jen na určitá omezená místa, lze to provést *maskováním* tak, že se nejdříve pokryje celá deska křemíku oxidem  $\text{SiO}_2$ , do něho se vyleptají na žádaných místech otvory a pak se provede difúze. Oxidová vrstva je tak tlustá, že jí difúzant nemůže projít, ovšem je nutno pamatovat na to, že difúze probíhá pod otvorem všemi směry, takže příměs se dostane i do jisté vzdálenosti pod okraj otvoru v oxidové masce.

*Implantace iontů* záleží v nastřelování iontů dopantu urychlených na energie 30 až 300 keV do podložky. Základní sestava aparatury obsahuje zdroj iontů (např.  $\text{B}^+$ ,  $\text{As}^+$ ), separátor na principu hmotnostního spektrometru, urychlovač s fokusačním zařízením a vychylovací zařízení, kterým se zaměřuje svazek iontů do žádaného místa na substrátu.

Energie iontů řídí hloubku proniknutí do polovodiče. Vysoce energetické ionty nejen že pronikají mezi atomy mřížce, ale vyrážejí mnoho atomů z jejich pravidelných poloh, takže vznikají prázdná místa (vakance) a atomy v meziuzlových polohách (intersticiály), což je nežádoucí. Je proto nutno implantované desky „vyhojit“ žíháním, tj. obnovit porušený pořádek v mřížce polovodiče. Žádoucí profil koncentrací lze získat volbou energie a celkové dávky implantace, popř. žíháním. Zatímco při difúzi koncentrace dopantu monotónně klesá s hloubkou, má při implantaci koncentrace maximum v určité hloubce pod povrchem. Jemu odpovídá průměrný doběh iontů dané energie a hmotnosti.

#### 4.4. Maskovací technika

Jak jsme se zmínili v předcházejícím odstavci, obvykle je třeba omezit difúzi nebo implantaci na ostře ohraničené ostrůvky a chránit ostatní povrch desky před účinky těchto procesů. Děje se tak pomocí maskování.

Maska je tvořena buď vrstvou oxidu křemičitého, nebo fotorezistem, a to podle toho, je-li následující proces vysokoteplotní (difúze) nebo nízkoteplotní (implantace, leptání).

*Fotorezist* je polymerní fotocitlivá látka, ve které se působením ultrafialového světla nebo svazku elektronů buď zvyšuje, nebo snižuje stupeň zesíťování. V prvním případě osvětlená místa se v rozpouštědle nerozpouštějí, v druhém případě se rozpouštějí snadno. První druh fotorezistu nazýváme negativní, druhý pozitivní. Osvětlená místa v negativním fotorezistu jsou nepropustná, v pozitivním naopak propustná. Vytváření struktur ve fotorezistových vrstvách se nazývá *folitolitografie*.

Při difúzi je prvním procesem oxidace. Na vrstvu oxidu se nanese vrstva fotorezistu a ten se exponuje skrze přiloženou masku s otvory na místech, která mají být odkryta pro následující leptání (*pozitivní fotorezist*) nebo naopak zakryta (*negativní fotorezist*). Po „vyvolání“ fotorezistu, tj. aplikaci rozpouštědla, se na žádaných místech objeví otvory odpovídající tvaru masky nebo jejímu negativu podle druhu fotorezistu.

Následuje leptání v leptadlech obsahujících kyselinu fluorovodíkovou, která rozpouští  $\text{SiO}_2$ . Leptá se tak dlouho, až se objeví povrch křemíku. Fotorezist se odstraní a provádí se difúze při vysoké teplotě. Difundant nemůže proniknout oxidovou vrstvou a difúze probíhá jen pod vyleptanými otvory. Při implantaci stačí i maska fotorezistová, pokud při dané energii ionty neproniknou nerozpuštěnou fotorezistovou vrstvou.

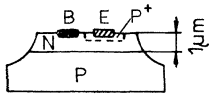
Maskovací technika se též využívá při zhotovování kovových vrstev pro kontakty a spoje metodou vypařování ve vakuu nebo naprašování ve výboji.

