

**Technologiekonzepte für
sub-100 nm SiGe-Hetero-Feldeffekttransistoren
zur Anwendung im Hochfrequenzbereich**

Marco Zeuner

**Technologiekonzepte für
sub-100 nm SiGe-Hetero-Feldeffekttransistoren
zur Anwendung im Hochfrequenzbereich**

**Vom Fachbereich Elektrotechnik und Informationstechnik
der Universität Hannover
zur Erlangung des akademischen Grades
Doktor-Ingenieur
genehmigte Dissertation
von**

**Diplom-Ingenieur Marco Zeuner
geboren am 4. Juni 1970 in Lemgo**

2003

1. Referent: Prof. Dr. Karl R. Hofmann

2. Referent: Prof. Dr.-Ing. Karl-Heinz Löcherer

3. Referent: Prof. Dr.-Ing. Ernst Gockenbach

Tag der Promotion: 11. Dezember 2002

Kurzfassung

Das Thema der vorliegenden Arbeit ist die wissenschaftliche Untersuchung und die experimentelle Entwicklung und Realisierung von Technologie- und Integrationskonzepten für SiGe-Hetero-Feldeffekttransistoren zur Anwendung im Hochfrequenzbereich. Darüber hinaus werden benachbarte Halbleiterkonzepte wie die Si-CMOS- und die III/V-Technologie beleuchtet, um maßgebliche Einflüsse auf die Entwicklung des SiGe-HFETs deutlich zu machen und sein Leistungspotential und seinen Stellenwert auf dem Markt der mikroelektronischen Bauelemente einordnen zu können.

Basierend auf dem Prinzip einer modulationsdotierten Heterostrukturordnung nutzt der SiGe-n-Kanal-HFET die verbesserten Transporteigenschaften eines zweidimensionalen Elektronengases (2DEG), das sich in einem zwischen relaxierten SiGe-Schichten eingebetteten, kristallografisch verspannten Si-Kanal bildet. Durch den Einsatz eines gradierten Pufferkonzeptes werden die Versetzungsdichten der Heterostruktur signifikant reduziert und gleichzeitig ein hoher Relaxationsgrad umgesetzt. Die gezielte Abstimmung der Dotierstoffkonzentrationen und der Schichtdicken im aktiven Teil des Schichtsystems führt zu hohen Elektronenbeweglichkeiten und großen Ladungsträgerdichten im 2DEG.

Im Rahmen der prozeßtechnischen Realisierung werden zwei Herstellungskonzepte im Detail betrachtet. Dabei wird dem konventionellen Technologieprozeß, dessen laterale Abmessungen fest durch das Maskenlayout vorgegeben sind, ein komplett neu entwickeltes, selbstjustierendes Integrationskonzept gegenübergestellt. Der Einsatz einer Replacement-Gate-Technik optimiert dabei die Lateralstruktur des Transistors und ermöglicht ein Skalieren der Gatelänge bis in den sub-100 nm Bereich. Außerdem umgeht eine rein lichtoptische Realisierung des T-Gates die zeit- und kostenintensive Elektronenstrahlolithografie und liefert zusätzliche Freiheitsgrade bei der Dimensionierung der Bauelementstruktur.

Die gezielte Optimierung des SiGe-Heteroschichtsystems und die Einführung des neuen selbstjustierenden Integrationskonzeptes führen zu einer maßgeblichen Verbesserung der elektrischen Eigenschaften des SiGe-HFETs im DC- und HF-Bereich. Reduzierte Leckströme und eine Verringerung der Unterschwellsteilheit im Abschnürbereich des Bauelementes, hohe Sättigungsströme von über 500 mA/mm und eine maximale Steilheit von 685 mS/mm kennzeichnen den Transistor im Gleichstrombetrieb. Im Hochfrequenzbereich markiert ein f_T von 90 GHz den derzeitigen Höchstwert für SiGe-basierte FETs, und f_{max} -Werte von bis zu 120 GHz unterstreichen ein hohes Potential bei der Leistungsverstärkung. Eine minimale Rauschzahl F_{min} von 0,3 dB und ein G_{ass} von 19 dB bei einer Frequenz von 2,5 GHz belegen die Konkurrenzfähigkeit des SiGe-HFETs zu anderen Hochfrequenzbauelementen.

Schlagworte: SiGe-HFET, Replacement-Gate, Hochfrequenztransistor

Abstract

The present work focuses on the research and development of new technology concepts for integrating and optimizing n-type SiGe hetero field-effect transistors used in high frequency applications. Additionally, other FET concepts like the Si CMOS- or the III/V-HEMT-technology are considered to clarify the crucial technical influences on the SiGe-HFET and to compare its performance to competitive microelectronic devices on the RF-market.

The SiGe n-channel HFET is based on a modulation-doped heterostructure with a biaxially strained Si channel embedded between unstrained SiGe layers. Thus, a quantum well is formed in the conduction band of the channel providing a two dimensional electron gas (2DEG) with advanced carrier transport properties, ideal for microwave applications. A compositionally graded SiGe buffer serves as a virtual substrate for the active layer stack. This concept enables very low defect densities and high degrees of relaxation. Optimizing carefully the configuration of the sophisticated active layer stack yields an elevated electron mobility and a high carrier density in the 2DEG.

The conventional technology concept of the SiGe-HFET consists of a mesa-isolated device structure with implanted ohmic contacts and a T-shaped Schottky-gate made by e-beam-lithography. Since the performance of this device type suffers from the substantial influence of various parasitic elements, a completely new integration concept is presented in this work. Based on a self-aligned transistor structure realized by a replacement-gate technique, this process offers a higher degree of lateral device optimization and a reproducible scalability towards sub-100 nm gate-lengths without using the expensive and time-consuming e-beam-lithography.

The electrical characterization of the developed and fabricated transistors confirms the promising microwave potential of the SiGe-HFET. Due to a modified configuration of the Si/SiGe hetero layer stack and the introduction of the new self-aligned technology concept the electrical performance of the transistor has been significantly improved. Low leakage currents in the pinch-off region, high saturation currents of about 500 mA/mm and a maximum transconductance of 685 mS/mm characterize the DC-behavior of the device. Essential figures of merit to specify the RF-performance of the transistor are the cut-off frequencies f_T and f_{max} . Maximum f_T values of about 90 GHz and up to 120 GHz for f_{max} demonstrate the high potential for microwave applications. Furthermore, the SiGe-HFET yields a low minimum noise figure F_{min} of 0.3 dB up to 2.5 GHz with a maximum in the associated gain G_{ass} of 19 dB.

Keywords: SiGe-HFET, replacement-gate, RF-transistor

Inhaltsverzeichnis

| | | |
|------------------|--|-----------|
| Kapitel 1 | Einleitung | 1 |
| 1.1 | Einführung und Motivation | 1 |
| 1.2 | Gliederung und Ziele dieser Arbeit | 3 |
| Kapitel 2 | SiGe – zwischen Si- und III/V-Technologie – Kompromiß oder Alternative? | 7 |
| 2.1 | Mainstream – Silizium-CMOS-Technologie | 7 |
| 2.1.1 | Die SIA-Roadmap – Taktgeber für die Halbleiterindustrie | 7 |
| 2.1.2 | CMOS – Stand der Technik und Zukunftsperspektive | 9 |
| 2.2 | Technologie der III/V-Halbleiter | 20 |
| 2.2.1 | Materialeigenschaften | 20 |
| 2.2.2 | Heterostruktur-Feldeffekttransistoren auf III/V-Basis | 22 |
| 2.3 | SiGe-Technologie – das Hochfrequenzpotential der Siliziumbranche | 28 |
| 2.3.1 | Entwicklung der SiGe-Technologie und Stand der Technik | 28 |
| 2.3.2 | Das Leistungspotential des SiGe-HFETs | 33 |
| Kapitel 3 | Aufbau und Funktion des SiGe-n-Kanal-Hetero-Feldeffekt- transistors | 40 |
| 3.1 | Grundlagen zum SiGe-Materialsystem | 40 |
| 3.1.1 | $\text{Si}_{1-x}\text{Ge}_x$ -Mischkristalle | 40 |
| 3.1.2 | Verspannte SiGe-Heterostrukturen | 42 |

| | |
|--|-----------|
| 3.2 Bauelementstruktur und Funktionsprinzip des SiGe-HFETs | 50 |
| 3.3 Eigenschaften des 2DEGs | 52 |
| 3.3.1 Ladungsträgerbeweglichkeit und Streuprozesse | 52 |
| 3.3.2 Ladungsträgerdichte | 55 |
| 3.4 Ladungsträgersteuerung im Gleichstromfall | 58 |
| 3.5 Kleinsignalersatzschaltbild | 64 |
| Kapitel 4 Dimensionierung und Konfiguration des SiGe-Hetero- schichtsystems | 69 |
| 4.1 Pufferschichten | 69 |
| 4.1.1 Klassische Pufferkonzepte | 70 |
| 4.1.2 Alternative Pufferkonzepte | 73 |
| 4.2 Dimensionierung und Auslegung der aktiven Schichten | 74 |
| 4.2.1 Einfluß des Ge-Gehaltes auf die Schichtparameter | 76 |
| 4.2.2 Die kritische Schichtdicke des Kanals | 77 |
| 4.2.3 Konfiguration der Dotierschichten | 79 |
| 4.2.4 Dimensionierung der Spacerdicke | 83 |
| 4.2.5 Funktion und Auslegung der Cap-Schichten | 85 |
| Kapitel 5 Integrationskonzepte des SiGe-Hetero-Feldeffekttransistors ... | 86 |
| 5.1 Das konventionelle Integrationskonzept | 87 |
| 5.1.1 Bauelementstruktur und Herstellungsablauf | 87 |
| 5.1.2 Modifikation und partielle Optimierung des Konzeptes | 89 |
| 5.2 Das selbstjustierende Integrationskonzept | 94 |
| 5.3 Vergleich der SiGe-HFET Technologievarianten | 99 |

| | | |
|------------------|---|------------|
| Kapitel 6 | Technologische Realisierung der Herstellungskonzepte | 103 |
| 6.1 | Niedertemperaturtechnologien für SiGe-HFETs | 104 |
| 6.2 | Optimierung der ohmschen Kontakte | 111 |
| 6.2.1 | Niederohmige Kontaktierung durch Ionenimplantation | 112 |
| 6.2.2 | Optimierung der Kontakte durch Silizidierung | 119 |
| 6.3 | Strukturierung des Dummy-Gates | 127 |
| 6.3.1 | Definition und trockenchemische Ätzung des Dummy-Gates | 128 |
| 6.3.2 | Naß- und trockenchemische Skalierung des Dummy-Gates | 137 |
| 6.4 | Planarisierungskonzepte des SiGe-HFETs | 141 |
| 6.5 | Gaterealisierung | 150 |
| 6.5.1 | Entwicklung temperaturstabiler Schottky-Übergänge auf Pt/Si-Basis | 151 |
| 6.5.2 | Gateherstellung mit Hilfe der e-Beam-Lithografie | 158 |
| 6.5.3 | Selbstjustierendes T-Gatekonzept | 162 |
| Kapitel 7 | Elektrische Charakterisierung und Simulation der SiGe-HFETs | 168 |
| 7.1 | Analyse des Gleichstromverhaltens | 169 |
| 7.2 | Charakterisierung im Hochfrequenzbereich | 182 |
| 7.2.1 | Verstärkungen und Grenzfrequenzen der Transistoren | 183 |
| 7.2.2 | Simulation und Extraktion des Kleinsignalersatzschaltbildes | 192 |
| 7.2.3 | Analyse des Rauschverhaltens | 200 |
| 7.3 | Das Leistungspotential der Transistoren bei kryogenischen Temperaturen | 206 |
| Kapitel 8 | Zusammenfassung und Ausblick | 212 |
| | Literaturverzeichnis | 218 |

Verzeichnis der verwendeten Symbole und Abkürzungen

Symbole:

| | |
|------------------------------------|--|
| A | Anisotropie |
| \ddot{A}_{Au} | elektrochemisches Äquivalent von Gold |
| A_G | Querschnittsfläche des Gates |
| $\underline{a}_i, \underline{b}_i$ | Amplituden der Leistungswellen ($i = 1, 2$) |
| a_{Si}, a_{Ge} | Gitterkonstante von Silizium bzw. Germanium |
| a_{SiGe} | Gitterkonstante von Silizium-Germanium Legierungen |
| a_{Sub} | Gitterkonstante des Substrates |
| BV_{CEO} | Durchbruchspannung zwischen Kollektor und Emitter |
| c | Tensor der Elastizitätskonstanten |
| C_0 | flächennormierte Eigenkapazität des 2DEGs |
| C_{ds} | Drain-Source Kapazität |
| C_{gd} | Gate-Drain Kapazität |
| C_{gdo} | Überlappkapazität des Gates auf der Drainseite |
| C_{gs} | Gate-Source Kapazität |
| C_{gso} | Überlappkapazität des Gates auf der Sourceseite |
| C_{pds} | Drain-Source Padkapazität |
| C_{pgd} | Gate-Drain Padkapazität |
| C_{pgs} | Gate-Source Padkapazität |
| d_{Cap} | Dicke der Cap-Schichten |
| d_d | Dicke der Dotierschicht |
| d_{Fox} | Dicke des Feldoxides |
| d_{GK} | Abstand zwischen Gateelektrode und Kanal |
| d_i | Dicke des Spacers zwischen Dotierschicht und Kanal |
| d_{krit} | kritische Schichtdicke |
| d_{Mesa} | Mesahöhe |
| D_n^{2D} | Zustandsdichte der Subbänder |
| d_{PR} | Dicke des Fotolackes |
| E | elektrische Feldstärke |
| E_F | Fermi-Energie |
| E_g | Bandabstand |

Symbole

| | |
|---------------------------------|--|
| E_i | i-tes Energiesubbandniveau |
| E_{krit} | kritische Feldstärke |
| E_L, E_V | Energieniveau des Leitungs- bzw. Valenzbandes |
| E_{Strahl} | Strahlungsenergie |
| $F(E)$ | Fermi-Dirac Verteilungsfunktion |
| f | Frequenz |
| f_C | Eckfrequenz des 1/f-Rauschens |
| f_{max} | Grenzfrequenz der Leistungsverstärkung |
| F_{min} | minimale Rauschzahl |
| f_T | Grenzfrequenz der Kurzschlußstromverstärkung |
| G_{ass} | mit F_{min} assoziierte Leistungsverstärkung |
| g_{ds} | Ausgangsleitwert |
| g_m | Steilheit |
| G_R | galvanische Abscheiderate |
| h | Plancksches Wirkungsquantum = $6,626 \cdot 10^{-34}$ Js |
| \underline{h}_{21} | Kurzschlußstromverstärkung |
| I_C | Kollektorstrom |
| I_D | Drainstrom |
| I_{Dsat} | Drainsättigungsstrom |
| I_G | Gatestrom |
| $I_{\text{on}}, I_{\text{off}}$ | definierte Stromwerte, die den On- und Off-Zustand des Transistors festlegen |
| J | Stromdichte |
| j_D | auf die Gateweite normierte Stromdichte |
| K | Anpassungsfaktor der Fukui-Formel |
| k | Boltzmann-Konstante = $1,381 \cdot 10^{-23}$ J/K |
| k | Stabilitätsfaktor |
| L_C | Breite des ohmschen Kontaktes |
| L_G | Gatelänge |
| L_{MK} | Abstand zwischen der Metallisierung und dem Kanalgebiet |
| L_s, L_d, L_g | Source-, Drain- und Gateinduktivität |
| L_{SD} | Abstand zwischen der Source- und Drainimplantation |
| L_T | Transferlänge |
| m | Steigung einer Geraden |
| m^* | effektive Masse |
| m_0 | Ruhemasse des Elektrons |
| m_t, m_l | transversale und longitudinale Masse |
| M_W | Molekülmasse |
| n | Idealitätsfaktor der Schottky-Diode |

Verzeichnis der verwendeten Symbole und Abkürzungen

| | |
|-----------------------|---|
| n | Anzahl der Gatefinger |
| n, p | n- bzw. p-leitender Ladungsträger- oder Halbleitertyp |
| N_0 | implantierte Dosis |
| N_A, N_D | Dotierstoffkonzentration an Akzeptoren bzw. Donatoren |
| n_d | Verarmungsladungsdichte der Dotierschicht |
| n_s, p_s | Flächendichte im 2DEG bzw. 2DHG |
| p | Druck |
| q | Elementarladung = $1,602 \cdot 10^{-19}$ C |
| Q_s | elektrische Ladung im Si-Kanal |
| r | Ätzrate |
| R_C | Kontaktwiderstand |
| R_{ds} | Drain-Source Widerstand |
| R_i | Widerstand zwischen Gateelektrode und Si-Kanal |
| R_{Impl} | Bahnwiderstand der Implantationsgebiete |
| R_K | Widerstand des gesteuerten Si-Kanals |
| R_{Kanal} | Widerstand des gesamten Si-Kanals |
| R_M | Metallwiderstand |
| R_n | äquivalenter Rauschwiderstand |
| R_p | projizierte Reichweite bei der Ionenimplantation |
| R_s, R_d, R_g | Source,- Drain- und Gatewiderstand |
| $R_{Schicht}$ | Schichtwiderstand |
| R_{Si} | Widerstand des ungesteuerten Si-Kanals |
| S | Ätzselektivität |
| $\underline{S}_{i,j}$ | Streuparameter ($i, j = 1, 2$) |
| T | Temperatur |
| T_{ox} | Gateoxiddicke |
| V | elektrisches Potential |
| v | Ladungsträgergeschwindigkeit |
| v_d | Driftgeschwindigkeit |
| V_{dd} | Versorgungsspannung |
| V_{DS} | Drain-Source Spannung |
| $V_{DS,sat}$ | Drain-Source Spannung im Sättigungsbereich |
| V_G | Gatespannung |
| V_{GS} | Gate-Source Spannung |
| V_P | Pinch-off Spannung |
| v_{sat} | Sättigungsdriftgeschwindigkeit |
| V_T | Schwellspannung |
| W_G | Gateweite |
| x, y | Molanteile der verschiedenen Halbleiterkomponenten |

| | |
|--------------------------|--|
| x, y, z | Ortskoordinaten |
| Y_i | Admittanzparameter ($i = 1..4$) |
| Z_i | Impedanzparameter ($i = 1..4$) |
| α | Anstellwinkel |
| α | Skalierungsfaktor |
| α_K | kathodische Effizienz des Galvanikprozesses |
| Γ_{opt} | optimaler Generatorreflexionsfaktor für F_{min} |
| Δd | Abstand des Ladungsträgerschwerpunktes des 2DEGs zum Potentialtopfrand |
| $\Delta E_L, \Delta E_V$ | Leitungsband- bzw. Valenzbanddiskontinuität |
| ΔR_p | Standardabweichung der projizierten Reichweite |
| ε | Dehnungstensor |
| ε_0 | Vakuumdielektrizitätskonstante = $8,854 \cdot 10^{-12}$ As/Vm |
| ε_r | relative Dielektrizitätskonstante |
| η | Gitterfehlpassung |
| μ_0 | Niederfeldbeweglichkeit |
| μ_n, μ_p | Elektronen-, Löcherbeweglichkeit |
| Ξ_d, Ξ_u, a, b | Deformationspotentiale |
| ρ_C | spezifischer Kontaktwiderstand |
| ρ_g | spezifischer Widerstand des Gates |
| $\rho_{M,Au}$ | spezifisches Gewicht von Gold |
| ρ_S | Schichtwiderstand des Implantationsgebietes |
| σ | Spannungstensor |
| τ | Zeitkonstante |
| τ_T | Transitzeit |
| Φ_B | Schottky-Barriere |
| Φ_M | Austrittsarbeit des Metalls |
| Ψ_n | n-te Wellenfunktion |
| ω | Kreisfrequenz ($\omega = 2\pi f$) |

Abkürzungen:

| | |
|------------|--|
| 2DEG, 2DHG | Zweidimensionales Elektronengas bzw. Löchergas |
| AFM | Atomic-Force Microscopy |
| AGC | Automatic Gain Control |
| Al | Aluminium |
| Al_2O_3 | Aluminiumoxid |
| AlGaAs | Aluminium-Galliumarsenid |
| AMPS | Advanced Mobile Phone Service |

| | |
|-----------------------|---|
| ASIC | Application Specific Integrated Circuit |
| Au | Gold |
| BiCMOS | B ipolar C omplementary M etal- O xide- S emiconductor |
| BJT | B ipolar J unction T ransistor |
| BPSG | B or- P hosphor- S ilikatglas |
| CCB | Constant Composition B uffer |
| CMOS | C omplementary M etal- O xide- S emiconductor |
| CMP | C hemical- M echanical- P olishing |
| Co, CoSi ₂ | Kobalt, Kobaltdisilizid |
| CPU | C entral P rocessing U nit |
| Cu | Kupfer |
| CVD | C hemical V apor D eposition |
| DC | D irect C urrent (Gleichstrom) |
| DECT | D igital E uropean C ordless T elephone |
| DGPS | D ifferential G lobal P osition S ystem |
| DMUX | D emultiplexer |
| DNS | D igital N etwork S witch |
| DRAM | D ynamic R andom- A ccess M emory |
| DSI | D oping by S econdary I on I mplantation |
| DSRC | D edicated S hort R ange C ommunication |
| DTI | D eep T rench I solation |
| DUV | D eep U ltraviolet |
| ETC | E lectronic T oll C ollection |
| EUVL | E xtrême U ltraviolet L ithography |
| FOX | F eldoxid |
| FUV | F ar U ltraviolet |
| GaAs | Galliumarsenid |
| GPS | G lobal P osition S ystem |
| GSM | G lobal S ystem for M obile C ommunication |
| HBT | H etero- B ipolartransistor |
| HCMOS | H etero C omplementary M etal- O xide- S emiconductor |
| HEMT | H igh E lectron- M obility T ransistor |
| HF | H ochfrequenz |
| HFET | H etero F ield- E ffect T ransistor |
| HSQ | H ydrogen- S ilsesquioxane |
| IC | I ntegrated C ircuit (integrierte Schaltung) |
| ICP | I nductively C oupled P lasma |
| ILD | I nterlevel- D ielectric |
| InAlAs | Indium-Aluminiumarsenid |

| | |
|----------------|--|
| InAlGaAs | Indium-Aluminium-Galliumarsenid |
| InAs | Indiumarsenid |
| InGaAs | Indium-Galliumarsenid |
| InP | Indiumphosphid |
| Ir, IrSi | Iridium, Iridiumsilizid |
| JFET | J unction F ield- E ffect T ransistor |
| LDD | L ightly D oped D rain |
| LNA | L ow- N oise A mplifier |
| LOCOS | L ocal O xidation of S ilicon |
| LPCVD | L ow P ressure C hemical V apor D eposition |
| LTE | L ow T emperature E pitaxy |
| M1, M2, M3, M4 | Metallisierungsebenen |
| MAG | M aximum- A vailable G ain |
| MBE | M olecular B eam E pitaxy |
| MESFET | M etal- S emiconductor F ield- E ffect T ransistor |
| MIBK | M ethylisobutylketon |
| MMIC | M onolithic M icrowave I ntegrated C ircuit |
| MOCVD | M etal- O rganic C hemical V apor D eposition |
| MODFET | M odulation- D oped F ield- E ffect T ransistor |
| MODQW | M odulation- D oped Q uantum- W ell |
| MOSFET | M etal- O xide- S emiconductor F ield- E ffect T ransistor |
| MPU | M icroprocessor U nit |
| MSG | M aximum- S table G ain |
| MUX | M ultiplexer |
| Ni, NiSi | Nickel, Nickelsilizid |
| OAI | O ff- A xis I llumination |
| PA | P ower A mplifier |
| PAE | P ower A dded E fficiency |
| PCS | P ersonal C ommunication S ervices |
| Pd | Palladium |
| PECVD | P lasma- E nhanced C hemical V apor D eposition |
| PI | P olyimid |
| PMMA | P oly m ethyl m ethacrylat |
| PSM | P hase- S hift M asking |
| Pt, PtSi | Platin, Platinsilizid |
| PVD | P hysical V apor D eposition |
| QW | Q uantum- W ell |
| RAM | R andom- A ccess M emory |
| RBS | R utherford B ackscattering S pectroscopy |

Verzeichnis der verwendeten Symbole und Abkürzungen

| | |
|--|---|
| REM | R asterelektronen m ikroskopie |
| RIE | R eactive- I on- E tch |
| RTA | R apid T hermal A nnealing |
| Sb | Antimon |
| Si | Silizium |
| Si ₃ N ₄ | Siliziumnitrid |
| SIA | S emiconductor I ndustry A ssociation |
| SiGe | Silizium-Germanium |
| SIMS | S econdary I on M ass S pectroscopy |
| SiO ₂ | Siliziumdioxid |
| SiOC | Carbon-Silikatglas |
| SiOF | Fluor-Silikatglas |
| SOG | S pin- O n- G las |
| SONET | S ynchronous O ptical N etwork |
| Ta ₂ O ₅ | Tantaloxid |
| TEM | T ransmission- E lectron- M icroscopy |
| TEOS | T etraethyl o rthosilikat |
| Ti, TiSi ₂ | Titan, Titandisilizid |
| TiN | Titannitrid |
| TiO _x | Titanoxid |
| TLM | T ransmission- L ine M odell |
| U | U nilateral G ain |
| VCO | V oltage- C ontrolled O scillator |
| WLAN | W ireless L ocal A rea N etwork |
| WLL | W ireless L ocal L oop |
| WTi | Wolframtitan |
| XRD | X - R ay D iffractometry |
| ZrO ₂ | Zirkoniumoxid |
| HF, BHF, HNO ₃ , H ₂ SO ₄ , H ₂ O ₂ , HCl, H ₃ PO ₄ , NH ₄ F | Summenformeln handelsüblicher Chemikalien |
| O ₂ , Ar, CF ₄ , SF ₆ , CHF ₃ | Summenformeln handelsüblicher Ätzgase |
| CO, CO ₂ , C ₂ H ₂ | gasförmige Ätzprodukte |
| PH ₃ , B ₂ H ₆ | Dotiergase beim CVD-Verfahren |

Kapitel 1

Einleitung

1.1 Einführung und Motivation

Die gesellschaftlichen und wirtschaftlichen Bedürfnisse im Informationszeitalter zeichnen sich durch eine steigende Forderung nach permanenter und ortsunabhängiger Verfügbarkeit von Informationen aus [1.1]. Darüber hinaus schafft die Digitalisierung eine Verflechtung bislang getrennter Medienformen, wie Text-, Sprach- und Bildinformationen, zu einer gemeinsamen multimedialen Basis. Die Informationsübertragung erfolgt dabei mit stetig ansteigenden Datenraten und zunehmend schnurlos über Satelliten, Mobilfunk und Radaranwendungen oder optisch über Glasfasernetze [1.2]. Die daraus erwachsende Flut von Datenmengen, die notwendig ist, um diesen Informationsaustausch zu gewährleisten, stellt hohe Ansprüche an die Kommunikationstechnik der Gegenwart und der Zukunft. Insbesondere die integrierten Schaltkreise (ICs) zur Verarbeitung der Informationen müssen den Anforderungen nach hohen Verstärkungen, niedrigen Rauschzahlen und steigenden Frequenzbereichen, bei gleichzeitig günstigen Herstellungskosten nachkommen.

In den Marktsegmenten der flüchtigen Speicher (DRAMs), der Mikroprozessoren und der Logikbausteine hat die Silizium-CMOS-Technologie ihre dominierende Stellung in den letzten Jahrzehnten permanent ausbauen können und bedient derzeit über 80 % des gesamten Halbleitermarktes [1.3]. Den Bereich der Hochfrequenzschaltungen hingegen teilen sich die III/V-Halbleiter und die SiGe-BiCMOS-Technologie [1.4]. Aufgrund der hohen Elektronenbeweglichkeit und der direkten Bandlücke sind die Verbindungshalbleiter für mikroelektronische Anwendungen bei hohen Frequenzen und in der Optoelektronik prädestiniert [1.5]. Zusätzlich begünstigt das semiisolierende Substrat die Integration der Einzelbauelemente zu anwendungsspezifischen MMICs höchster Güte. Trotz schlechterer mechanischer und thermischer Eigenschaften und einem deutlich

höheren Material- und Herstellungspreis im Vergleich zum Silizium, führen die überragenden physikalischen Vorteile der III/V-Halbleiter zu einer Vormachtstellung im Bereich der Hochfrequenzanwendungen oberhalb von 50 GHz.

Bestrebungen, die etablierte, kostengünstige und prozeßtechnisch gut beherrschbare Siliziumtechnologie für den Hochfrequenzbereich nutzbar zu machen, führten zur Implementierung des Germaniums in das Si-Materialsystem. Die Entwicklung des SiGe-Heterobipolartransistors Mitte der 80er Jahre ist als eine Folge der fortschreitenden schaltungsorientierten Optimierung des Bipolartransistors in der Siliziumtechnologie zu sehen. Hohe Stromverstärkungen und herausragende Rausch- und Hochfrequenzeigenschaften resultierten in der kommerziellen Markteinführung des Bauelementes bereits Mitte der 90er Jahre und eröffneten ihm ein breites Anwendungspotential im Telekommunikationsmarkt [1.6]. Die hohe Kompatibilität des Herstellungsprozesses zur CMOS-Technologie machte eine Verschmelzung der beiden Technologien naheliegend. Die SiGe-BiCMOS-Technologie profitiert dabei sowohl von den guten Hochfrequenzeigenschaften des HBTs als auch von den Vorzügen des CMOS-Prozesses für Logikschaltungen [1.7]. Dadurch sind analoge und digitale Funktionen auf einem Chip, mit einer ausgereiften und hoch integrierbaren Technologie umsetzbar.

Die intensive Erforschung des Si/SiGe-Heterosystems in den 80er Jahren führte zu der Idee, die verbesserten Transporteigenschaften der Ladungsträger auch für Feldeffekttransistoren zu nutzen. Basierend auf dem Prinzip des ersten modulationsdotierten GaAs/AlGaAs-Heterostruktur-Feldeffekttransistors von Mimura et al. [1.8] aus dem Jahre 1980, entwickelten Dämbkes et al. 1985 den ersten n-MODFET im Si/SiGe-Materialsystem [1.9]. Ein kristallografisch verspannter, zwischen relaxierten SiGe-Schichten eingebetteter Si-Kanal bildet dabei ein von Dotierschichten gespeistes, zweidimensionales Elektronengas (2DEG), mit deutlich erhöhter Ladungsträgerbeweglichkeit im Vergleich zum Inversionskanal von reinen Si-MOSFETs. Ein analoges Vorgehen beim SiGe-p-MODFET nutzt einen pseudomorphen SiGe- oder reinen Ge-Kanal, um einen ausgeprägten Quantensprung im Valenzband zu realisieren und dadurch erhöhte Löcherbeweglichkeiten zu erzielen [1.10]. Grundlage dieser neuartigen Transistortypen sind hochentwickelte und optimal abgestimmte Si/SiGe- bzw. Ge/SiGe-Schichtstrukturen, mit abrupten Heteroübergängen und scharfen Dotierstoffgradienten.

Erst die Weiterentwicklung kristalliner Wachstumsverfahren, wie der Molekularstrahlepitaxie (MBE) und der Chemical Vapor Deposition (CVD), in den 70er

und 80er Jahren legte den Grundstein für ein atomares Schichtwachstum mit hoher Kristallqualität. Einen weiteren Meilenstein zur vertikalen Optimierung der Transistorstruktur setzte die Einführung des relaxierten und gradierten SiGe-Puffers 1991 [1.11]. Dieser zeichnet sich im Gegensatz zur Pufferschicht mit konstantem Ge-Gehalt durch drastisch reduzierte Versetzungsdichten in der aktiven Schichtstruktur aus. Daraus resultiert direkt eine Verbesserung der Ladungsträgerbeweglichkeit im Kanal, mit den höchsten derzeit veröffentlichten Elektronenbeweglichkeiten für modulationsdotierte Schichtstrukturen von $2830 \text{ cm}^2/\text{Vs}$ bei 300 K und $18000 \text{ cm}^2/\text{Vs}$ bei 77 K [1.12].

Neben der vertikalen Optimierung des Transistorschichtdesigns spielt die laterale Bauelementstruktur eine entscheidende Rolle für die Leistungsfähigkeit des FETs. Da die Anfangszeit der SiGe-HFETs sowohl im Schichtaufbau als auch im Layout und im Herstellungsprozeß stark von den III/V-HEMTs geprägt wurde, zeigen sich hier viele Parallelen. So zeichnet sich das typische konventionelle Technologiekonzept des SiGe-HFETs durch eine mesaisolierte, passivierte Bauelementstruktur aus, die mit Hilfe von elektronenstrahlgeschriebenen Multifinger-Schottky-Gates gesteuert wird [1.13]. Zwar hat die Weiterentwicklung der Elektronenstrahlithografie die Skalierung der Gatelänge in den letzten Jahren stark vorangetrieben, jedoch nimmt dieses konventionelle Konzept, aufgrund eines relaxierten Transistordesigns, zu hohe parasitäre Einflüsse in Kauf, die das Leistungspotential des Transistors mindern. Mit steigender Integrationsdichte gewinnen diese parasitären Elemente zunehmend an Bedeutung für das elektrische Transistor- und Schaltungsverhalten, so daß die Anforderungen an das verwendete Integrationskonzept stetig ansteigen. Ein lateraloptimiertes Bauelementdesign verlangt deshalb nach selbstjustierenden Integrationskonzepten, die auch bei Gatelängen im sub-100 nm Bereich eine hohe Leistungsausbeute realisieren und gleichzeitig die Prozeßkosten, durch Einsparung zeit- und kostenintensiver Herstellungsschritte, drastisch reduzieren.

1.2 Gliederung und Ziele dieser Arbeit

Die vorliegende Arbeit beschäftigt sich in Kapitel 2 zunächst mit der Si-CMOS-Technologie als dem dominierenden Marktführer in der Halbleiterbranche. Anhand der SIA-Roadmap wird der Stand der Technik dargestellt und eine Prognose für die zukünftige Entwicklung bis ins Jahr 2016 vorgenommen. Die Beschreibung eines charakteristischen CMOS-Herstellungsprozesses zeigt die

Grundprinzipien dieses Technologiekonzeptes und beleuchtet die Probleme der fortschreitenden Miniaturisierung. Der nächste Abschnitt beschäftigt sich zuerst mit den charakteristischen Materialeigenschaften von III/V-Halbleitern. Danach wird am Beispiel der speziellen Technologiekonzepte der III/V-Hetero-Feldeffekttransistoren die gezielte Ausrichtung des Materialsystems auf Hochfrequenzanwendungen dargestellt. Im Anschluß daran erfolgt die Einführung in die SiGe-Technologie. Anhand des SiGe-HBTs werden die Vorteile gegenüber dem konventionellen Si-Bipolartransistor erarbeitet und das Leistungs- und Anwendungspotential, insbesondere in Verbindung mit der Standard-Si-CMOS-Technologie als SiGe-BiCMOS Konzept, beleuchtet. Abschließend soll mit Hilfe von Simulationen und experimentellen Ergebnissen das Leistungspotential von SiGe-HFETs bestimmt und demonstriert werden. Außerdem geben Kostenabschätzungen Auskunft über die Wirtschaftlichkeit des Bauelementkonzeptes im direkten Vergleich zur Si- und GaAs-Technologie.

Kapitel 3 gibt eine Einführung in die physikalischen Grundlagen des SiGe-Materialsystems und beschreibt die spezifischen Eigenschaften von Si/SiGe-Heteroschichtsystemen. Anschließend erfolgt die Darstellung des Aufbaus und der Wirkungsweise des SiGe-HFETs, unter Herleitung einer analytischen Strom-Spannungs-Beziehung. Für den Kleinsignalbetrieb werden die Grenzfrequenzen der Strom- und Leistungsverstärkung definiert und ein entsprechendes Ersatzschaltbild vorgestellt.

Den thematischen Schwerpunkt des Kapitels 4 bildet die Dimensionierung und Konfiguration des heterogenen Transistorschichtsystems. Dazu werden zunächst verschiedene SiGe-Pufferkonzepte vorgestellt und bewertet, bevor die Auslegung der aktiven Schichten erfolgt. Diese vertikale Optimierung des Bauelementes verfolgt dabei die Zielsetzung, das elektrische Potential des Transistors für Hochfrequenzanwendungen zu maximieren.

Für die laterale Optimierung des SiGe-HFETs werden in Kapitel 5 zwei Integrationskonzepte vorgestellt, die in ihren Prinzipien komplett unterschiedliche Ursprünge aufweisen (Abbildung 1.1). Die konventionelle Technologie ist historisch von den III/V-HEMTs abgeleitet und basiert auf der Umsetzung des Schottky-Gates mittels Elektronenstrahlolithografie. Das komplett neu entwickelte selbstjustierende Konzept nutzt eine Replacement-Gate-Struktur zur maskenlosen Ausrichtung der ohmschen Implantationsgebiete und bedient sich dabei nur optischer Lithografieschritte.