#### 4.5. Vývoj tranzistorových struktur

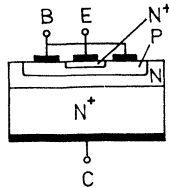
Slitinový tranzistor brzy nevyhovoval požadavkům na aplikace v přijímačích pro krátké a později velmi krátké vlny, neboť jeho technologie neumožňovala snížit tloušťku báze pod asi  $20\ \mu\text{m}$ , a tím byl omezen frekvenční rozsah na frekvence řádu 1 MHz. Skutečným pokrokem se koncem šedesátých let stala technologie *epitaxně difúzní*, která posunula frekvenční mez do gigahertzového pásma. Příklad takové struktury germaniového tranzistoru ukazuje obr. 11. Báze je zde velmi tenká, řádově  $1\ \mu\text{m}$  a provádí se difúzí

do epitaxní kolektorové vrstvy. Masivní substrát (podložka) je silně dopován, aby měl malý sériový odpor. Emitor je proveden difúzí nebo slévací technikou. Kovové kontakty obsahují kovy (Al, Sb) odpovídající svými příměšovými charakteru typům příslušných částí tranzistoru (P, N). Odleptáním boků se získá tvar „meza“, čímž se odstraní porušené oblasti přechodu báze-kolektor na obvodu. („Mesa“ je španělsky stůl, přeneseně stolová hora.)

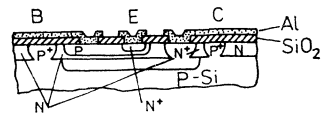
Začátek šedesátých let znamenal prudký vývoj integrovaných obvodů a naprostou nadvládu křemíku. Tento vývoj umožnila především technologie planárních bipolárních tranzistorů. První *planární tranzistory* měly strukturu schematicky naznačenou na obr. 12. Hlavním rysem je, že okraj přechodu B-C je na horním povrchu pokryt pasivační vrstvou  $\text{SiO}_2$ , takže tranzistor získal na stabilitě.



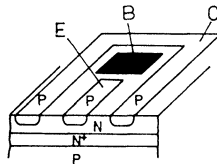
Obr. 11. Tranzistor meza PNP.



Obr. 12. Planární struktura PNP.



Obr. 13. Planární tranzistor NPN s ponořenou vrstvou  $\text{N}^+$ .



Obr. 14. Laterální (boční) tranzistor PNP.

Provedení důsledně planární, kdy všechny přívody k vrstvám E, B, C jsou na horním povrchu, vyžaduje konstrukci vedoucí ke snížení sériového odporu slabě dotované vrstvy C. Takovým řešením je dnes všeobecně používaná *ponořená vrstva* silně dotovaná na stejný typ jako kolektor (obr. 13). Tato vrstva, kterou je nutno vytvořit difúzí před pěstováním epitaxní vrstvy (kolektorové), slouží jako dobře vodivý přívod ke kolektoru C s krátkou dráhou v méně vodivé vrstvě mezi kontaktním ostrůvkem  $\text{N}^+$  a ponořenou vrstvou.

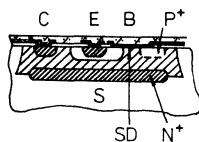
Jinou strukturou užívanou v integrovaných obvodech je *laterální (boční) tranzistor* (obr. 14). V tomto tranzistoru jsou části P, N, P uspořádány vedle sebe v jedné úrovni. Spodní vrstva  $\text{N}^+$  opět slouží ke snížení dráhového odporu, tentokrát však báze N.

Jednou z nevýhod bipolárního tranzistoru klasické struktury je jeho pomalý přechod ze stavu sepnutého do stavu vypnutého. Tyto stavy představují v logických obvodech realizaci logických úrovní 0 a 1 v a počítačové technice se požaduje co největší rychlost přechodu z jednoho stavu do druhého.

Potíž je, že v tranzistoru v sepnutém stavu (vodivém), kdy mezi emitorem a kolekto-

rem je napětí jen asi 0,2 V, je kolektor vůči bázi pólován ve směru propustném ( $\sim 0,7$  V) a kolektorový přechod je zaplaven injektovanými menšinovými nosiči. Tyto nosiče při přechodu do stavu nevodivého mizí jen pomalu, převážně rekombinací a vodivost přechodu jen pomalu klesá.

Aby nedocházelo k této situaci, zapojuje se paralelně ke kolektorovému přechodu PN tzv. Schottkyho dioda (kov-polovodič) se stejným usměrňovacím účinkem jako přechod PN. Při silně vodivém stavu tranzistoru je Schottkyho kontakt vodivější než přechod, převezme téměř všechny proud tranzistorem a nedojde k nahromadění menšinových nosičů v přechodu. Přepnutí do nevodivého stavu je rychlé, protože proud v Schottkyho diodě zprostředkují jen většinové nosiče. Struktura takového „Schottkyho tranzistoru“ je na obr. 15. Schottkyho diodu (SD) tvoří kovová elektroda z takového materiálu, který vytváří s bází ohmický a s kolektorem usměrňující kontakt.