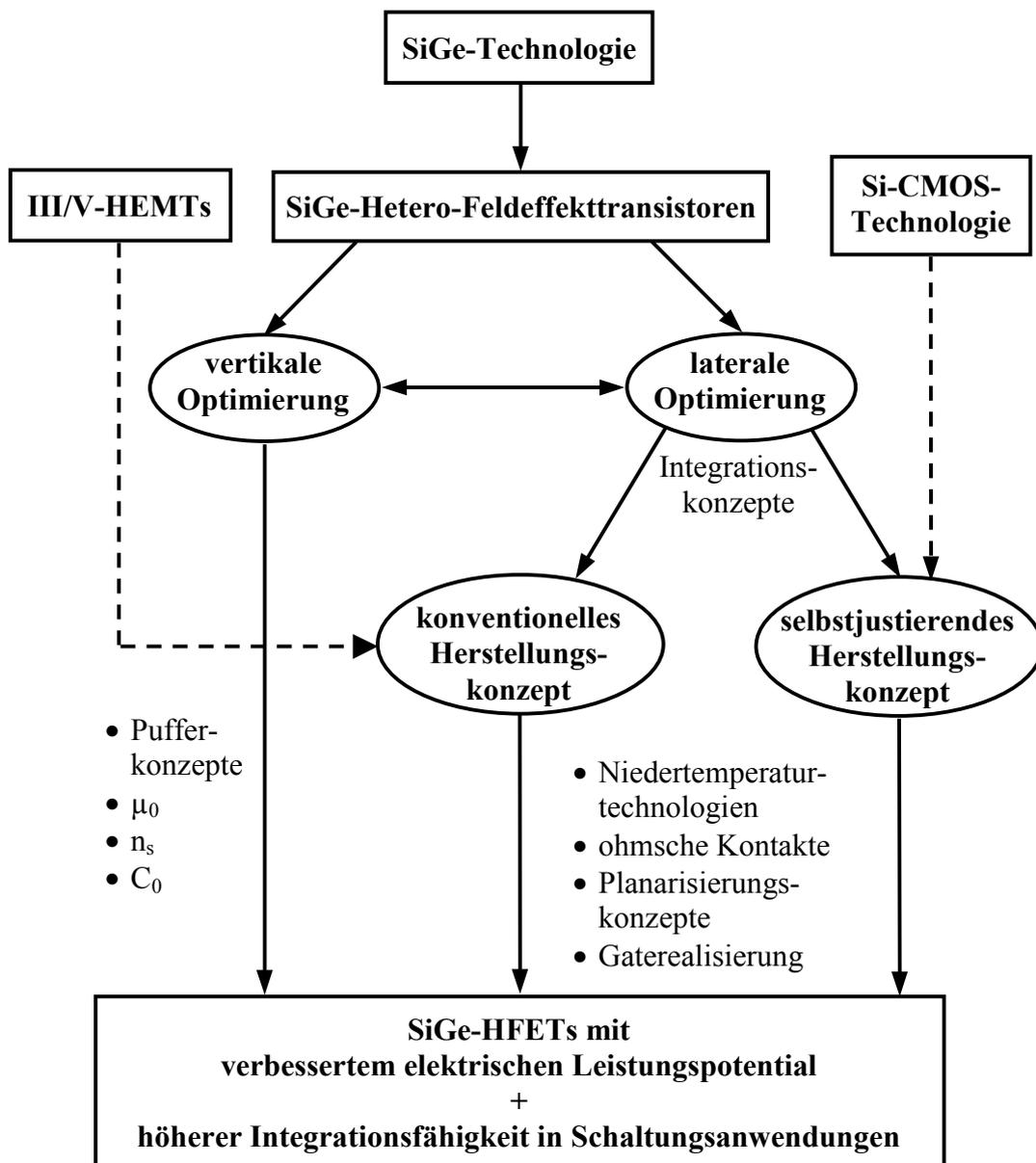


Abbildung 1.1: Funktionale Struktur der vorliegenden Arbeit unter Berücksichtigung der Einflußfaktoren der Silizium-CMOS-Technologie und der III/V-HEMT-Technologie

Die zur Realisierung der vorgestellten Technologiekonzepte erforderlichen Prozeßmodule und Einzelkomplexe beschreibt Kapitel 6. Auf der Basis des experimentell ermittelten maximalen Temperaturbudgets für den Herstellungsprozeß werden hier zuerst Optimierungsverfahren für die ohmsche Kontaktierung des

Transistors erarbeitet. Neben der Charakterisierung von Einzelprozessen, wie z. B. Lithografiesequenzen und trocken- oder naßchemischen Ätzprozessen, kommen auch komplette Integrationsmodule zur Darstellung (Planarisierungskonzepte, Realisierung der Gatestrukturen). Dabei werden stets verschiedene technologische Lösungsansätze vorgestellt und deren Vor- und Nachteile diskutiert, bevor schließlich das, unter den gegebenen Randbedingungen, wirtschaftlichste und technisch geeignetste Verfahren zur Anwendung kommt.

Ein objektiver Leistungsvergleich, der im Rahmen dieser Arbeit hergestellten SiGe-HFETs, steht im Mittelpunkt von Kapitel 7. Hier erfolgt nicht nur eine Gegenüberstellung der Technologiekonzepte hinsichtlich ihrer DC- und HF-Leistungskennlinien, sondern es wird auch ein Algorithmus zur Extraktion der Ersatzschaltbildparameter erarbeitet.

Kapitel 8 schließt diese Arbeit mit einer Zusammenfassung der wichtigsten Ergebnisse ab und gibt einen Ausblick für nachfolgende Aufgabenstellungen zu diesem Thema.

Die dargestellte Gliederung macht die primären Ziele dieser wissenschaftlichen Arbeit deutlich; durch eine vertikale Optimierung des SiGe-Heteroschichtsystems und eine laterale Verbesserung der Transistorstruktur soll das Leistungspotential des SiGe-HFETs für den Hochfrequenzbereich erweitert und maximiert werden. In einem ersten Optimierungsschritt erfolgt eine anwendungsorientierte Konfiguration und Dimensionierung des SiGe-Heteroschichtsystems, nach ausgesuchten Zielgrößen des Kleinsignalersatzschaltbildes (g_m , R_s , R_d , C_{gs} , C_{ds}), die elementaren Einfluß auf das Hochfrequenzverhalten des Transistors haben. Danach soll durch die Modifikation des bestehenden Herstellungsprozesses und durch die Einführung eines komplett neu entwickelten Integrationskonzeptes die Lateralstruktur des SiGe-HFETs nachhaltig verbessert werden. Die Implementierung von selbstjustierenden Prozeßmodulen ermöglicht und realisiert eine fortschreitende Miniaturisierung des Bauelementes bis in den sub-100 nm Bereich hinein und macht gleichzeitig kosten- und zeitintensive Technologien, wie die Elektronenstrahlithografie überflüssig. Außerdem sollen neue Integrationstechniken eine hohe Kompatibilität des Herstellungsprozesses zu Standard-CMOS-Verfahren gewährleisten und die Einbindung des SiGe-HFETs in Schaltungen erleichtern.

Kapitel 2

SiGe – zwischen Si und III/V-Technologie – Kompromiß oder Alternative?

2.1 Mainstream – Silizium-CMOS-Technologie

2.1.1 Die SIA-Roadmap – Taktgeber für die Halbleiterindustrie

Als Gordon E. Moore vor mehr als 35 Jahren das exponentielle Wachstum der Halbleiterkomponenten pro Chip voraussagte, hätte niemand erwartet, daß diese Vision auch heute noch Gültigkeit hat und sich über die Bedeutung einer Voraussage hinaus sogar zur Leitlinie für die Halbleiterindustrie entwickelt [2.1]. Seine damaligen Prognosen stimmen überein mit den Richtlinien der „International Technology Roadmap for Semiconductors“ (ITRS), die im Abstand von zwei Jahren von der Semiconductor Industry Association (SIA) herausgegeben wird. Aufgrund der Globalisierung der Halbleiterindustrie in den letzten zwei Jahrzehnten war es nötig geworden, die Interessen der nationalen Mikroelektronik- und Halbleiterorganisationen zu bündeln und in einer internationalen Roadmap zusammenzufassen. Im Laufe der Jahre ist die Roadmap zu einem anerkannten Taktgeber für die Halbleiterindustrie geworden, der nicht nur das Schrittmaß für die Miniaturisierung der Strukturgrößen vorgibt, sondern auch mögliche Lösungsansätze für technologische Probleme bereitstellt.

Die Tabelle 2.1 zeigt einen Auszug aus der Roadmap von 2001 und bezieht gleichzeitig die erzielten Werte von 1997 und 1999 mit ein [2.2]. Sie beschreibt den aktuellen Stand der Technik für DRAMs, MPUs und ASICs und gibt eine Prognose für die Technologieentwicklung bis ins Jahr 2016 ab. Eine maßgebende Rolle in der Roadmap spielt die Einführung einer neuen Produktgeneration in der

DRAM-Technologie, die einen Produktanteil von über 40 % am Halbleitermarkt des Jahres 2000 hatte (siehe Abbildung 2.1) [2.3], [2.4], [2.5]. Alle vier Jahre steigt hier die Speicherkapazität um das vierfache. Dabei wächst die Chipgröße jedoch nur um das 1,2fache pro Generation, was durch eine entsprechende Reduktion der Speicherzellengröße realisiert wird. Das $\frac{1}{2}$ -Pitch Maß bei den DRAMs, als halber Abstand der Metalleitbahnen in der ersten Verbindungsebene, gilt als repräsentative Größe für den Entwicklungsfortschritt. Hier findet im 2- bzw. 3-Jahresrhythmus eine Skalierung um 70 % statt. Die Kosten pro Bit entwickeln sich umgekehrt proportional zu der Speichergröße und bescheren dem Kunden somit alle vier Jahre eine vierfache Speicherkapazität bei konstantem Preis.

| ITRS 2001 inklusive der Daten für 1999 und 1997 | | | | | | | | | |
|---|------|------|------|------|------|-------|------|-------|-------|
| Year of Production | 1997 | 1999 | 2001 | 2003 | 2005 | 2007 | 2010 | 2013 | 2016 |
| DRAM Production | | | | | | | | | |
| DRAM $\frac{1}{2}$ Pitch [nm] | 250 | 180 | 130 | 100 | 80 | 65 | 45 | 32 | 22 |
| Generation at prod. | 64M | 256M | 512M | 1G | 2G | 4G | 8G | 32G | 64G |
| Gbits/cm ² at prod. | 0,06 | 0,20 | 0,42 | 0,91 | 1,46 | 2,35 | 4,75 | 14,35 | 28,85 |
| DRAM cost/bit (packaged μ cents) | 36 | 15 | 7,7 | 3,8 | 1,9 | 0,96 | 0,34 | 0,12 | 0,042 |
| High-Performance MPU and ASIC Production | | | | | | | | | |
| MPU/ASIC Gate Length [nm] | 200 | 140 | 90 | 65 | 45 | 35 | 25 | 18 | 13 |
| MPU (million tr./cm ²) | 8 | 19,7 | 38,6 | 61,2 | 97,2 | 154,3 | 309 | 617 | 1235 |
| Wiring levels | 6 | 7 | 7 | 8 | 9 | 9 | 10 | 10 | 10 |
| Mask count (max.) | 22 | 24 | 25 | 25 | 25 | 27 | 27 | 29 | 29 |
| MPU cost (μ cents/transistor) | 510 | 245 | 97 | 49 | 24 | 12 | 4,31 | 1,52 | 0,54 |

Tabelle 2.1: ITRS 2001 DRAMs, MPUs und ASICs [2.2]

Ursprünglich galten die DRAM-Produkte als das Zugpferd für die gesamte Halbleiterindustrie, da sich die Logik-ICs (z. B. MPUs) langsamer entwickelten als die Speicherbausteine. Seit Anfang der 90er Jahre beschleunigte sich der Fortschritt auf diesem Gebiet jedoch enorm, so daß heute DRAMs, MPUs und „high Performance“-ASICs gemeinsam die Führungsrolle tragen. Äquivalent zum $\frac{1}{2}$ -Pitch Maß bei den DRAMs ist bei den Mikroprozessoren und ASICs die Gatelänge die entscheidende Größe, an der die Entwicklung gemessen wird. Eine Strukturverkleinerung von 70 % erfolgt hier, wie auch bei den DRAMs, von 1997 bis 2001 in einem 2-Jahresrhythmus und alle drei Jahre in der folgenden Zeit. Die Funktionen pro Chip werden dabei in jedem Technologieabschnitt verdoppelt, während sich die Kosten pro Transistor halbieren. Die Anzahl der Metallisierungs- und Maskenebenen steigt nahezu parallel und ist ein Maß für die Durchlaufzeit und die Komplexität des Herstellungsprozesses.

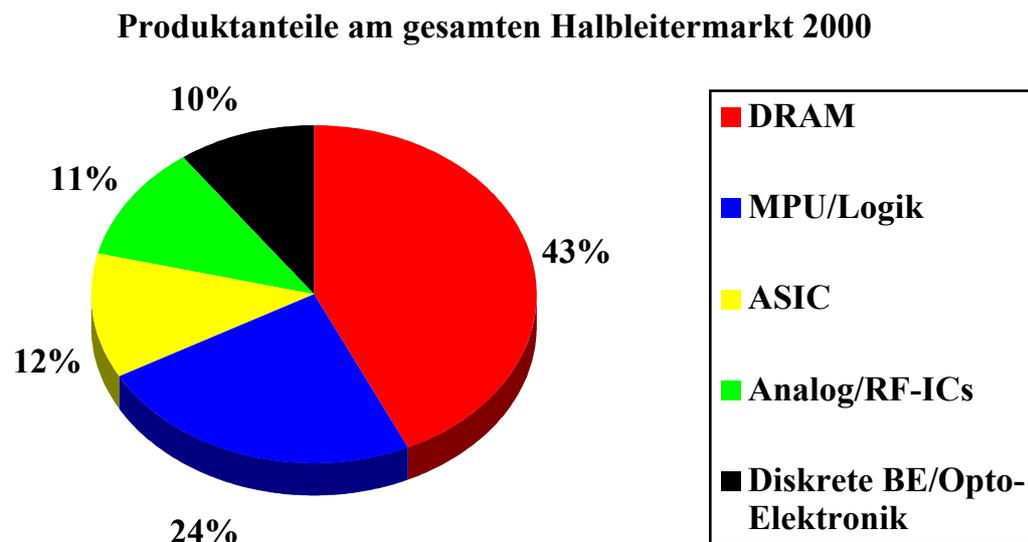


Abbildung 2.1: Produktanteile am gesamten Halbleitermarkt 2000 [2.3]

2.1.2 CMOS – Stand der Technik und Zukunftsperspektive

In Abbildung 2.2 sind die Anteile der CMOS-, Bipolar-, BiCMOS- und GaAs-Technologie am Gesamtumsatz der Halbleiterindustrie in der Welt dargestellt [2.6], [2.7]. Es ist deutlich die dominierende Stellung der CMOS-Technologie zu erkennen, die im Jahr 2000 einen Anteil von 83 % hatte. Demgegenüber stehen 8 % Bipolar-, 7 % BiCMOS- und 2 % GaAs-Produkte.

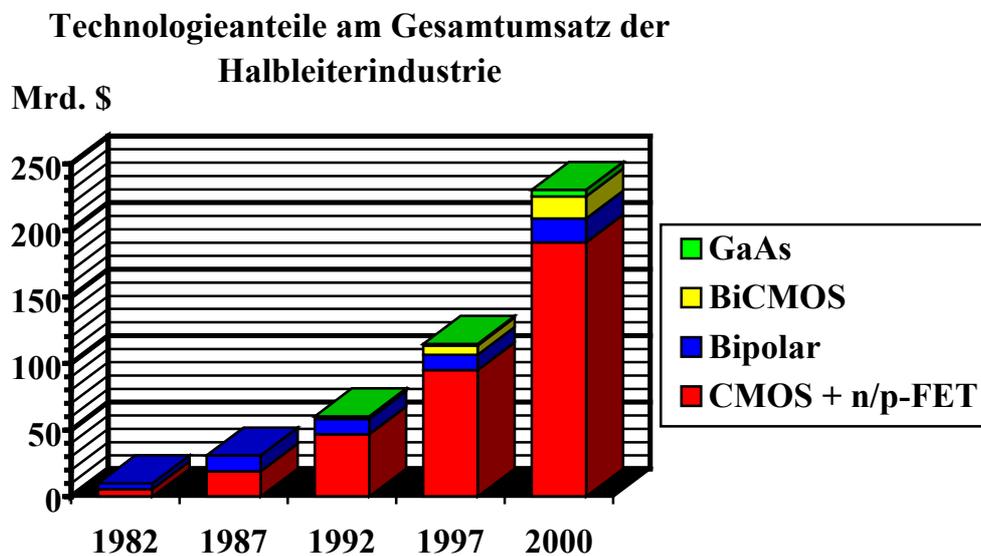


Abbildung 2.2: Technologieanteile am Gesamtumsatz der Halbleiterindustrie
[2.6], [2.8]

Der Erfolg der CMOS-Technologie ist auf zahlreiche Faktoren zurückzuführen:

- gute Skalierbarkeit durch die symmetrische Anordnung der komplementären Transistoren
- hohe Packungsdichte
- extrem niedrige Ruheverlustleistung
- hoher Signalausgangspegel (zwischen 0 V und voller Betriebsspannung)
- hohe Immunität gegen Rauschen und Soft Errors
- unkompliziertes Layout und einfaches Design durch reguläre Schaltungsstrukturen
- Integrationsfähigkeit von analogen und digitalen Schaltungen auf einem Chip
- hohe Technologieausbeute bei niedrigem Preis

Um die charakteristischen Merkmale der CMOS-Technologie zu demonstrieren, soll im folgenden der technologische Ablauf eines 0,25 μm CMOS-Prozesses näher betrachtet werden [2.9]. Es sei an dieser Stelle darauf hingewiesen, daß der beschriebene Herstellungsprozeß beispielhaften Charakter hat und es durchaus unterschiedliche Ausführungsformen der verschiedenen Hersteller gibt. In Abbildung 2.3 ist der schematische Querschnitt durch einen 0,25 μm CMOS-

Gesamtprozeß mit vier Metallisierungsebenen dargestellt [2.10]. Dieser Prozeß läßt sich mit 16 Maskenebenen realisieren und hat eine nominelle Betriebsspannung von 2,5 V. Die Schwellspannung der n- und p-MOSFETs ist symmetrisch auf $\pm 0,5$ V mit einer Variation ΔV_T von ± 50 mV eingestellt. Bei einem maximalen Transistorleckstrom von $1 \text{ nA}/\mu\text{m}$ ergibt sich für den n-MOSFET ein Sättigungsstrom $I_{D\text{sat}}$ von $755 \mu\text{A}/\mu\text{m}$ und für den p-MOSFET von $350 \mu\text{A}/\mu\text{m}$ [2.11].