Obr. 15. Tranzistor s Schottkyho diodou B-C.

Tranzistory se významně uplatňují i ve výkonové elektronice, zejména ve spojitosti s automatizací a robotizací ve strojírenství. Další vývoj tranzistorů se ubíral i jiným směrem, a to k vícevrstvovým strukturám (např. PNP), které mohou sloužit ke kontrolovanému usměrnění střídavých proudů nebo k bezztrátové regulaci stejnosměrného proudu; jde o tyristory, triaky atd. Jejich výklad již vybočuje z naší tematiky.

## 5. Polem řízené tranzistory

Tento typ tranzistorů obecně označovaný FET (Field Effect Transistor) pracuje na principu kontroly vodivosti vodivého kanálu mezi dvěma ohmickými kontakty pomocí třetí elektrody, hradla, která vytváří *pole kolmé na směr proudu*. U nás byly zavedeny pro tyto elektrody termíny *emitor-kolektor-hradlo*; zde se emitorem rozumí ta elektroda, ze které pramení nosiče proudu v kanálu. V mezinárodní literatuře se označují elektrody symboly S, D, G podle anglického Source, Drain, Gate (zdroj, výtok, brána). Těchto symbolů se nadále přidržíme.

Proud je realizován jedním druhem nosičů, a to většinových. Proto se tomuto druhu tranzistorů říká *unipolární*.

Ohmické kontakty S a D mohou téměř neomezeně dodávat a odebírat nosiče do kanálu a z kanálu bez zratelného úbytku napětí na rozhraní s kanálem. Nejde tedy o žádnou analogii injekce menšinových nosičů z emitoru, která je základním jevem v bipolárním tranzistoru a naše označení emitor místo zdroje je dost nešťastné. Polem řízené prvky se ukázaly jako velmi výhodné pro konstrukci integrovaných obvodů s vysokou hustotou integrace. Jejich technologie je jednodušší než u bipolárních tranzistorů, obvykle

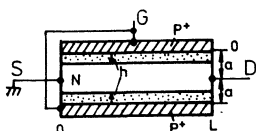
mají menší příkony a v jistých zapojeních nepotřebují ani rezistory ani kondenzátory. Původní jejich nevýhoda, totiž nižší rychlost, se zdokonalením výroby vytratila.

Podle uspořádání řídicí elektrody-hradla se rozeznávají tyto typy:

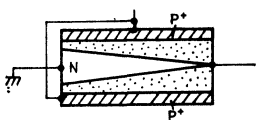
- a) tranzistor s přechodovým hradlem, JFET (Junction FET), kde hradlo tvoří se substrátem přechod PN nebo NP;
- b) tranzistor s Schottkyho hradlem, MESFET, (Metal FET) s hradlem struktury kov-polovodič v závěrném režimu;
- c) tranzistor s izolovaným hradlem, MISFET (Metal-Insulator-Semiconductor FET), kde struktura hradla je nejčastěji kov-oxid-polovodič (MOSFET).

### 5.1. Tranzistor s přechodovým hradlem

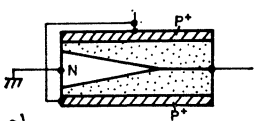
První analýzu realizovatelného tranzistoru s přechodovým hradlem provedl v r. 1950 Shockley [6]. Navržená struktura je tvořena tenkou podlouhlou destičkou, např. typu N, na jejichž obou stěnách jsou jako hradla nadifundovány vrstvy P opatřené kovovými



a)



b)



c)

Obr. 16. Dvojhradlový polní tranzistor s přechodovými hradly (JFET). (Podle Shockleye [6])

při  $U_G = \text{konst.}$

- a) pro  $U_D = 0$ ,
- b) pro  $U_D = U_P$ ,
- c) pro  $U_D > U_P$ .

kontakty navzájem spojenými (obr. 16a). Na koncových ploškách destičky jsou kontakty tvořící emitor a kolektor. Jestliže položíme potenciál emitoru  $U_S = 0$ , pak kolektor bude za provozu na kladném potenciálu  $U_D > 0$  a obě propojená hradla na záporném potenciálu  $U_G < 0$ . Pod hradlovými vrstvami se vytvoří oblasti ochuzené o elektrony a ty jsou nevodivé. Jejich tloušťka je přibližně úměrná odmocnině z napětí  $U_G$ , pokud je  $U_D \ll |U_G|$ . Proud mezi S a D prochází tedy kanálem, jehož výška je tím menší, čím větší je napětí na hradle. Při větších napětích na kolektoru  $U_D$  je nutno respektovat, že rozdíl potenciálů mezi vrstvami P a substrátem N se mění podél kanálu, takže směrem ke kolektoru se kanál zužuje. To vede k tomu, že v závislosti na  $U_D$  (při  $U_G = \text{konst.}$ ) proud  $I_D$  nejdříve roste lineárně, pak pomaleji, až při  $U_D = U_P$  (obr. 16b), kdy kanál se na konci D uzavře, nastává nasycení a proud dále nestoupá.

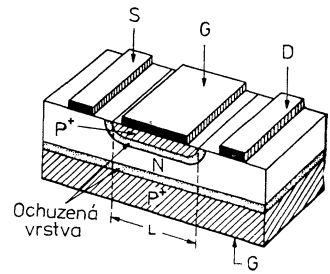
Shockley odvodil rovnici pro proud v závislosti na  $U_G$  a  $U_D$ , jakož i podmínky pro uzavření kanálu a pro hodnotu nasyceného proudu.



Skutečné provedení prvních tranzistorů typu JFET je na obr. 17. Mají-li hradla mít plnou kontrolu nad výškou kanálu při nepříliš vysokých napětích, musí být vrstva N velmi tenká a je obvykle provedena jako epitaxní vrstva. Na takové vrstvičce ovšem nelze realizovat čelní kontakty, a proto je užito planární techniky. Horní a dolní hradla jsou vytvořena difúzí jakožto vrstvy P<sup>+</sup>. Kontakty S, G, D jsou naneseny vypařením vhodných kovů ve vakuu.

Dnešní struktura JFET používá plně planární technologie. Jediné hradlo působí jen jednostranně na vodivý kanál v epitaxní vrstvě. Tato vrstva je vypěstována na podložce ze semi-izolačního polovodiče stejného nebo i jiného složení jako e-vrstva. Vysoký odpor podložky ( $\rho$  až  $10^7 \Omega \text{ cm}$ ) ohraničuje zespodu vodivý kanál.

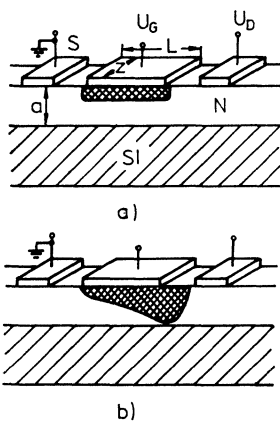
Kontakty emitoru a kolektoru jsou provedeny jako ostrůvky silně vodivého materiálu téhož typu, jako je vodivý kanál. Takové oblasti vytvářejí dobrý kontakt jak k polovodiči, tak i ke kovovým přívodům.



Obr. 17. Praktické provedení struktury JFET (podle Sze [8]).

## 5.2. Tranzistor s Schottkyho hradlem [MESFET]

Princip tohoto tranzistoru i jeho teorie jsou velmi podobné předcházejícímu JFET. Hradlo je zde provedeno jako usměrňující kontakt mezi kovem a polovodičem (obr. 18).



Obr. 18. Struktura MESFET

- a)  $U_D < U_P$
- b)  $U_D = U_P$ .

První podmínka takového kontaktu je, aby výstupní práce z polovodiče typu N pro elektrony byla menší než výstupní práce z kovu. Pak přecházejí snadněji elektrony z polovodiče do kovu a v polovodiči se vytvoří oblast ochuzená o elektrony, tedy s kladným

nepohyblivým prostorovým nábojem. Tento prostorový náboj tvoří donory zbavené elektronů. Jestliže se přiloží na kov záporné napětí, elektrony jsou odpuzovány do vzdálenějších částí polovodiče, ochuzená oblast se rozšiřuje. Situace je zcela analogická jako na přechodu  $P^+N$ , kde část  $P^+$  zastupuje kov. Rozšiřování ochuzené oblasti do vodivé vrstvy reguluje průřez vodivého kanálu, a tím ovlivňuje celkovou vodivost.

Struktury MESFET jsou významné u polovodičů, u nichž lze obtížně vyrábět přechody PN nebo struktury MIS (viz další část 5.3). To je např. případ *arzenidu galitého*, který je jinak vhodný pro mikrovlonnou techniku. Struktury MESFET na GaAs jsou dnes nejběžnějšími aktivními prvky pro integrované obvody v oblasti mikrovl.