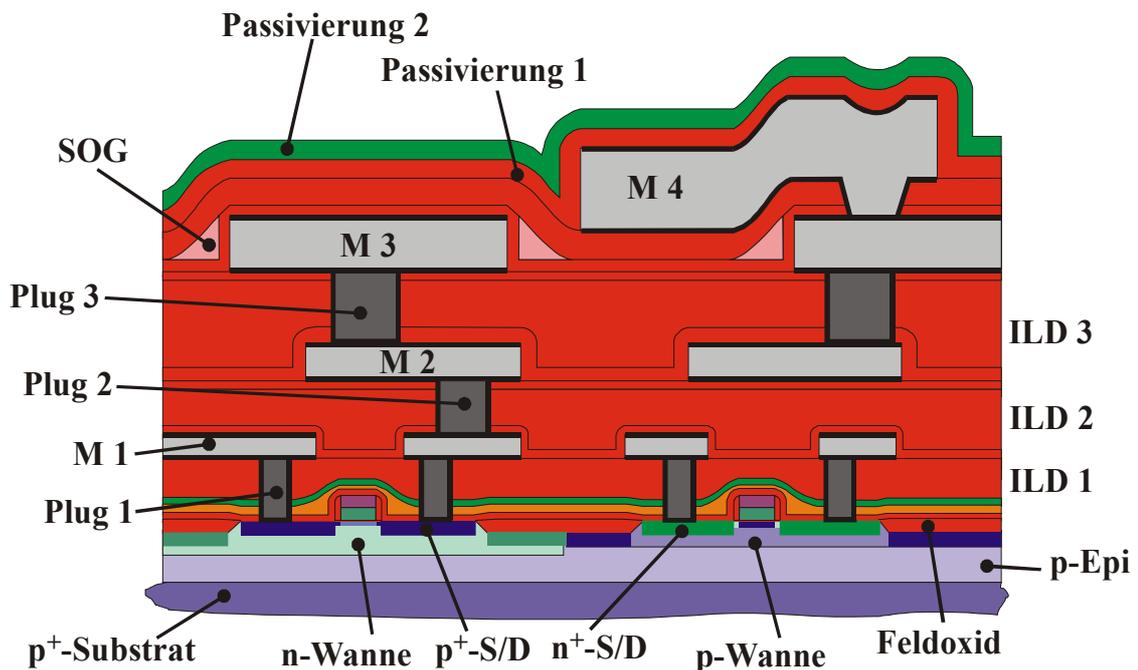


Abbildung 2.3: Schematischer Querschnitt eines $0,25 \mu\text{m}$ CMOS-Gesamtprozesses mit vier Metallisierungsebenen

Der Herstellungsprozeß läßt sich sinnvoll in die folgenden drei Abschnitte teilen: Substratvorbereitung und Bauelementseparation, Transistorintegration und Prozessierung der Metallisierungsebenen.

Substratvorbereitung und Bauelementseparation:

Standardmäßig startet der CMOS-Prozeß auf p^+ -Substrat mit einem spezifischen Widerstand von $0,01 \Omega\text{cm}$, auf dem eine p^- -Epitaxieschicht von $2,0 - 4,0 \mu\text{m}$ Dicke mit einem Widerstand von $10 - 20 \Omega\text{cm}$ abgeschieden wird. Für eine vielseitig einsetzbare Basistechnologie kommt der Twin-Well Prozeß zum Einsatz, bei dem zwei separate, reterograde Wannen unabhängig voneinander durch Io-

nenimplantation gesteuerte Dotierverläufe in die Epi-Schicht eingelassen werden. Durch die retero-grade Dotierung der Wannen sinken die Sperrschichtkapazitäten der pn-Übergänge, ebenso wie der Body-Effekt und die Gefahr eines Latch-up wird verringert [2.12]. Der zusätzliche Einsatz von langsam diffundierenden Dotierstoffen (Arsen für die n-Wanne, Indium für die p-Wanne) unterstützt einen ausgeprägteren retero-graden Verlauf der Dotierkonzentration, was die Kurzkanaleffekte mindert und die Kanalbeweglichkeit der Ladungsträger erhöht [2.13], [2.14].

Die Separation und Isolation der aktiven Transistorgebiete erfolgt mit der klassischen LOCOS-Technik [2.15]. Zum Schutz vor parasitären Kanälen zwischen den Wannen wird zusätzlich eine Channel-Stopper Implantation durchgeführt. Alternativ kann die Wannenisolation auch mit der platzsparenden, aber aufwendigeren Trench Isolation realisiert werden. Damit ist die Substratvorbereitung abgeschlossen und die Prozessierung der aktiven Bauelemente kann beginnen.

Transistorintegration:

Zuerst wird, abgestimmt auf die Gatelänge des Transistors von $0,25\ \mu\text{m}$, ein $4\ \text{nm}$ dickes Gateoxid in trockener O_2 -Atmosphäre bei 900°C aufgewachsen. Direkt danach erfolgt die Abscheidung einer $200\ \text{nm}$ dicken Polysiliziumschicht mit dem LPCVD-Verfahren. Durch die Anwendung der Dual-Gate Technik mit einem n^+ -Polysiliziumgate für den n-Kanal-Transistor und einem p^+ -Polysiliziumgate für den p-Kanal-Transistor kann eine symmetrische Schwellspannung und Dotierung für beide MOSFETs eingestellt werden. Außerdem treten dadurch keine vergrabenen Kanäle auf, was Punch-Through-Ströme verhindert und Kurzkanaleffekte reduziert [2.16].

Die Gatelithografie erfolgt standardmäßig mit DUV-Steppern und einer Lichtwellenlänge von $248\ \text{nm}$, aber auch eine $365\ \text{nm}$ I-Line Belichtung mit Phase-Shift-Masking (PSM) ist denkbar. Nach der Strukturierung des Gates und vor der Source/Drain Implantation wird eine Si-Implantation durchgeführt, die das Substrat und die Polysiliziumoberfläche amorphisiert [2.11]. Alle nachfolgenden Implantationen zeigen dadurch ein abrupteres Dotierprofil ohne den sogenannten Channeling-Tail. Die Gefahr der Bor-Ausdiffusion des p^+ -Gates in das Gateoxid und in den Kanalbereich wird damit ebenfalls eingedämmt. Als nächstes werden unter großem Winkel Halo-Implantationen unter den Gatebereichen eingebracht, die den gleichen Dotierstofftyp wie die Wannen haben. Sie erzeugen dort ein zweidimensionales Dotierprofil und sollen dadurch Kurzkanaleffekte verringern,

die Schwellspannung der Transistoren einstellen und ihr I_{on}/I_{off} -Verhältnis erhöhen [2.17], [2.18]. Anschließend werden die LDD-Gebiete mit einer Eindringtiefe von 50 - 100 nm selbstjustierend zum Polysiliziumgate und dem Feldoxid erzeugt. Sie stellen das Bindeglied zwischen dem Inversionskanal und den tiefen Source/Drain-Implantationen dar und reduzieren die hohe Kanalfeldstärke vor dem Drainbereich. Dadurch verringern sie eine Oxiddegradation durch heiße Elektronen und erhöhen somit die Lebensdauer des Bauelementes. Die laterale Definition der LDD-Gebiete zu den tiefen Source/Drain-Bereichen erfolgt mit Hilfe von SiO_2 -Spacern. Nachfolgend findet die Implantation der Polysiliziumgates und der Source/Drain-Gebiete statt, die etwa 100 - 150 nm in die Wanne reichen [2.9]. Um den Schichtwiderstand der ohmschen Kontakte und der Gates auf etwa 2 - 4 Ω/\square zu reduzieren, wird abschließend ein selbstjustierender $TiSi_2$ -Prozeß durchgeführt mit einer nominellen Ti-Schichtdicke von 20 - 30 nm [2.9], [2.19], [2.20].

Prozessierung der Metallisierungsebenen:

Bei einer CMOS-Technologie für Logikanwendungen mit vier Metallisierungsebenen dienen die Layer M1, M2 und M3 zur Verschaltung der Logikzellen und der statischen RAM-Bereiche. Die Ebene M4 zeichnet sich durch eine dickere Metallisierung aus und verfügt gleichzeitig über ein größeres Pitch-Maß. Sie verbindet die Schaltungsteile mit der Versorgungsspannung und realisiert lange Signalleitungen zwischen den funktionalen Einheiten des Chips. Die Separation der aktiven Transistorgebiete von den Metallisierungsebenen wird mit einer Schichtfolge aus dünnem TEOS-Oxid, BPSG und Si_3N_4 ausgeführt und mit einer dicken TEOS-Schicht abgeschlossen [2.9]. Nach der CMP-Planarisierung werden die Kontaktlöcher zur Verbindung der Transistorebene mit der ersten Metallisierung geätzt und mit Wolfram im CVD-Verfahren aufgefüllt. Als Barrierenmaterial sowohl zwischen den Wolfram-Plugs und dem Oxid als auch zwischen den AlCu-Metallebenen und dem Interlevel-Dielectric (ILD) dient eine etwa 100 nm dicke Ti/TiN-Schichtfolge. Nach der Wolfram-Abscheidung findet mit Hilfe von CMP der Abtrag des überschüssigen Materials statt, um wieder eine planare Oberfläche zu erzeugen.

Die erste Metallisierungsebene ist 500 nm dick und besteht aus gesputtertem oder mit dem CVD-Verfahren abgeschiedenen Al mit einem Anteil von 1 % Cu. Eingebettet in die Ti/TiN-Barrieren wird die strukturierte Metallebene mit einem 500 nm dicken TEOS-Oxid überzogen. Es schließt sich eine CMP-Planarisierung an und die Ätzung der Kontaktlöcher für die zweite Aluminiumebene. Dieser

Ablauf wiederholt sich für die drei ersten Metallisierungslayer, die jeweils eine Dicke von 500, 700 und 900 nm haben. Vor dem Aufbringen der vierten Aluminiumschicht mit einer Dicke von 1800 nm erfolgt die Planarisierung mit Silikat-Spin-On-Glass (SOG). Die Kontaktlöcher werden hier nicht mehr mit Wolfram aufgefüllt, sondern direkt mit dem aufgesputterten Al. Abschließend sorgt eine Schichtfolge aus CVD TEOS-Oxid und Si_3N_4 für die Passivierung des Chips.

| ITRS 2001 inklusive der Daten für 1999 und 1997 | | | | | | | | | |
|---|--------|-------|-------|-------|-------|-------|------|------|------|
| Year of Production | 1997 | 1999 | 2001 | 2003 | 2005 | 2007 | 2010 | 2013 | 2016 |
| Process Integration, Devices, Structures | | | | | | | | | |
| DRAM $\frac{1}{2}$ Pitch [nm] | 250 | 180 | 130 | 100 | 80 | 65 | 45 | 32 | 22 |
| MPU/ASIC Gate Length [nm] | 200 | 140 | 90 | 65 | 45 | 35 | 25 | 18 | 13 |
| Equivalent physical T_{ox} [nm] | 4-5 | 2,5 | 1,5 | 1,4 | 1,2 | 1,0 | 0,7 | 0,5 | 0,45 |
| S/D extension junction depth [nm] | 50-100 | 36-60 | 27-45 | 19-31 | 13-22 | 10-17 | 7-12 | 5-9 | 4-6 |
| Wafer diameter [mm] | 200 | 200 | 200 | 300 | 300 | 300 | 300 | 300 | 450 |
| Mixed-Signal NMOS RF Device | | | | | | | | | |
| f_T [GHz] | 15 | 20 | 132 | 183 | 264 | 372 | 541 | 744 | 1082 |
| f_{max} [GHz] | 20 | 25 | 160 | 170 | 180 | 190 | 215 | 245 | 275 |
| Noise figure [dB] | 1,5 | 1,5 | 1,5 | 1,2 | 1,2 | 1,0 | < 1 | < 1 | < 1 |

Tabelle 2.2: ITRS 2001 Prozeßintegration und Hochfrequenzbauelemente [2.2]

Tabelle 2.2 zeigt einen Auszug der Roadmap des Jahres 2001 für den Bereich Prozeßintegration und stellt außerdem den angestrebten Fortschritt der Hochfrequenzeigenschaften im Mixed-Signal Bereich dar [2.2]. Für die mittel- und langfristige Entwicklung des CMOS-Prozesses spielen unterschiedliche Schlüsseltechnologien eine wichtige Rolle. Diese werden im folgenden Abschnitt näher beleuchtet und sind in Abbildung 2.4 nochmals kompakt zusammengefaßt.

Ein zentrales Thema bei der ständig fortschreitenden Miniaturisierung der Strukturgrößen ist die Lithografie, die allein einen Anteil von mehr als 35 % an den Gesamtherstellungskosten trägt [2.21]. Mit der Einführung der 130 nm-Technologie im Jahre 2001 vollzog sich auch der Übergang zu einer neuen Lithografiegeneration, basierend auf Belichtungssystemen mit einer Wellenlänge von 193 nm. Damit kam zum ersten Mal in der Geschichte der Halbleiterindustrie eine neue Lithografiertechnik zum Einsatz, deren Wellenlänge größer ist als der aktuelle Technologiestatus der SIA-Roadmap. Durch die konsequente Weiterentwicklung sogenannter RET-Verfahren (resolution enhancement techniques) ist es der Halbleiterindustrie in den letzten Jahren gelungen, die Auflösung der Belichtungssysteme permanent zu erweitern. Verbesserte numerische Aperturen und Verfahren wie OAI (off-axis illumination), OPC (optical proximity correction) und PSM (phase-shift masking) erlauben es, Strukturgrößen weit unterhalb der eingesetzten Wellenlänge zu realisieren. Laut Experten reicht die Auflösung der 193 nm-Systeme bis zur 90 nm-Technologie aus und wird dann von der 157 nm F_2 -Technik abgelöst, die sich zur Zeit noch im Entwicklungsstadium befindet. Im Jahr 2010, wenn Strukturgrößen deutlich unter 50 nm erforderlich sind, stößt die optische Lithografie an ihre Grenzen (Abbildung 2.4) [2.22]. Als Alternativen zu den derzeitigen Verfahren sind weiterführende Technologien wie die Röntgenlithografie, die Elektronenprojektionsmethode oder die EUVL in der Diskussion. Bislang gibt es jedoch noch keine Einigung bei der Halbleiterindustrie und den Lithografiezulieferern auf eine bestimmte Technik, so daß die Zukunft ungewiß ist [2.23].

Ein weiterer Problempunkt in der CMOS-Technologie ist die permanente Reduzierung der Oxiddicke T_{ox} . Hier steuert die Entwicklung, getrieben von der Verringerung der Gatelänge, kontinuierlich auf Werte unter 1 nm zu. Da Gateoxid-dicken kleiner als 1,5 nm zu massiven Tunnelströmen führen, die wiederum signifikante Einbußen bei der Langzeitstabilität und der Lebensdauerleistung des Bauelementes nach sich ziehen, können thermische Oxide bei diesen Strukturgrößen nicht mehr eingesetzt werden [2.24]. Um die Steigerung der Transistorleistungsfähigkeit auch in Zukunft voranzutreiben, ist ein Übergang zu neuen Gatedielektrika notwendig. Hier sind Materialien mit hoher Dielektrizitätskonstante (high- k) gefragt (z. B. Ta_2O_5 , Al_2O_3 , ZrO_2 oder TiO_2), die bei gegebener physikalischer Dicke, die elektrische Wirkung einer Oxidschicht mit erheblich geringeren Abmessungen aufweisen [2.21], [2.13]. Dadurch werden die Tunnelströme und Kurzkanaleffekte reduziert und die Steuerwirkung des Gates auf den Kanal erhöht. Leider bringen diese Materialien eine Neuorganisation des gesamten Gatekonzeptes mit sich, weil sie hohe Grenzflächenzustände beim Übergang

zum Silizium verursachen und das Polysiliziumgate durch ein Metallgate ersetzt werden muß. Da es zur Lösung dieser weitreichenden Probleme unterschiedliche Ansätze gibt und sich alle potentiellen Materialalternativen noch im Forschungs- bzw. Entwicklungsstadium befinden, konnte hier noch kein Konsens gefunden werden. Die ITRS setzt deshalb, bis zur geplanten Einführung der high- k Dielektrika im Jahre 2005, auf den Einsatz von Oxinitriden oder reinen Nitridschichten, die aufgrund ihrer höheren Dielektrizitätskonstante eine stärkere Skalierung als Oxid zulassen.

Um den Einfluß der Kurzkanaleffekte bei Transistoren mit ständig geringer werdenden Gatelängen zu minimieren, müssen auch die flachen LDD-Implantationen der Source- und Draingebiete in ihrer Ausdehnung an die neuen Strukturgrößen angepaßt werden. Insbesondere die Implantationstiefe hat einen großen Einfluß auf die Ladungsverteilung unter dem Gate und wird deshalb proportional zur Gatelänge verkleinert. Die Dotierstoffkonzentration muß den veränderten Abmessungen der Implantationsgebiete angeglichen werden, um den Widerstand der Source/Draingebiete niedrig zu halten und die angestrebten Sättigungsströme von $900 \mu\text{A}/\mu\text{m}$ (NMOS) und $450 \mu\text{A}/\mu\text{m}$ (PMOS) zu erreichen [2.25], [2.26]. Die technische Realisierung mit geringeren Implantationsenergien, gut kontrollierbaren RTA-Prozessen, Preamorphisierung des Materials und alternativen, langsamer diffundierenden Dotierstoffen sollte beim heutigen Stand der Technik auf diesem Gebiet kein Problem darstellen. Darüber hinaus spielt ein wohl angepaßtes, in zwei Dimensionen optimiertes Dotierprofil in den Wannen eine entscheidende Rolle bei der Unterdrückung von Kurzkanaleffekten, weil das elektrische Feld und der Potentialverlauf im Kanalbereich dadurch nachhaltig beeinflußt werden. Hier kann eine stark reterograde Wannenimplantation in optimal abgestimmter Kombination mit flachen LDD-Gebieten, Preamorphisierung des Substrats, Halo-Implantationen und einer Kanaldotierung zur Justage der Schwellspannung signifikante Verbesserungen der Transistorcharakteristik bringen [2.13], [2.17], [2.19].

Ein gewaltiger Entwicklungsschritt hat sich in den letzten Jahren im Bereich der Hochfrequenztransistoren vollzogen. Prognostizierte die SIA-Roadmap 1999 noch Grenzfrequenzen f_T und f_{max} im Bereich von 20 GHz, wurden zwei Jahre später bereits die siebenfachen Werte erreicht. Ein maßgeblicher Grund dafür ist der extreme Anstieg der Grenzfrequenzen bei Gatelängen unter 100 nm. So ist es IBM-Forschern gelungen, mit einem 40 nm CMOS-Transistor f_T - und f_{max} -Werte von über 150 GHz zu erzielen [2.27]. Darüber hinaus kommen alternative Materialien und Integrationskonzepte zum Einsatz, die eine deutliche Reduzierung der

parasitären Elemente bewirken und damit eine weitere Leistungssteigerung nach sich ziehen. Eine effektive Methode zur Verringerung der kapazitiven Substratverluste ist der Einsatz der SOI-Technik. Durch die Einführung einer isolierenden Oxidschicht zwischen dem aktiven Transistorbereich und dem Si-Substrat kommt es zu einer signifikanten Reduzierung der Sperrschichtkapazitäten, die von den Source/Drain-Gebieten ausgehen. Daraus resultierend verkürzen sich die Schaltzeiten der Bauelemente, gleichzeitig eliminiert sich der Body-Effekt, was die „Low-Power“ Eigenschaften verbessert. Mit Hilfe der SOI-Technik ist es gelungen, die Leistungsfähigkeit der Transistoren, im Vergleich zu einem Standard CMOS-Prozeß, um 35 % zu steigern [2.28].

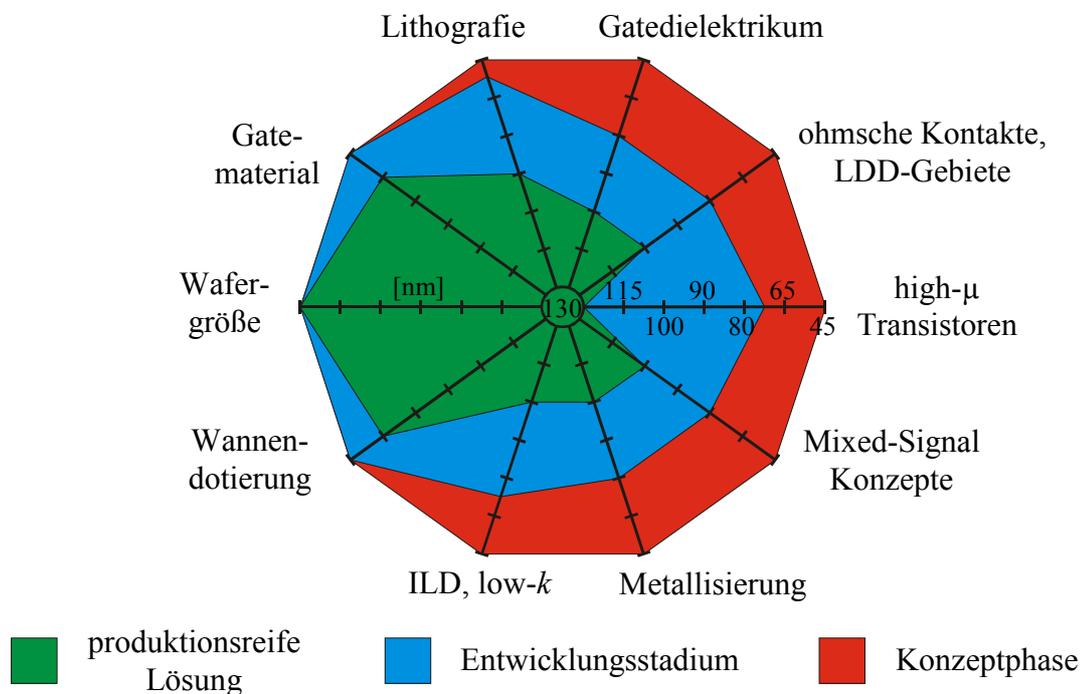


Abbildung 2.4: Status verschiedener Schlüsseltechnologien im Entwicklungsprozess der CMOS-Technologie

Neben den parasitären Kapazitäten spielt der Gatewiderstand eine entscheidende Rolle für das Hochfrequenzverhalten. Aufgrund der beschleunigten Skalierung der Gatelänge in den letzten Jahren reduziert sich über die Querschnittsfläche auch der Widerstand des Gates. Da auch die Dotierstoffkonzentrationen der Polysiliziumgates an ihre Grenzen stoßen und zunehmend Probleme mit der Bor-Ausdiffusion in den Kanal auftreten, ist bis 2007 der Übergang zu Metallgates geplant [2.2]. Für den speziellen Einsatz im Hochfrequenzbereich ist sogar eine T-förmige Metallelektrode mit extrem niedrigen Widerständen denkbar.

Über den Hebel der Gatelänge allein sind jedoch in Zukunft die angestrebten Leistungssteigerungen im Hochfrequenzbereich nicht zu realisieren. Aus diesem Grund werden zur Zeit Konzepte erarbeitet, die eine Implementierung von kristallografisch verspannten Si- und SiGe-Kanälen vorsehen. Die über den Ge-Gehalt einer Pufferschicht steuerbare mechanische Deformation der Gitterkonstanten des Halbleiterkanals resultiert in einer verbesserten Ladungsträgerbeweglichkeit, die bis zu mehrere 100 % betragen kann. Dadurch kommt es zu einer drastisch reduzierten Ladungsträgerlaufzeit zwischen Source und Drain und zu einer Steigerung der Stromtreiberfähigkeit. Der Einsatz dieser neuen Materialkonzepte stellt jedoch erhöhte Ansprüche an die zukünftigen Integrationstechniken. Aufgrund einer verringerten thermischen Stabilität der verspannten Halbleiterschichten und insbesondere auch der metallischen Gateelektroden müssen die derzeitigen Temperaturbudgets der CMOS-Prozesse von bis zu 1000 °C signifikant herabgesetzt werden. Des weiteren bietet sich eine Neuorganisation des Herstellungs Konzeptes an, bei der mit einer Replacement- oder Dummy-Gate Anordnung operiert wird, um den selbstjustierenden Charakter des Prozesses zu erhalten [2.2]. Damit steuert die CMOS-Technologie Zukunftsperspektiven an, die im Rahmen dieser Arbeit am Beispiel des SiGe-HFETs schon realisiert wurden, dies unterstreicht zusätzlich die Aktualität der hier betrachteten Thematik.

Die bisherigen Betrachtungen konzentrierten sich ausschließlich auf die Optimierung der Transistorstruktur, ohne dabei die Metallisierungsebenen mit einzubeziehen. Vergleicht man diese beiden Elemente hinsichtlich ihrer Verzögerungszeiten, wird deutlich, daß bei heutigen Strukturgrößen von 0,13 µm der limitierende Geschwindigkeitsfaktor bei Schaltungen in der Metallisierung liegt (Abbildung 2.5). Eine Verbesserungsmöglichkeit dieses Mißverhältnisses wird zur Zeit durch den Übergang von Aluminium- zu Kupferleitbahnen angestrebt, was eine Widerstandsverringerung von 36 % einbringt. Eine Materialalternative nach Kupfer wird noch nicht in Betracht gezogen.

Im Bereich der ILDs (Interlevel-Dielectrics) gibt es verstärkte Entwicklungsaktivitäten auf dem Gebiet der low- k Isolatoren, deren relative Dielektrizitätskonstante im Bereich von eins bis vier liegt [2.2], [2.21]. Diese können entweder durch CVD-Verfahren aufgebracht werden (SiOF, SiOC, α -CF) oder bieten die Möglichkeit des Aufschleuderns (HSQ, Polymere, Xerogel). Hauptziel ist eine Reduzierung der Schaltgeschwindigkeit durch das Herabsetzen der Dielektrizitätskonstante. Da ein low- k Material aber auch immer eine geringe Wärmeleitfähigkeit mit sich bringt, muß hier ein Mittelweg gefunden werden. Eine duale Lö-

sung mit einem low- k IDL als Isolierung innerhalb einer Metallisierungsebene und einem higher- k IDL zwischen den Metallebenen ist denkbar [2.29].

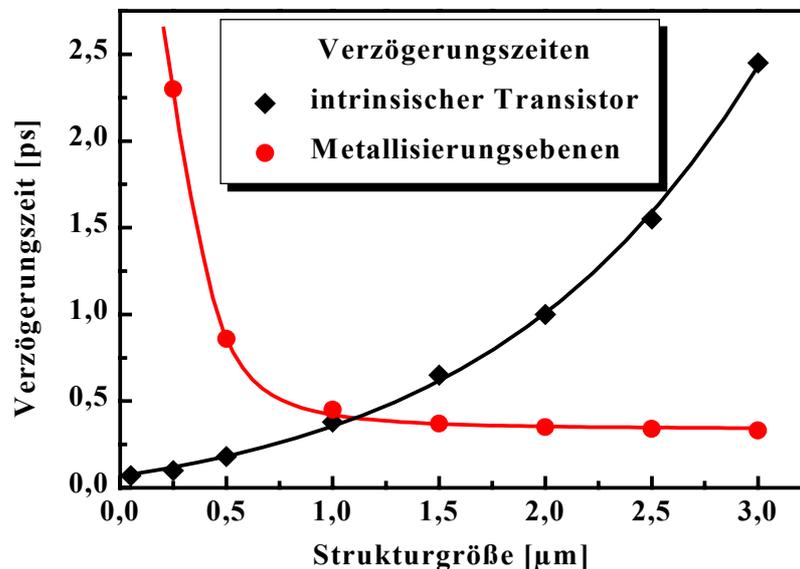


Abbildung 2.5: Vergleich der Verzögerungszeiten des intrinsischen Transistors u. der Metallisierungsebenen in Abhängigkeit der Strukturgröße

Abschließend soll mit dem Waferdurchmesser eine ganz entscheidende Größe der Roadmap diskutiert werden. Um weiterhin eine profitable Herstellung von Halbleiterprodukten zu realisieren und eine Kostenreduktion pro Chipfunktion von 25 % - 30 % im Jahr zu erzielen, hat 2001 der Übergang auf Halbleitersubstrate mit 300 mm Durchmesser begonnen. Im Vergleich zum 200 mm Siliziumsubstrat ergibt sich eine 2,4 mal höhere Chipausbeute pro Wafer. Da die meisten Kostenfaktoren bei der Produktion proportional zur Waferanzahl und weniger zur Wafergröße sind, können bei gleichem Durchsatz pro Zeit und unter Einbeziehung einer Leistungssteigerung der Herstellungsmaschinen die Produktionskosten pro Flächeneinheit der Produktionsstätte (Fab) konstant gehalten werden. Damit sich die enormen Kosten einer neuen Fab von derzeit ca. 2 Milliarden Euro schnell amortisieren, ist es entscheidend, die „Ramp Time“, also die Dauer bis alle neuen Prozesse implementiert sind, so kurz wie möglich zu halten. Gleichzeitig sollte die Produktionsausbeute (Yield) möglichst nah an 100 % liegen, um die hohen Betriebskosten durch einen maximalen Output wirtschaftlich zu halten [2.30].

2.2 Technologie der III/V-Halbleiter

Neben der dominierenden Silizium-Technologie mit einem Anteil von 98 % am Gesamtumsatz der Halbleiterindustrie ist es den III/V-Halbleitern im Jahr 2001 gelungen, mit steigender Tendenz 2 % des Marktes für sich zu gewinnen [2.6]. Insbesondere die Bereiche der Informationstechnik, der mobilen Telekommunikation, der Optoelektronik und der Satelliten- und Automobiltechnologie haben sich als Antriebsmotoren für die steigende Nachfrage herauskristallisiert. Bauelemente und Schaltungen auf der Basis von Galliumarsenid (GaAs) decken dabei den größten Teil des III/V-Marktes ab. Zum Durchbruch verhalf hier der GaAs-MESFET, der sich sowohl durch hohe Grenzfrequenzen und niedrige Rauschzahlen als auch durch große Leistungsdichten auszeichnet. Die Einführung des GaAs-HEMTs sorgte für eine Erweiterung des Frequenzspektrums auf über 100 GHz, bei gleichzeitig verbesserten Rauscheigenschaften [2.31]. Mit dem Einsatz von Indiumphosphid (InP) als Substratmaterial für die HEMTs konnte dieser Trend erneut verstärkt werden, wodurch die Transitfrequenz f_T über 470 GHz stieg und gleichzeitig das Rauschen noch vermindert wurde [2.32].

2.2.1 Materialeigenschaften

Die Gründe für die überragende Stellung der III/V-Halbleiterbauelemente, insbesondere im Bereich Hochfrequenz- und Rauscheigenschaften, liegen in deren exzellenten Transporteigenschaften. Tabelle 2.3 zeigt wichtige physikalische Materialkonstanten für verschiedene III/V-Halbleiterkombinationen und die jeweiligen Vergleichswerte bei Silizium [2.33]. Die auffällige Kombination aus großem Bandabstand, hoher Elektronensättigungs- und Maximalgeschwindigkeit und einer Elektronenbeweglichkeit, die um ein vielfaches größer ist als bei Silizium, prädestinieren diese Materialien für Mikrowellen- und Leistungsanwendungen. Eine weitere besondere Eigenschaft der III/V-Halbleiter ist ihre große Vielfalt an kristallografischen Kombinationsmöglichkeiten. Dabei gibt es kristallengepaßte Kombinationen zwischen Materialien gleicher oder ähnlicher Gitterkonstante und das bewußte Mischen unterschiedlicher Gitterkonstanten, um verspannte Kristallschichten zu erzeugen. Die daraus entstehenden binären, ternären oder quaternären Heterostrukturen bieten ein weites Feld, den Bandabstand des Endmaterials zu beeinflussen und durch „band gap engineering“ die Materialeigenschaften auf die jeweilige Anwendung anzupassen.

| Material-eigenschaft | Si | GaAs | $\text{Al}_x\text{Ga}_{1-x}\text{As}$ | InP | $\text{In}_x\text{Ga}_{1-x}\text{As}$ | $\text{In}_x\text{Al}_{1-x}\text{As}$ | InAs |
|---|------|------|---------------------------------------|------|---------------------------------------|---------------------------------------|-------|
| Bandabstand [eV] (25 °C) | 1,12 | 1,42 | 1,81 | 1,35 | 0,73 | 1,45 | 0,35 |
| Elektronen-Sättigungsgeschwindigkeit [10^7 cm/s] | 1,0 | 0,83 | 0,86 | 1,1 | - | - | - |
| max. Elektronengeschwindigkeit [10^7 cm/s] | 1,0 | 2,2 | - | 3,0 | 2,7 | 6,6 | - |
| Elektronenbeweglichkeit μ_n [cm^2/Vs] | 1500 | 8500 | 3000 | 4500 | 12000 | 1500 | 33000 |
| Löcherbeweglichkeit μ_p [cm^2/Vs] | 450 | 450 | 100 | 150 | 450 | 75 | 400 |
| relative Dielektrizitätskonstante ϵ_r | 11,9 | 12,6 | 11,8 | 12,4 | - | - | 14,6 |
| thermische Leitfähigkeit [W/cmK] | 1,5 | 0,46 | 0,70 | - | - | - | 0,26 |

Tabelle 2.3: Spezifische Eigenschaften von III/V-Halbleitermaterialien und Silizium [2.33]

Abbildung 2.6 gibt den Zusammenhang zwischen dem Bandabstand E_g und der Gitterkonstante a für verschiedene Halbleiter mit direktem und indirektem Bandübergang wieder. Diese Darstellung macht die Variationsvielfalt der Materialien im III/V-Halbleiterbereich deutlich, auf die bei den verschiedenen Bauelementenanwendungen noch eingegangen wird. Nicht minder wichtig ist der semi-isolierende Charakter des GaAs- und InP-Substrates für die Entwicklung zum Marktführer im Hochfrequenzbereich. Zum einen kann dadurch eine einfache Bauelementisolierung realisiert werden, zum anderen dient es gleichzeitig als Dielektrikum für die Leitungen der Mikrowellenschaltkreise.

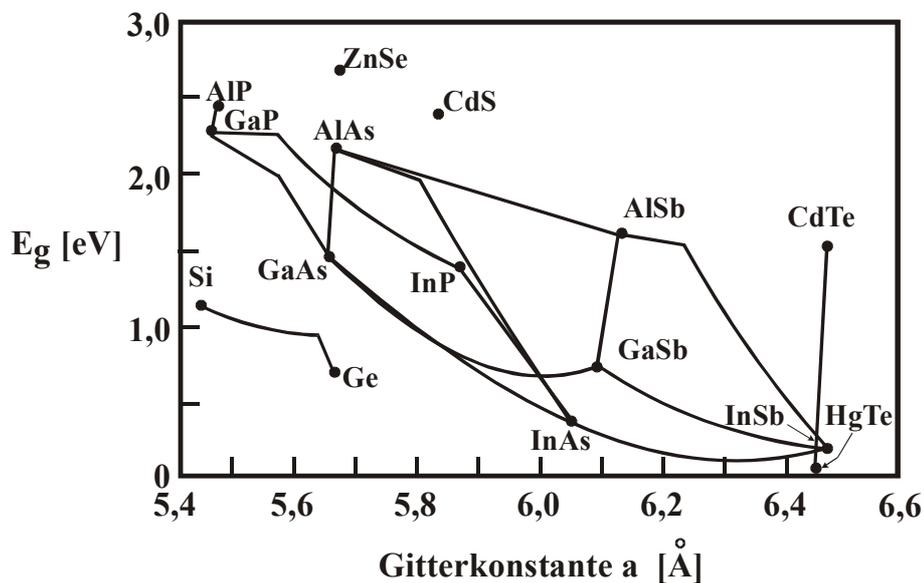


Abbildung 2.6: Zusammenhang zwischen dem Bandabstand E_g und der Gitterkonstanten a für verschiedene Halbleiter

2.2.2 Heterostruktur-Feldeffekttransistoren auf III/V-Basis

Die Erforschung und Weiterentwicklung epitaktischer Wachstumsverfahren wie MBE (molecular beam epitaxy) und MOCVD (metal organic chemical vapor deposition) in den 70er Jahren legte den Grundstein für eine neue Bauelementgeneration auf der Basis von Halbleitermaterialübergängen mit unterschiedlichen Gitterkonstanten und Bandabständen. Im Jahre 1979 konnten Störmer und Dingle an einem selektiv dotierten AlGaAs/GaAs Heteroübergang erstmals eine erhöhte Beweglichkeit der Ladungsträger nachweisen [2.34]. Bereits ein Jahr später gelang es Mimura und Delagebeaudeuf, das Prinzip auf einen Feldeffekttransistor zu übertragen [2.35], [2.36]. Der Grundgedanke dabei ist, einen Heteroübergang zwischen einem Halbleiter mit kleiner Bandlücke und einem Halbleiter mit größerem Bandabstand zu erzeugen, um an der Grenzschicht einen Potentialtopf zu bilden. Bei einer selektiven Dotierung (Modulationsdotierung) des Halbleitermaterials mit dem großen E_g diffundieren die freien Ladungsträger der ionisierten Donatoren bzw. Akzeptoren in den energetisch günstigeren Potentialtopf und bilden ein zweidimensionales Ladungsträgergas (2DEG bzw. 2DHG). Aufgrund der räumlichen Trennung zwischen den ionisierten Atomrümpfen der Dotieratome und den Ladungsträgern im Quantum Well erfolgt der Stromtransport mit

erhöhter Beweglichkeit μ . Da bei den bewährten III/V-Halbleiterkombinationen die Elektronenbeweglichkeit μ_n meistens mehr als eine Zehnerpotenz größer ist als die Löcherbeweglichkeit μ_p , findet ausschließlich eine Bauelementkonfiguration mit Potentialtopf im Leitungsband und 2DEG Anwendung. Eine weitere Steigerung der Beweglichkeit und insbesondere verbesserte Rauscheigenschaften können erzielt werden, wenn der Halbleiter mit der geringeren Bandlücke eingebettet wird zwischen Halbleitern mit höherem Bandabstand und größerer Gitterkonstante. Der eingeschlossene Kanal, in dem sich das Elektronengas bildet, wird dabei zusätzlich kristallografisch verspannt, was verbesserte elektrische Eigenschaften mit sich bringt. Für die beschriebenen Bauelemente gibt es eine Vielzahl von Namen. Durchgesetzt haben sich im wesentlichen die Bezeichnungen HEMT (High Electron Mobility Transistor), HFET (Heterostructure Field-Effect Transistor) und MODFET (MODulation Doped Field-Effect Transistor).