### 5.3. Tranzistory s izolovaným hradlem

Princip tohoto druhu tranzistorů je stejně starý jako bipolární tranzistor. Shockley jej původně (1948) formuloval jako „modulace vodivosti povrchovým nábojem“.

Mezinárodní označení tohoto typu tranzistoru je MISFET. V r. 1960 se objevil první MISFET vyrobený pomocí přirozené oxidace povrchu křemíku a pro tento nejběžnější typ se ujal termín MOSFET.

Hlavním rysem MOSFETu je úplná galvanická izolace vstupu (hradla) od vodivého kanálu, a tím i od výstupu (kolektoru). Ve vysokofrekvenční technice to znamená, že vstupní impedance je téměř čistě kapacitní, a tedy při nepříliš vysokých frekvencích velmi vysoká. Schematická struktura s vodivým kanálem N-typu a různé stavy MOSFETu jsou na obr. 19. Komplementární strukturu s kanálem P si čtenář snadno dovede odvodit. Substrát je typu P. Před oxidací se vytvoří difúzí nebo implantací silně dotované ostrůvky typu  $N^+$ , tj. elektrody S a D. Po oxidaci se vrstva  $SiO_2$  všude odleptá až na plochu pod budoucí hradlovou elektrodou a pak se nanese kovové elektrody S, G a D. Uvažme nejdříve stav bez napětí mezi emitorem a kolektorem, tj. při  $U_D = 0$ , v ideální struktuře MOS, kdy zanedbáváme vliv nepohyblivých nábojů na rozhraní i v oxidu a rozdíl výstupních prací. Potenciál emitoru definujeme  $U_S = 0$ .

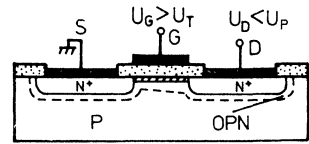
Zvětšujeme-li napětí  $U_G$  od nuly do kladných hodnot, jsou díry odpuzovány do objemu substrátu a elektrony přitahovány k rozhraní. Tím se zvětšuje oblast prostorového náboje. Koncentrace děr, původně menšinových nábojů, stoupá pod povrchem, až při jistém mezním napětí na hradle  $U_G = U_T$  je u povrchu právě tolik elektronů, jako tam bylo původně děr. Při dalším růstu  $U_G > U_T$  převládají v tenké vrstvě pod povrchem polovodiče elektrony, vzniká *inverze vodivosti*. Tato inverzní vrstva tvoří mezi ostrůvky  $N^+$  vodivý kanál typu N, do něhož může emitor dodávat bez překážek a téměř neomezeně elektrony. V ideálním případě závisí  $U_T$  jen na koncentraci akceptorů v substrátu a na tloušťce oxidu. Ve skutečnosti přítomnost kladných nábojů v oxidu i na rozhraní a rozdíl výstupních prací polovodiče a kovu hradlové elektrody působí posun  $U_T$  k záporným hodnotám.

Jak nyní probíhá proud  $I_D$  mezi emitorem a kolektorem v idealizovaném případě při různých konstantních napětích na hradle ?

1. Při napětí  $U_G < U_T$  vodivý kanál neexistuje a při vložení jakéhokoli napětí na kolektor je jeden z přechodů  $N^+P$  pólován v závěrném směru a proud je velmi malý, sotva měřitelný.

2. Při  $U_G \geq U_T$  se vytvoří vodivý kanál, proud elektronů může procházet mezi emitorem a kolektorem.
- 2a. Při malém napětí  $U_D \ll U_G$  nemá toto napětí vliv na vodivost kanálu a proud  $I_D$  je úměrný  $U_D$  (obr. 19a).
- 2b. Při stoupajících napětích  $U_D$  se mění napětí mezi hradlem a kanálem a klesá podél délky kanálu. Vodivost kanálu směrem ke kolektoru klesá, kanál se „zužuje“. Proud stoupá postupně pomaleji než lineárně. Při jistém napětí  $U_D = U_P$  se kanál úplně uzavře u kolektoru (obr. 19b), neboť napětí mezi hradlem a kolektorem už nestačí k inverzi vodivosti, proud se stává nasycený  $I_D = I_{D_{sat}}$ .
- 2c. V dalším průběhu při  $U_D > U_P$  se kanál zkracuje (obr. 19c) a proud zůstává téměř konstantní, rovný  $I_{D_{sat}}$ .

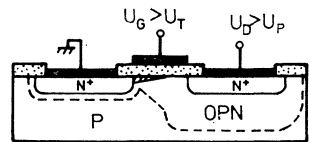
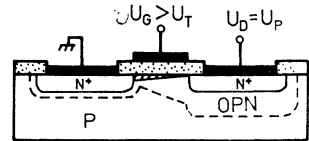
Při velmi vysokých napětích  $U_D \gg U_P$  může dojít k průrazu mezi kolektorem a emitorem, zvláště u rychlých tranzistorů s krátkým hradlem.



Obr. 19. Struktura MOSFET a vývoj vodivého kanálu při  $U_G = \text{konst.} > U_T$ ; čárkovaně značena ochuzená oblast.

- a)  $U_D \ll U_G$ ,
- b)  $U_D = U_P$ ,
- c)  $U_D > U_P$ .

OPN značí oblast prostorového náboje.



Technologie MOSFETů je obvykle jednodušší než technologie bipolárních tranzistorů, jejich příkony jsou nižší, takže se výborně hodí pro vysokou integraci. Rychlost přenosu signálu mezi vstupem (G) a výstupem (D) závisí především na rychlosti, jakou se může nabíjet kondenzátor tvořený hradlem a kanálem. Ukazuje se, že rozhodující je co největší vodivost kanálu a co nejkratší délka  $L$ . Odtud plyne snaha, aby pohyblivost, nasycená driftová rychlost a koncentrace nosičů v kanálu byly co nejvyšší. To platí obecně o všech druzích FETů.

Délka kanálu je určena rozlišovací schopností použité maskovací techniky. Fitolitografické postupy končí u rozměru  $L$  blízko k  $1 \mu\text{m}$ . Pomocí *rentgenové a elektronové litografie* je dnes možno se dostat do submikronové oblasti, takže délka  $L = 0,25 \mu\text{m}$  je už bezpečně realizovatelná. Některé prognózy předpovídají posunutí hranice rozměrů na  $0,1 \mu\text{m}$  do konce století. S těmito parametry se lze dostat do oblasti spínacích rychlostí

desítek až několika jednotek pikosekund. Tím se stávají různé typy FETů hlavním prvkem rychlých polovodičových pamětí s velmi vysokým stupněm integrace. Pokud se týká driftové rychlosti, má GaAs přednost před křemíkem, zvláště při teplotě kapalného dusíku, avšak za cenu několikanásobných technologických nákladů.

#### 5.4. Tranzistor s vysokou pohyblivostí elektronů

Snaha po zvýšení mezní driftové rychlosti vyústila v poslední době v aplikaci tzv. *dvojrozměrného elektronového plynu*. Tento jev se stal posledním „hitem“ polovodičové fyziky, neboť znamená průlom do dosavadních představ založených na modelu ideálního krystalu. Jak se k tomuto pojmu došlo? Molekulární epitaxie umožňuje vytvářet přesně kontrolované, nepředstavitelně tenké vrstvy různých složení v libovolných sledech. Tloušťku lze bezpečně kontrolovat až do rozměru řádu desítek až jednotek atomových rovin.

Taková velmi tenká vrstva např. GaAs uložená mezi vrstvami  $\text{Ga}_{0,65}\text{Al}_{0,35}\text{As}$  je pro elektrony potenciální jámou o příčných rozměrech srovnatelných s kvantově mechanickou vlnovou délkou elektronu a ukládá na pohyb elektronu nové kvantové podmínky. Hlavní z nich je, že energie elektronů může nabývat jen určitých hodnot, podobně jako je tomu v izolovaném atomu, avšak namísto diskrétních hladin jde o úzké energetické intervaly (pásky) s velmi vysokou hustotou energetických stavů. Pokud jsou ohraničující prostředí dosti rozměrná, jsou elektrony nuceny pohybovat se ve vrstvě téměř dvourozměrně; odtud je i název dvojrozměrný elektronový plyn. Pokud se velmi tenké vrstvy různých materiálů střídají, mohou tzv. tunelovým jevem elektrony procházet potenciálovými bariérami, a to napříč vrstvami. Jejich pohyb podél vrstvy je podobný pohybu v trojrozměrném krystalu, jenže mu jeden rozměr chybí.