Der klassische Schichtaufbau eines GaAs-HEMTs startet mit einem semiisolierenden GaAs-Substrat, auf das eine undotierte GaAs-Schicht epitaktisch aufgebracht wird (Abbildung 2.7). Am Heteroübergang folgt zuerst ein 2 - 5 nm dünner intrinsischer AlGaAs-Spacer, der den Potentialtopf räumlich von der Silizium dotierten AlGaAs-Schicht trennen soll. Dadurch ergibt sich eine verminderte Coulomb-Wechselwirkung zwischen den ionisierten Donatoratomen und dem 2DEG und damit eine erhöhte Beweglichkeit [2.37]. Mit einer Dotierung von $1 - 2 \cdot 10^{18} \text{ cm}^{-3}$ bei einer Dicke der Dotierschicht von bis zu 40 nm kann eine Ladungsträgerkonzentration n_s im Kanal von bis zu $5 \cdot 10^{12} \text{ cm}^{-2}$ erzielt werden [2.38], [2.39]. Abschließend wird die Struktur mit einer n^+ -GaAs-Schicht abgeschlossen. Sie hat die Aufgabe, die ohmschen Kontakte zu verbessern und damit die parasitären Widerstände an Source und Drain klein zu halten. Die Vorteile der HEMT-Struktur gegenüber dem klassischen MESFET spiegeln sich in einer erhöhten Ladungsträgerbeweglichkeit von bis zu $8000 \text{ cm}^2/\text{Vs}$ und einer verzehnfachten Ladungsträgerdichte im Kanal wider. Daraus resultieren Grenzfrequenzen deutlich über 100 GHz und minimale Rauschzahlen, die 40 % unter denen des MESFETs liegen [2.40].

Um die elektrischen Eigenschaften des GaAs-HEMTs weiter zu steigern, stellten Rosenberg et al. 1985 einen Transistor mit InGaAs-Kanal vor [2.41]. InGaAs zeichnet sich im Vergleich zu GaAs durch eine erhöhte Elektronenbeweglichkeit bei kleinerem Bandabstand aus. Dadurch kommt es zur Ausbildung eines ausgeprägteren Potentialtopfes bei gleichzeitig erhöhtem μ_n im 2DEG. Da die Gitterkonstante ebenfalls mit dem Indiumanteil ansteigt, kann dieser nur solange gesteigert werden, bis die kritische Schichtdicke für den Kanal erreicht ist. Im Sy-

stem $\text{Al}_x\text{Ga}_{1-x}\text{As}/\text{In}_y\text{Ga}_{1-y}\text{As}$ auf GaAs-Substrat kann daher bei einer Kanaldicke von 10 nm der Indiumanteil y maximal 25 % betragen [2.42]. Die sonstige Schichtkonfiguration stimmt mit der des klassischen GaAs-HEMTs überein. Durch die bessere Ladungsträgerlokalisierung im 2DEG und die Anwendung eines verspannten InGaAs-Kanals erzielt der pseudomorphe GaAs-HEMT maximale Grenzfrequenzen f_{max} bei der Leistungsverstärkung von bis zu 350 GHz und unterbietet die minimalen Rauschzahlen des konventionellen GaAs-HEMTs [2.43]. Er ist dadurch der beste GaAs-basierte Transistor, der für diese Bereiche verfügbar ist.

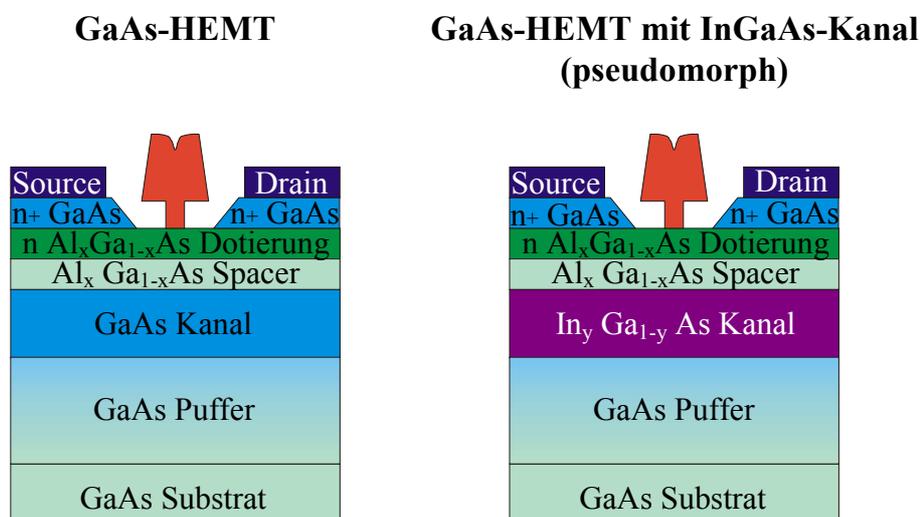


Abbildung 2.7: Schematischer Querschnitt durch die Schichtstruktur eines konventionellen GaAs-HEMTs und eines pseudomorphen HEMTs mit InGaAs-Kanal

Um den Indiumanteil im Kanal und die daraus erwachsenen Vorteile für das Bauelement weiter zu steigern, ist es notwendig, das Substratmaterial zu wechseln und auf Indiumphosphid überzugehen. Als epitaktische Pufferschicht wird dabei intrinsisches $\text{In}_y\text{Al}_{1-y}\text{As}$ verwendet, was bei einem Anteil von $y = 52$ % gitterangepaßtes Wachstum sicherstellt, aber gleichzeitig einen höheren Bandabstand hat. Im System $\text{In}_x\text{Ga}_{1-x}\text{As}/\text{In}_{0.52}\text{Al}_{0.48}\text{As}$ findet für $x = 53$ % ein spannungsfreier Übergang statt, so daß ein komplett gitterangepaßtes Schichtsystem entsteht. Für höhere In-Anteile bis ca. 80 % werden pseudomorphe Kanäle realisiert [2.44]. Die intrinsischen InAlAs-Spacer haben je nach Anwendung des Bauelementes eine Dicke von 3 - 8 nm, und die Dotierstoffkonzentration der nachfolgenden Dotierschicht liegt im Bereich von $2 - 8 \cdot 10^{18} \text{ cm}^{-3}$. Vor dem Ab-

schluß der Schichtstruktur mit einer hochdotierten n^+ -InGaAs-Deckschicht kann noch eine dünne InP-Stopschicht für den Gate-Recess eingeführt werden.

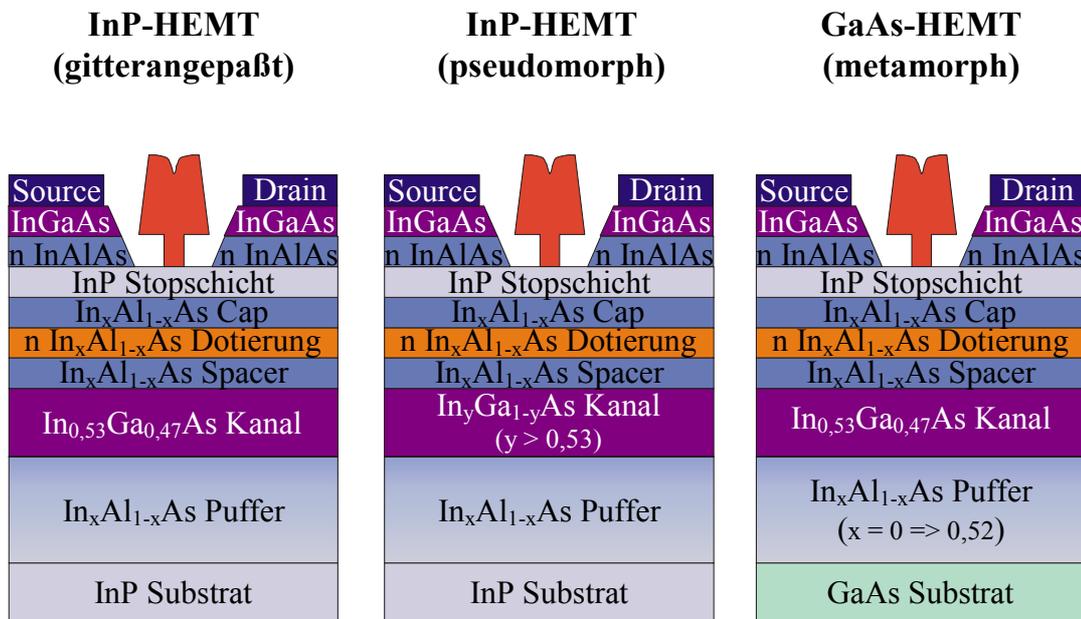


Abbildung 2.8: Schematischer Querschnitt durch einen gitterangepaßten und einen pseudomorphen InP-HEMT und einen metamorphen HEMT auf GaAs Substrat mit InGaAs-Kanal

Durch ihre überragenden Grenzfrequenzen von bis zu 600 GHz f_{max} [2.45] und 472 GHz f_T [2.32] und ihre niedrigen Rauschzahlen, die bei 60 GHz noch unter 1 dB liegen [2.46], haben die gitterangepaßten und pseudomorphen InP-HEMTs ihre Ausnahmestellung bei Mikrowellenbauelementen deutlich gemacht. Im Vergleich zum GaAs-Substrat sind InP-Wafer jedoch sechsmal teurer, weniger robust in der Handhabung, noch nicht in großen Durchmessern als Substrat verfügbar und weisen eine weniger ausgereifte Prozeßtechnologie auf [2.47], [2.48]. Deshalb gibt es in den letzten Jahren einen starken Trend, sogenannte metamorphe HEMTs mit InGaAs-Kanälen auf normalen GaAs-Substraten zu wachsen. Dazu wird ein gradiertes Puffer aus InAlAs oder InAlGaAs auf das Substrat epitaktisch aufgebracht, dessen Indiumanteil linear von 0 % auf die gewünschte Komposition ansteigt, damit anschließend der Kanal gitterangepaßt realisiert werden kann. Die Auslegung der Spacer, der Dotierschichten und der hochdotierten Deckschichten erfolgt analog zum HEMT auf InP-Substrat.

Die Prozeßtechnologie zur Herstellung der HEMTs ist prinzipiell für alle Transistorvariationen gleich. Zuerst müssen die verschiedenen Bauelemente, die auf einem Wafer gefertigt werden sollen, voneinander getrennt werden, da sie bei gleicher Schichtstruktur alle über den 2DEG-Kanal miteinander verbunden sind. Dazu bieten sich die Verfahren der Isolation durch Mesaätzung und die Ionenimplantation an. Bei der Mesaätzung wird der spätere aktive Transistorbereich zuerst in einem Lithografieprozeß maskiert und anschließend das Halbleitermaterial zwischen den Bauelementen trocken- oder naßchemisch entfernt. Wichtig dabei ist, daß die niederohmigen Schichten bis in den semiisolierenden Bereich abgetragen werden, um Leckströme so gering wie möglich zu halten. Bei der Ionenimplantation erfolgt kein Materialabtrag, sondern eine bewußte Degradierung der Leitfähigkeit durch Ionen- bzw. Protonenbeschuß. Die ursprüngliche Kristallstruktur wird zerstört und es bilden sich Kristalldefekte und teilweise Reaktionskomplexe zwischen dem Implantationsstoff und dem Halbleitermaterial [2.49]. Die Vorteile der Implantation sind die gezielte, punktgenaue Erzeugung hochohmiger Halbleiterbereiche und damit die Realisierung höherer Packungsdichten und die Erhaltung der planaren Materialoberfläche, was für hochintegrierte Schaltungen mit mehreren Metallisierungsebenen sehr wichtig ist.

Nach der Bauelementisolation werden die ohmschen Kontakte mit Hilfe der Lift-off Technik erzeugt. Für GaAs- und InP-basierte HEMTs kommen Materialkombinationen aus GeAu/Ni/Au zum Einsatz, die in RTA-Systemen bei 350 °C bis 400 °C legiert werden und dabei in Kombination mit einer hochdotierten Deckschicht standardmäßig Kontaktwiderstände kleiner als 0,1 Ω mm erzielen [2.48].

Zur Realisierung kurzer Gatelängen und kleiner Gatewiderstände steht heute standardmäßig die Elektronenstrahlithografie zur Verfügung. Mit Hilfe von Mehrlagenlacksystemen und der Lift-off-Technik wurden bereits T-förmige Gates mit 50 nm Länge umgesetzt, auch wenn der Produktionsstandard für Hochfrequenztransistoren derzeit bei etwa 100 nm liegt. Als Materialkombination zur Erzeugung eines reproduzierbaren und thermisch stabilen Schottkyüberganges wird bei GaAs- und InP-Transistoren Ti/Pt/Au gewählt. Vor dem Aufdampfen des Kontaktmaterials müssen zuerst die Oberflächenpassivierung, meistens bestehend aus Siliziumnitrid, und die hochdotierte Deckschicht für die ohmschen Kontakte unter dem Gate entfernt werden. Epitaktisch eingefügte Stopschichten kontrollieren den Ätzprozeß und definieren gleichzeitig den Abstand zwischen Gate und Kanal und damit einen entscheidenden Parameter zur Beeinflussung der Transistorcharakteristik. Abschließend wird die gesamte Bauelementstruktur mit einer Isolationsschicht passiviert.

Tabelle 2.4 gibt einen Überblick über die Leistungskennndaten und Anwendungsgebiete von GaAs und InP basierten HEMTs. Dabei zeigen die Werte in Klammern an, bei welcher Gatelänge bzw. bei welcher Frequenz die Werte ermittelt wurden.

| Bauelement- typ | g_m [mS/mm] (Gatelänge) | f_T [GHz] f_{max} [GHz] F_{min} [dB] | Leistungs- dichte [W/mm] | Anwendungs- gebiete |
|----------------------------------|---------------------------------|--|---------------------------------------|--|
| GaAs-HEMT | 600 (250 nm) [2.40] | 113 (100 nm) 190 (250 nm) 1,8 (60 GHz) [2.50, 51, 40] | 0,41 (60 GHz, 250 nm) [2.51] | mobile Kommuni- kationssysteme, Mikrowellenan- wendungen mit mittlerem Lei- stungsbedarf, opti- sche Übertragungs- systeme |
| GaAs-HEMT mit InGaAs-Kanal | 930 (80 nm) [2.52] | 150 (80 nm) 350 (80 nm) 1,6 (60 GHz) [2.53, 43, 54] | 0,84 (60 GHz, 150 nm) [2.54] | |
| InP-HEMT MM | 1100 (180 nm) [2.48] | 204 (180 nm) 300 (150nm) 1,1 (20 GHz) [2.55] | - | |
| InP-HEMT LM | 2050 (30 nm) [2.56] | 472 (30 nm) - 0,9 (60 GHz) [2.32],[2.54] | - | optische Kommuni- kationssysteme, Satellitenanwen- dungen, Radar- und Sensorsysteme, Mikrowellenan- wendungen für mittlere Leistungen |
| InP-HEMT PM | 1400 (100 nm) [2.57] | 340 (50 nm) 600 (100 nm) 0,7 (60 GHz) [2.58, 45, 59] | 0,5 (60 GHz) [2.60] | |

Tabelle 2.4: Leistungskennndaten und Anwendungsgebiete von III/V-Heterostruktur-Feldeffekttransistoren

2.3 SiGe-Technologie – das Hochfrequenzpotential der Siliziumbranche

2.3.1 Entwicklung der SiGe-Technologie und Stand der Technik

Die junge Geschichte der SiGe-Technologie ist stark geprägt worden von dem kommerziellen Durchbruch des SiGe-HBTs und der daraus resultierenden SiGe-BiCMOS-Technologie. In der Tat kommt dem SiGe-HBT eine Pionierrolle in der Evolution der SiGe-Technologie zu. Denn nur wenige Jahre nach seiner ersten technischen Realisierung durch S. S. Iyer [2.61] kam es bereits 1996, dank massiver industrieller Entwicklungsaktivitäten (IBM, DaimlerChrysler/TEMIC, NEC, Hitachi), zur kommerziellen Einführung des SiGe-HBTs [2.62].

Bereits Ende der 80er Jahre stößt der Si-Bipolartransistor (Si-BJT), der aufgrund seiner guten Hochfrequenzeigenschaften als Zugpferd für die analoge Schaltungstechnologie gilt, an seine physikalischen Leistungsgrenzen. Um die Grenzfrequenzen f_T und f_{max} und die Rauscheigenschaften des Transistors weiter zu verbessern, ist sowohl eine Verkürzung der Basisweite als auch eine Erhöhung der Dotierung in der Basis notwendig. In Kombination mit einer aus Gründen der Transistoreffizienz ebenfalls hohen Dotierstoffkonzentration im Emitter führt diese Entwicklung zu einem großen elektrischen Feld am Emitter-Basis-Übergang, einem schwer kontrollierbaren Dotierprofil und einer Degradation des Bauelementes durch Tunnelströme [2.63].