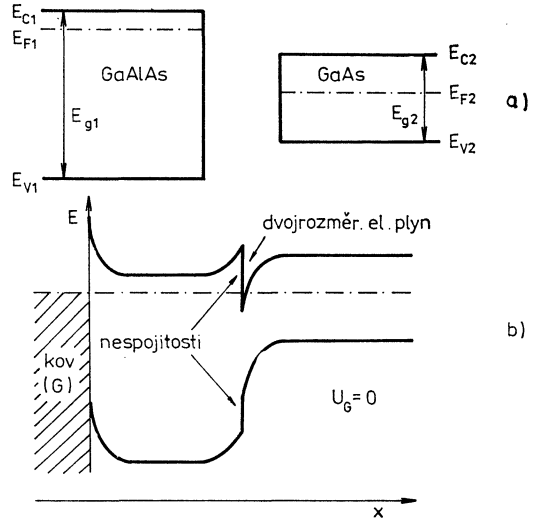
Dvojrozměrný elektronový plyn může vzniknout i na rozhraní dvou masivních polovodičů nebo jako inverzní vrstva v tranzistoru MIS, jestliže jsou vytvořeny podmínky hluboké a úzké potenciální jámy.

Pro vysvětlení tohoto jevu musíme sáhnout k *pásovému modelu* polovodiče. Hlavním jeho rysem je existence ohraničených intervalů dovolených energií elektronů, mezi nimiž je pás (ZP) zakázaných energií. Jeho spodní hranici tvoří okraj ( $E_v$ ) valenčního pásu, který je téměř zaplněn elektrony. Prázdné hladiny blízko u horního okraje valenčního pásu představují kladné díry.

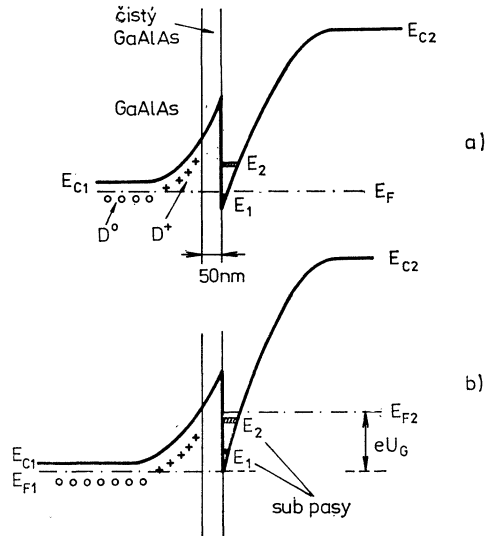
Horní hranici ZP tvoří spodní okraj ( $E_c$ ) vodivostního pásu, který je téměř prázdný až na některé hladiny blízko u okraje. Elektrony obsazující tyto hladiny, stejně jako díry ve valenčním pásu, se chovají jako volné nosiče a zprostředkují elektrický proud. V ZP mohou existovat pouze *lokalizované hladiny* související s poruchami a příměsovými atomy v krystalu. Tyto hladiny dodávají do pásu elektrony (to jsou donory), nebo odebírají z valenčního pásu elektrony a dodávají tak díry (to jsou akceptory).

Pro nás bude zajímavý případ tzv. *heteropřechodu*, tj. koherentního navázání dvou chemicky různých polovodičů s různou šířkou ZP,  $E_{g_1}$  a  $E_{g_2}$ . Budeme předpokládat, že  $E_{g_1} > E_{g_2}$  a že první polovodič je typu N a druhý velmi slabě dotovaný typu N nebo P. Na obr. 20a jsou zakreslena pásová schémata nezávislých polovodičů. Čerchovaná čára naznačuje tzv. Fermiho mez, o které si nyní nepřesně řekneme jen tolik, že je to energie

Obr. 20. Vytvoření heteropřechodu GaAlAs-GaAs:  
 a) pásová schémata oddělených polovodičů,  
 b) pásová schéma po vytvoření heteropřechodu; stav za tepelné rovnováhy (rozdíl šířky pásů  $E_{g1} - E_{g2}$  je přehnaný).



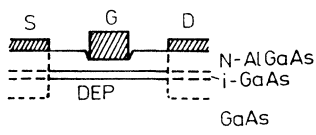
Obr. 21. Detail kvantové jámy:  
 a) bez předpětí,  
 b) s předpětím  $U_G > 0$ .



tická úroveň taková, že energetické hladiny ležící nad ní jsou prakticky prázdné a hladiny pod ní jsou prakticky obsazené. Její význačnou vlastností je, že za tepelné rovnováhy probíhá vodorovně na stejné úrovni i ve složeném systému.