Durch die Einführung des SiGe-HBTs gewinnt das Transistordesign zusätzliche Freiheitsgrade bei der vertikalen Optimierung des Bauelementes und ermöglicht so eine Leistungssteigerung weit über das Niveau des Si-Bipolartransistors hinaus. Erst die Weiterentwicklung kristalliner Wachstumsverfahren, wie der Molekularstrahlepitaxie (MBE) und der Chemical Vapor Deposition (CVD) in den 70er und 80er Jahren, legte den Grundstein für ein atomares Schichtwachstum mit hoher Kristallqualität und damit für die Erfolgsgeschichte des HBTs und der SiGe-Technologie. Der kontrollierte epitaktische Einbau von Germanium in die Basisschicht des Transistors und die präzise Implementierung beliebiger Dotierstoffverteilungen erlauben ein gezieltes „band-gap-engineering“ mit exakt ausgerichteten Heteroübergängen zum Emitter und zum Kollektor und die Realisierung scharfer Dotierprofile. Der Ge-Gehalt in der Basis liegt, abhängig vom Herstellungsverfahren und der Auslegung des Transistors, zwischen 10 % und 30 % [2.64], [2.65]. Ein gradiertes Ge-Verlauf mit linearem Konzentrationsanstieg

vom Emitter zum Kollektor führt dabei zu den entscheidenden physikalischen Vorteilen des SiGe-HBTs. Durch die mit anwachsendem Ge-Gehalt abnehmende Bandlücke E_G kann die Injektion von Minoritätsträgern aus dem Emitter in die Basis erhöht werden. Gleichzeitig reduziert sich durch einen zunehmenden Valenzbandsprung ΔE_V die Rückinjektion von Minoritätsträgern in den Emitter, was zu einem drastischen Anstieg der Stromverstärkung β führt. Die abnehmende Bandlücke in der Basis prägt ihr gleichzeitig ein permanentes elektrisches Feld ein, was eine Beschleunigung der Minoritätsträger in Richtung Kollektor zur Folge hat. Dadurch kann zusätzlich die Basislaufzeit der Ladungsträger nahezu halbiert werden [2.63]. Aufgrund der erhöhten Valenzbanddifferenz ΔE_V des Emitter-Basis-Überganges ist es außerdem möglich, den Grad der Emitterdotierung zu reduzieren und dadurch, trotz einer zunehmenden Skalierung der Basis, die Durchbruchspannung zu steigern und parasitäre Kapazitäten zu senken.

Die oben beschriebenen Maßnahmen im vertikalen Schichtaufbau des Bauelementes resultieren in einem signifikanten Leistungssprung des SiGe-HBTs im Vergleich zum Si-BJT. Insbesondere bei den Grenzfrequenzen, die wichtige Indikatoren für des Hochfrequenzpotential des Transistors darstellen, können durch eine gezielte Skalierung der Basisweite bei gleichzeitiger Erhöhung der Dotierstoffkonzentration große Fortschritte verzeichnet werden (Abbildung 2.9).

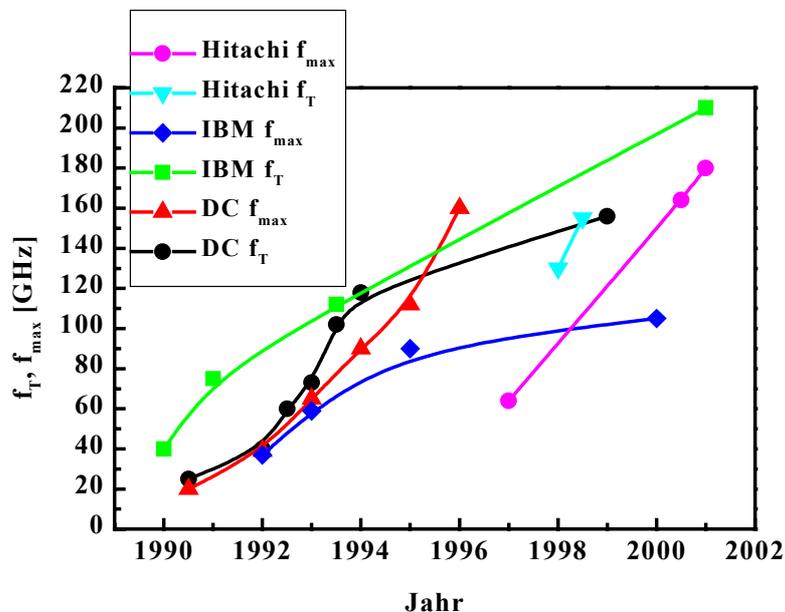


Abbildung 2.9: Entwicklung der Grenzfrequenzen f_T und f_{max} verschiedener Hersteller für den SiGe-HBT [2.66]

Für den Einsatz in Low Noise Amplifiers (LNAs) oder anderen sensitiven Anwendungen im Telekommunikationsbereich ist die Güte des Rauschverhaltens entscheidend. Mit niedrigen minimalen Rauschzahlen F_{\min} von 0,4 dB bei 5 GHz und 0,9 dB bei 10 GHz und einer Eckfrequenz f_c von unter 300 Hz für das 1/f-Rauschen unterstreicht der SiGe-HBT zusätzlich sein Hochfrequenzpotential [2.67]. Gute Ergebnisse können auch auf dem stetig wachsenden Marktsegment der analogen „low-power“-Anwendungen erzielt werden. Durch die Leistungsverbesserungen im Hochfrequenzbereich erreicht der SiGe-HBT ein 4- bis 5-mal höheres f_T/I_C -Verhältnis als der Si-BJT (Abbildung 2.10), was insbesondere im Bereich der mobilen Kommunikation von großer Bedeutung ist [2.68].

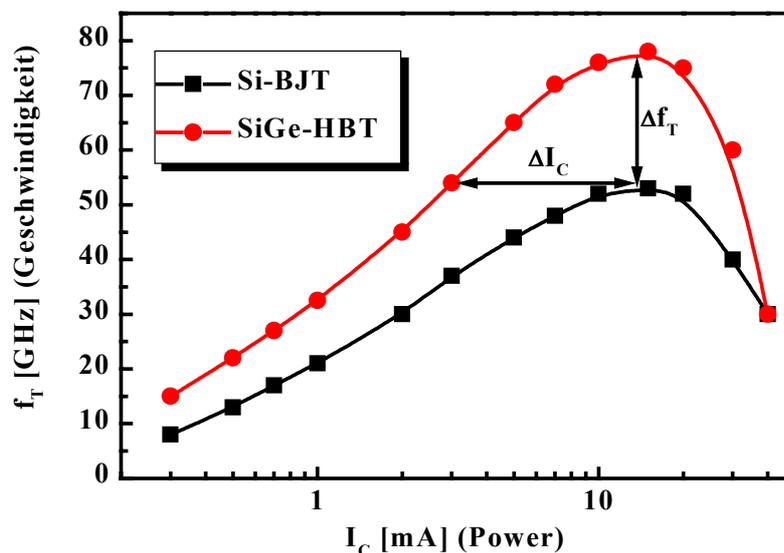


Abbildung 2.10: Das Leistungspotential des HBTs im Hochfrequenzbereich kann optional in „low-power“-Potential transferiert werden

Große Leistungsreserven im Hochfrequenzbereich, niedrige Rauschkenndaten, eine exzellente Linearität und eine „power-added efficiency“ (PAE) von bis zu 70 % [2.69] eröffnen dem SiGe-HBT schnell einen breiten Einsatzbereich in der Telekommunikationsbranche und bei leistungsstarken Netzwerkanwendungen. Ein entscheidender Vorteil, der dem SiGe-HBT zum endgültigen Durchbruch verhalf, ist seine große Kompatibilität zur Si-CMOS-Technologie [2.63]. Durch die Integration des SiGe-HBTs in einen Standard CMOS-Prozeß entsteht die SiGe-BiCMOS Technologie, die dabei die charakteristischen Merkmale beider Konzepte miteinander vereinigt. Dadurch wird es möglich, analoge Funktionen

und Hochfrequenzkomponenten mit digitalen Logik- bzw. Speicherelementen zu koppeln und auf einem Chip hochkompakt zu integrieren. Vorreiter auf dem Gebiet der SiGe-BiCMOS-Technologie ist die Firma IBM, die bereits 1998, zwei Jahre nach der kommerziellen Einführung des SiGe-HBTs, die erste SiGe-BiCMOS-Generation präsentierte [2.70] (Tabelle 2.5).

| BiCMOS Technologie | 1. Generation 1998 | 2. Generation 1999 | 3. Generation 2001 |
|--|-----------------------|-----------------------|-----------------------|
| CMOS L_G [μm] | 0,5 | 0,25 | 0,18 |
| f_T (HBT) [GHz] (high BV/high RF) | 29/50 | 29/50 | 30/120 |
| f_{max} (HBT) [GHz] | 50/60 | 50/60 | 50/100 |
| F_{min} [dB] | 0,8 | 0,8 | 0,4 |
| Breakdown Voltage [V] BV_{CEO} (HBT) | 5,5/3,3 | 5,5/3,3 | 5,0/2,1 |
| CMOS Gate-Delay [ps] | 90 | 50 | 33 |
| Technologie | base during gate | base after gate | base after gate |
| Metallisierung | Al | Al | Cu |

Tabelle 2.5: Zeitliche Entwicklung und Stand der Technik bei der SiGe-BiCMOS-Technologie (IBM) [2.71]

Das Integrationskonzept der SiGe-BiCMOS-Technologie basiert auf einem Standard CMOS-Prozeß, der üblicherweise für Logikanwendungen eingesetzt wird. Die Trennung der analogen und digitalen Schaltungskomponenten erfolgt mit Hilfe der „deep-trench isolation“ (DTI), um eine möglichst effektive und kompakte Isolation der unterschiedlichen Elemente zu erzielen. Der SiGe-HBT Prozeß hat aufgrund der epitaktisch abgeschiedenen, pseudomorphen Basisschicht ein strenger limitiertes Temperaturbudget als das CMOS-Verfahren. Aus diesem Grund verfolgt das Integrationskonzept eine gezielte Modularisierung der kritischen Einzelkomponenten. Danach bildet die Front-End-Prozessierung der CMOS-Bestandteile das erste abgeschlossene Modul. Anschließend erfolgt die

Realisierung der temperatursensitiven SiGe-HBT-Komponenten. Dieses „base after gate“-Verfahren führt zu einer technologischen Entkopplung der beiden Bauelemente und macht so den HBT-Prozess unabhängig von Änderungen in der CMOS-Technologie [2.72]. Die Umsetzung von Schaltungskomponenten für den Hochfrequenzbereich erfordert zusätzlich die Integration passiver Elemente wie Kapazitäten, Induktivitäten und Widerstände. Für die Realisierung qualitativ hochwertiger Spulen sind extrem dicke Metallisierungslayer notwendig, die ein Standard CMOS-Prozess nicht bereitstellt. Deshalb wird bei der SiGe BiCMOS-Technologie eine zusätzliche mehrere μm dicke Metallebene eingeführt. Trotz dieser Modifikationen hat der Gesamtprozess mehr als 90 % Kompatibilität zum Standard CMOS-Verfahren und erzielt dabei eine Ausbeute, die ebenfalls deutlich über 90 % liegt [2.63].

| Marktsegment / Frequenzband | Anwendungsbereich | Schaltung/ IC |
|--|--|---|
| Mobile Kommunikation, Basis-Stationen 1 - 2 GHz | GSM, DECT, AMPS, Personal Communication Services (PCS), GPS, DGPS | Low Noise Amplifier (LNA), Transceiver, Receiver, Voltage Controlled Oscillator (VCO), Mischer, Automatic Gain Control Amplifier (AGC), Frequenz-Synthesizer, Digital Network Switch (DNS), Power Amplifier (PA), Digital-Analog/ Analog-Digital-Wandler, Multiplexer (MUX), Demultiplexer (DMUX), GPS-Chipsätze, ASICs |
| Drahtlose Datenübertragung 2 - 6 GHz | Bluetooth, Dedicated Short Range Communication (DSRC), Electronic Toll Collection (ETC), Wireless LAN, Wireless Local Loop (WLL) | |
| Optische Datenübertragung 10 - 40 GBit/s | Synchronous Optical Network (SONET) | |

Tabelle 2.6: Übersicht der Marktsegmente, Frequenzbereiche und Schaltungsanwendungen für die SiGe-HBT- und BiCMOS-Technologie

Eine Zusammenstellung der derzeitig verfügbaren Schaltungsanwendungen auf Basis der SiGe-BiCMOS-Technologie zeigt Tabelle 2.6. Dabei ergibt sich ein deutlicher Schwerpunkt im Marktsegment der Telekommunikationsbranche mit einem angestrebten Frequenzbereich von 2 - 30 GHz. Ziel ist es dabei, den

Marktanteil der SiGe-Technologie im Bereich der Hochfrequenz-ICs bis ins Jahr 2005 auf 10 % zu steigern, was einem Umsatz von 1,8 Milliarden Dollar entspricht [2.68].

2.3.2 Das Leistungspotential des SiGe-HFETs

Die enormen Leistungssteigerungen beim Si-Bipolartransistor durch die Einführung des Germaniums und die hohe Kompatibilität der SiGe-Technologie zum Standard-CMOS-Prozeß sind die Hauptmotivationsfaktoren für eine entsprechende Entwicklung beim SiGe-Hetero-Feldeffekttransistor. Aufgrund der spezifischen physikalischen Eigenschaften des Si/SiGe-Materialsystems lassen sich in einer gitterangepaßten, zugverspannten Si-Schicht Elektronenbeweglichkeiten erzielen, die deutlich über dem Maximalwert des Bulk-Materials von $1500 \text{ cm}^2/\text{Vs}$ (300 K) liegen. Durch die zusätzliche Einbettung dieser Si-Schicht in einen SiGe-Halbleiter entsteht im Leitungsband des Siliziums ein Potentialtopf, der als Strompfad genutzt werden kann. Die für den Stromfluß notwendigen Ladungsträger stammen dabei aus Dotierschichten, die außerhalb des Si-Kanals angeordnet sind. Das energetisch abgesenkte Leitungsbandniveau in der Si-Schicht sorgt dafür, daß sich die Elektronen im Potentialtopf sammeln und ein hochmobiles Elektronengas bilden.

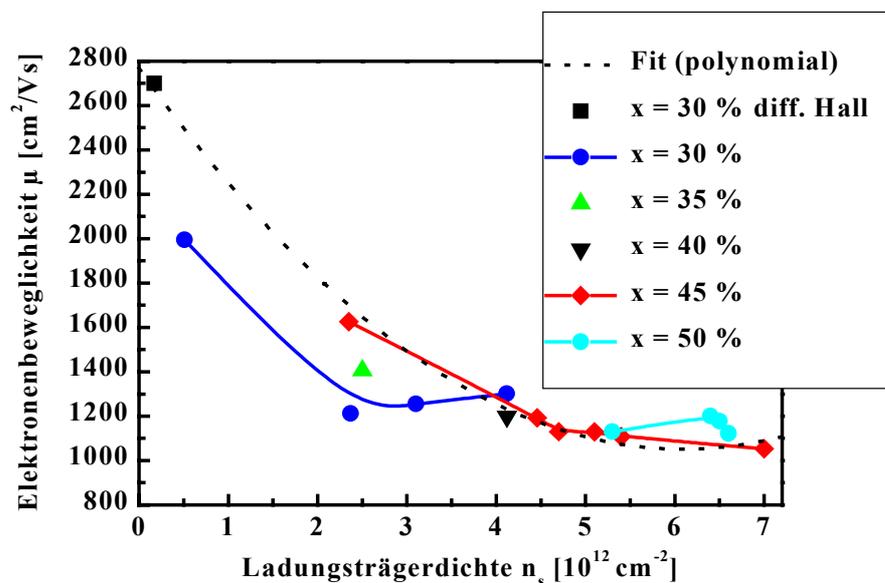


Abbildung 2.11: Erhöhte Elektronenbeweglichkeiten in SiGe-Heterostrukturen mit verspannten Si-Kanälen (2DEG) [2.73]

Die höchste bisher nachgewiesene Beweglichkeit μ für modulationsdotierte Transistorstrukturen mit Si-Kanal beträgt $2830 \text{ cm}^2/\text{Vs}$ [2.74] und ist damit mehr als sechsmal höher als die von Takagi ermittelte „universal mobility“ in konventionellen Si-MOSFETs, die sich im Bereich von $500 - 600 \text{ cm}^2/\text{Vs}$ bewegt [2.75]. Monte-Carlo Simulationen gehen sogar davon aus, daß Beweglichkeitswerte von $3490 \text{ cm}^2/\text{Vs}$ im verspannten Si möglich sind [2.76]. Für den realen Transistor ist das Produkt aus Beweglichkeit μ und Ladungsträgerdichte n_s eine entscheidende Größe, die maßgeblich das elektrische Leistungspotential des SiGe-HFETs bestimmt. Abbildung 2.11 zeigt den Zusammenhang zwischen μ und n_s für modulationsdotierte $\text{Si}_{1-x}\text{Ge}_x$ -HFETs bei unterschiedlichen Ge-Gehalten x .

Neben diesen vielversprechenden Beweglichkeitswerten konnte in Simulationen nachgewiesen werden, daß auch in verspannten Si-Schichten ballistischer Ladungsträgertransport möglich ist. Dabei lassen sich nach Miyata et al. [2.77] transiente Driftgeschwindigkeiten der Elektronen als Antwort auf einen elektrischen Impuls erzielen, die je nach Verspannungsgrad bis zu 100 % über den Werten von relaxiertem Si liegen (Abbildung 2.12).