Jestliže vyrobíme (molekulární epitaxi) velmi strmý heteropřechod z těchto polovodičů a stranu N opatříme Schottkyho kontaktem, vytvoří se ochuzené oblasti a v důsledku nespojitosti  $E_c$  na rozhraní přechodu vznikne v průběhu  $E_c$  na straně GaAs ostré minimum (obr. 20b). Ochuzené oblasti jsou tam, kde se okraj pásu  $E_c$  vzdaluje od Fermiho meze. Minimum  $E_c$  vytváří podmínky jednak pro vznik dvojrozměrného plynu s úzkými pásy přípustných energií ( $E_1$ ,  $E_2$ ), jednak pro vznik inverzní vrstvy v P-GaAs (obr. 21a). Jestliže na kovovou elektrodu vložíme kladné (závěrné) předpětí, posune se

levá část diagramu dolů. Jak je vidět na detailu obr. 21b, posune se tím kvantová jáma dolů tak, že Fermiho mez protne úzký pás  $E_1$ , popř. jej překročí. Tím způsobem můžeme předpětím regulovat obsazení pásů  $E_1$ , popř. i  $E_2$  elektrony, což je funkce FET. Výhoda dvojrozměrného plynu je v tom, že může nabýt velmi vysoké hustoty elektronů (až  $10^{12} \text{ cm}^{-2}$ ) při neobyčejně vysoké pohyblivosti (až  $10^6 \text{ cm}^2/\text{Vs}$  při teplotě kapalného dusíku). Vlastní funkce tohoto typu tranzistoru HEMT (High Electron Mobility Transistor) je složitá, neboť je to přechodný typ mezi MISFET a MESFET (obr. 22).



Obr. 22. Polní tranzistor s vysokou pohyblivostí elektronů (HEMT); DEP — dvojrozměrný elektronový plyn.

## Závěr

V článku nejsou ani zdaleka vyčerpávajícím způsobem probrány různé etapy vývoje tranzistorů a jim podobných součástek.

Doufám přesto, že letný pohled do dějin tranzistoru dává podnět k zamyšlení o spojení teorie a praxe, o jejich vzájemném ovlivňování a o významu nedocenitelných výsledků základního i aplikovaného výzkumu pro pokrok společnosti. Zároveň je vidět, že pokrok v elektronice, právě tak jako v jiných vědách, je založen na mezioborové spolupráci. Vezměme si např. zásluhu chemie na přípravě čistých materiálů, ohromný přínos vysokovakuové techniky a konečně i automatizovaného řízení experimentů. To jsme se ani nezmiňovali o matematických modelech, které tvoří podklad k optimalizaci konstrukce polovodičových součástek a které jsou s postupující integrací neustále náročnější. Stojí za zmínku i ekonomická stránka polovodičové elektroniky. Celkový obrat v tomto odvětví se odhaduje celosvětově na 54 miliard dolarů v r. 1989 a má stoupnout na 76 mld. dolarů do r. 1992. Jen jako příklad lze uvést, že do výrobního zařízení pro výrobu základních desek pro integrované obvody bylo investováno v r. 1988 4,2 mld. dolarů. Současný trend je takový, že malé firmy zanikají a jsou pohlcovány průmyslovými giganty. Jako příklad uvádíme, že koncern Hyundai Electronics v Již. Koreji měl jen v dokončovacím závodě v r. 1987 kapacitu montáže a testování 500 mil. součástek.

Obstojíme v této světové soutěži?

## Literatura

- [1] J. BARDEEN, W. H. BRATTAIN: Phys. Rev. 74 (1948), 230.
- [2] J. BARDEEN, W. H. BRATTAIN: Bell Syst. Tech. J. 28 (1949), 239.
- [3] J. H. SHLVE: Phys. Rev. 75 (1949), 294.
- [4] W. SHOCKLEY: Bell Syst. Tech. J. 28 (1949), 435.
- [5] H. FRANK, V. ŠNEJDAR: *Principy a vlastnosti polovodičových součástek*. SNTL Praha, 1976.
- [6] W. SHOCKLEY: *Electrons and Holes in Semiconductors*. Van Nostrand, N. York, 1950.
- [7] E. SPENKE: *Elektronische Halbleiter*. Springer Verlag, Berlin, 1965 (2. vyd.).
- [8] S. M. SZE: *Semiconductor Devices, Physics and Technology*. J. Wiley, N. York 1985.