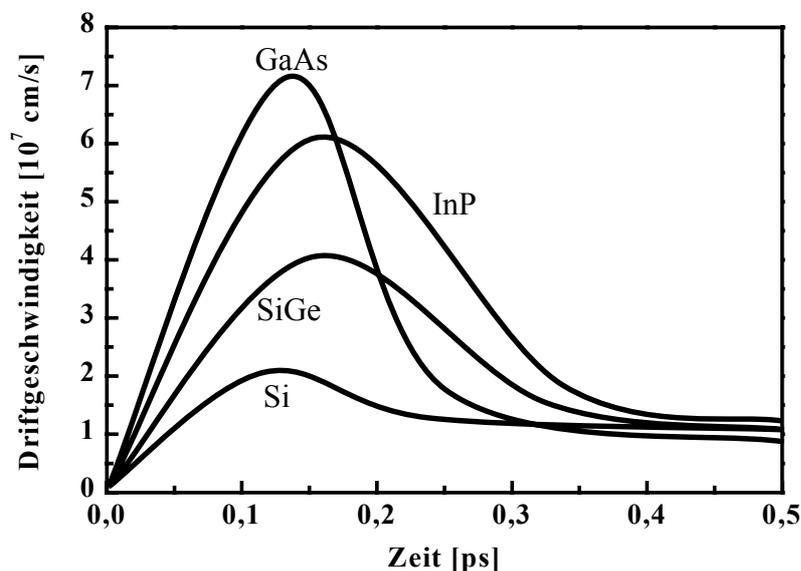


Abbildung 2.12: Simulierter „velocity overshoot“ der Driftgeschwindigkeit für Heterostrukturen verschiedener Materialsysteme [2.77], [2.78]

Basierend auf diesen Annahmen sind Leistungsabschätzungen für das Hochfrequenzverhalten von n-Kanal-SiGe-HFETs entstanden. Abbildung 2.13 vergleicht die Grenzfrequenz der Kurzschlußstromverstärkung f_T eines n-Kanal SiGe-

HFETs und eines Si-n-MOSFETs unter Einbeziehung eines ballistischen Ladungsträgertransports für kurze Gatelängen L_G . Eine Gegenüberstellung der erzielbaren Maximalwerte ergibt, daß durch die erhöhte Elektronenbeweglichkeit des HFETs eine Steigerung der Grenzfrequenz von bis zu 100 % möglich ist. Ergänzende Arbeiten von Paul und Ismail gehen sogar von einem noch optimistischeren Hochfrequenzpotential aus und stellen den n-Kanal-SiGe-HFET mit dem AlGaAs/GaAs-HEMT auf eine Stufe [2.79], [2.80].

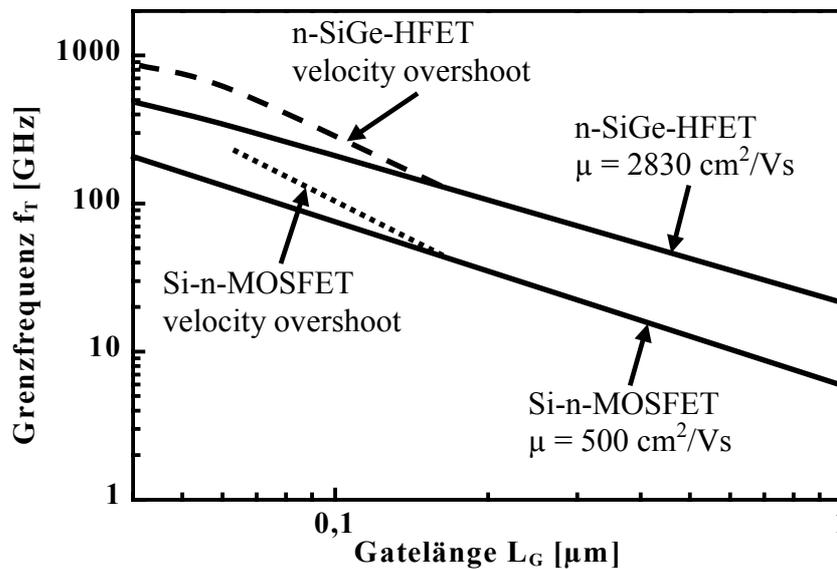


Abbildung 2.13: Vergleich der Hochfrequenzpotentiale (f_T) eines Standard Si-n-MOSFETs und eines n-SiGe-HFETs mit und ohne Berücksichtigung des ballistischen Ladungsträgertransports [2.67]

Ein weiteres wichtiges Gütekriterium für Hochfrequenzbauelemente ist das Rauschverhalten. Aufgrund des vergrabenen Kanals des SiGe-HFETs und der örtlichen Trennung des 2DEGs von den Donatorrümpfen in den Dotierschichten deuten bereits die physikalischen Randbedingungen auf ein hohes Potential hin. Erste Untersuchungen zum Hochfrequenzrauschverhalten des Transistors haben ein exzellentes F_{\min} von 0,3 dB bei 2,5 GHz erzielt und bestätigen damit diese Vermutung (für eine detaillierte Darstellung der Rauscheigenschaften sei an dieser Stelle auf Kapitel 7 dieser Arbeit verwiesen) [2.81]. Im Bereich des niederfrequenten Rauschens haben Messungen von Siemens und der Universität Warwick an SiGe-MOSFETs übereinstimmend Eckfrequenzen f_C zwischen 20 und 100 Hz erzielt [2.82]. Damit liefert der SiGe-HFET sogar niedrigere f_C -Werte als

der SiGe-HBT und liegt deutlich besser als die III/V-Bauelemente, die sich im kHz-Bereich bewegen.

Ein maßgeblicher Vorteil der SiGe-HFETs gegenüber der Si-CMOS-Technologie und insbesondere den III/V-Halbleitern ist die nahezu gleiche Ladungsträgerbeweglichkeit in n- und p-Kanal-Transistoren. Die Umsetzung eines p-Kanal-SiGe-HFETs erfordert ein hochmobiles, zweidimensionales Löchergas (2DHG). Durch die Einbettung einer dünnen, gitterangepaßten SiGe-Epitaxieschicht mit hohem Ge-Gehalt bzw. eines reinen Ge-Kanals in SiGe-Schichten mittleren Ge-Gehaltes kann ein ausgeprägter Potentialtopf im Valenzband des Kanals realisiert werden, in dem sich dann das 2DHG bildet. Die maximale Löcherbeweglichkeit, die mit diesen Heterostrukturen erzielt worden ist, übersteigt mit $3500 \text{ cm}^2/\text{Vs}$ sogar die Werte des n-Kanal-Transistors [2.83]. Die daraus resultierenden Grenzfrequenzen erreichen mit $f_T = 70 \text{ GHz}$ [2.84] und $f_{\text{max}} = 135 \text{ GHz}$ [2.85] das gleiche Niveau des n-HFETs. Die Kombination des n- und des p-Kanal-Transistors in einer komplementären Schaltungslogik führt zu einem neuen SiGe-Hetero-CMOS-Konzept (HCMOS), das die Vorteile der CMOS-Technologie mit dem vielversprechenden Hochfrequenzpotential des SiGe-HFETs vereinigt. Aufgrund der vergleichbaren Beweglichkeiten können die Einzeltransistoren beim HCMOS-Konzept gleich dimensioniert werden. Dadurch erhöht sich die Packungsdichte der Chips, und das Schaltungsdesign vereinfacht sich. Außerdem lassen sich Logikanwendungen und Hochfrequenzkomponenten mit einem Herstellungsprozeß auf einem Chip integrieren.

Erste Integrationskonzepte der HCMOS-Technologie gehen auf König und Sadek zurück [2.86], [2.87]. Sie beruhen auf Schichtstrukturen, die den p- und den n-Kanal Transistor in einem SiGe-Heterosystem übereinander anordnen. Eine experimentelle Realisierung dieser Konzepte steht allerdings noch aus. Jedoch zeigen Simulationen von HCMOS-Ringoszillatoren eindrucksvoll eine vier- bis sechsfache Reduktion des Power-Delay-Produktes im Vergleich zu CMOS-Oszillatoren mit denselben Design-Rules (Abbildung 2.14). Eine Gegenüberstellung der experimentell ermittelten Gatterlaufzeiten (Delays) von Inverterstufen auf der Basis von Si-n-MOS-Transistoren und SiGe-n-HFETs zeigt, daß die verbesserten Transporteigenschaften der Elektronen im SiGe-HFET zu kürzeren Gatterlaufzeiten führen, die einem Vorsprung von etwa 1 - 2 Technologiegenerationen in der SIA-Roadmap entsprechen (Abbildung 2.15) [2.82]. Dadurch ist es möglich, bei gleichem Transistordesign deutlich erhöhte Leistungspotentiale zu erzielen und dabei gleichzeitig über große Verstärkungen, hohe Grenzfrequenzen und gutes Rauschverhalten zu verfügen.

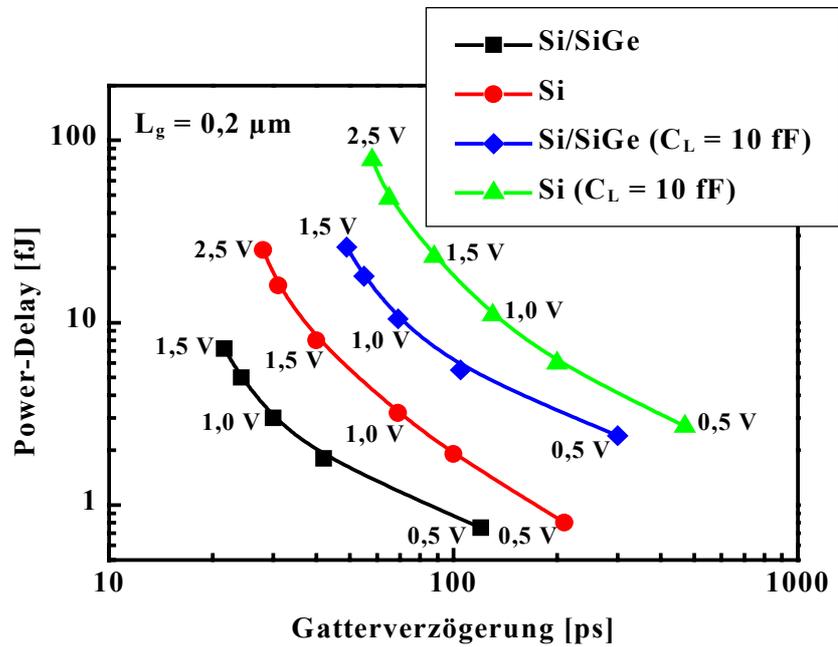


Abbildung 2.14: Power-Delay-Produkt in Abhängigkeit von der Gatterlaufzeit für SiGe-HCMOS- und Si-CMOS-Ringoszillatoren [2.87]

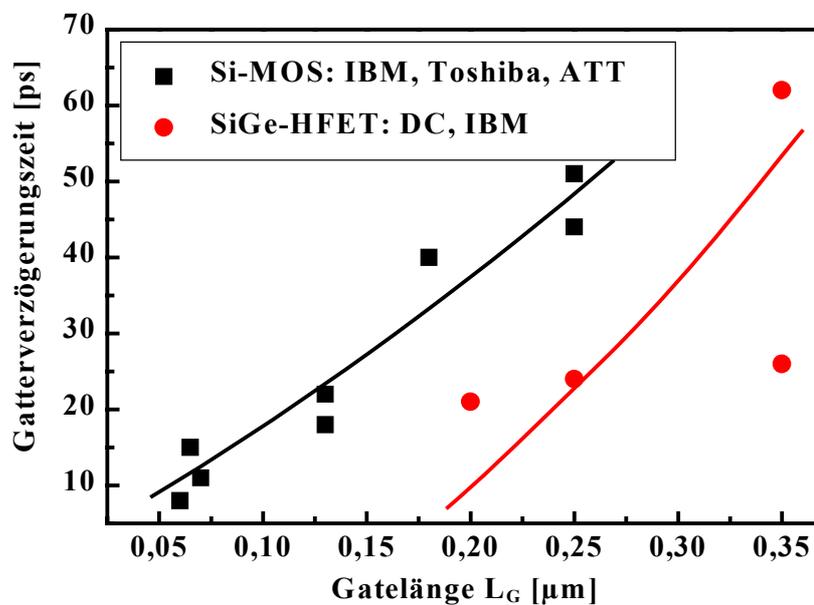


Abbildung 2.15: Vergleich der Gatterverzögerungszeiten von SiGe-n-HFET- und Si-n-MOS-Invertern [2.82]

Neben dem Leistungspotential eines Bauelementes sind die Produktionskosten das entscheidende Kriterium, ob es zu einer industriellen Einführung kommt oder das Konzept verfällt. Der große Vorteil der SiGe-Technologie gegenüber den III/V-Halbleitern ist die hohe Kompatibilität zu den Si-Standardprozessen. Wie bereits bei der SiGe-BiCMOS-Technologie dargestellt, kann trotz der zusätzlichen Implementierung von Prozeßmodulen für die SiGe-Epitaxie und die passiven Bauelemente eine große Übereinstimmung von etwa 90 % erhalten bleiben. Der Kostenaufwand für die Modifikation der bestehenden Silizium-Produktionslinie ist damit gering. Eine im vorhinein für den SiGe-HBT gemachte konservative Abschätzung der Kostenerhöhung von 20 - 30 % [2.67] ist bei der Einführung der Technologie sogar noch durch die Halbleiterindustrie mit einem Anstieg von nur 10 % unterboten worden.

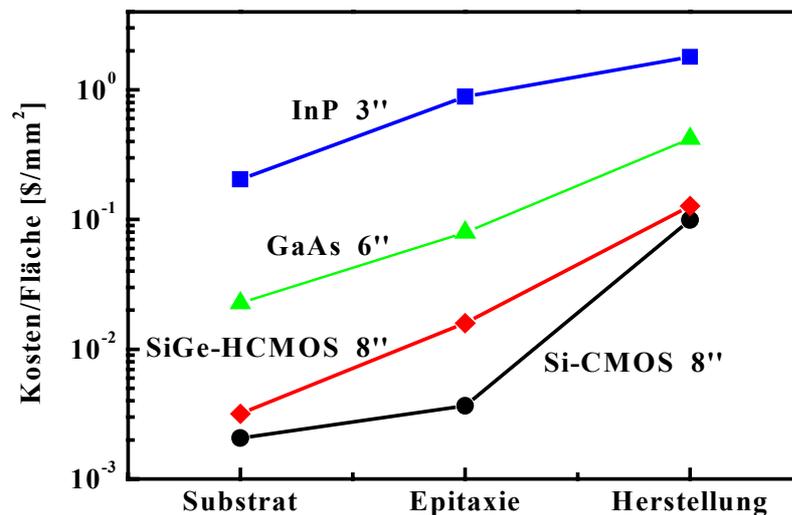


Abbildung 2.16: Kostenabschätzung für ein SiGe-HCMOS-Konzept und Vergleich mit Technologien anderer Materialsysteme

Für die Realisierung einer SiGe-HFET-Technologie ist der Epitaxieanteil höher, da hier ein SiGe-Puffer die Basis für die aktiven Bauelemente darstellt (Abbildung 2.16). Dadurch steigert sich auch der Kostenbeitrag für das Schichtwachstum im Vergleich zum HBT auf etwa den fünffachen Wert [2.88]. Bei GaAs- oder InP-Bauelementen liegt bereits der Preis für einen Substratwafer deutlich über dem des Siliziums. Dabei muß beachtet werden, daß die Größe der GaAs-Substrate etwa zwei Generationen dem Siliziummarkt hinterherhinkt. Durch zusätzlich höhere Epitaxie- und Prozessierungskosten fallen die Chipkosten pro mm^2 Waferfläche signifikant höher aus, selbst im Vergleich zum SiGe-

HFET. Bei einer Gegenüberstellung des Frequenz/Kosten-Verhältnisses für eine spezifische Schaltungsanwendung liegt die SiGe-Technologie sogar eindeutig im Vorteil gegenüber den Si- und den GaAs-Halbleitern (Abbildung 2.17).

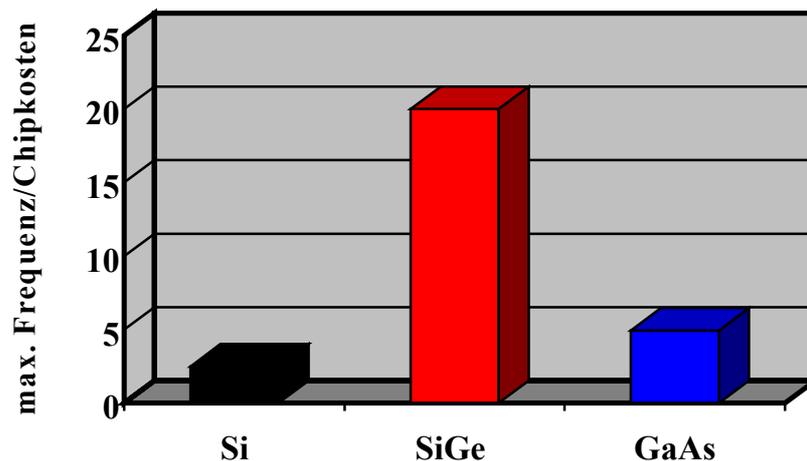


Abbildung 2.17: Gegenüberstellung des Frequenz/Chipkosten-Verhältnisses für Si-, SiGe- und GaAs-Anwendungen [2.68]

Die Erfolgsgeschichte des SiGe-HBTs, der kommerzielle Durchbruch der SiGe-BiCMOS-Technologie und das hohe Leistungspotential des SiGe-HFETs sind deutliche Beweise dafür, daß sich die SiGe-Technologie zu einem festen Bestandteil des Halbleitermarktes, als Hochfrequenzsegment der Siliziumbranche, entwickeln wird. Das exzellente Hochfrequenzpotential des SiGe-Materialsystems und die hohe Kompatibilität zur etablierten Standard Si-Technologie versprechen dabei eine Ausweitung des Si(Ge)-Anwendungsbereiches bis hin zu Frequenzen von 50 GHz. Entscheidender Vorteil gegenüber den III/V-Halbleitern ist dabei die Verfügbarkeit ausgereifter und hoch optimierter Integrationskonzepte bei deutlich geringeren Herstellungs- und Materialkosten.

Diese Arbeit soll einen Beitrag dazu leisten, das kommerzielle Erfolgspotential des SiGe-HFETs zu erhöhen und dadurch einen vergleichbaren Entwicklungsverlauf wie beim SiGe-HBT zu beschleunigen. Dazu ist es notwendig, durch eine gezielte Optimierung des Heteroschichtsystems und des Technologiekonzeptes die theoretischen Potentialabschätzungen experimentell zu verifizieren und die Integrationsfähigkeit, auch in Hinblick auf spätere Schaltungsanwendungen, zu steigern.

Kapitel 3

Aufbau und Funktion des SiGe-n-Kanal-Hetero-Feldeffekttransistors

3.1 Grundlagen zum SiGe-Materialsystem

3.1.1 $\text{Si}_{1-x}\text{Ge}_x$ -Mischkristalle

Die Elementhalbleiter Silizium (Si) und Germanium (Ge) gehören der IV. Hauptgruppe des Periodensystems der Elemente an und kristallisieren im Diamantgitter. Die Einheitszelle des Diamantgitters läßt sich durch zwei kubisch flächenzentrierte Gitter beschreiben, die um ein Viertel der Raumdiagonale gegeneinander verschoben sind [3.1]. Die Kantenlänge der Einheitszelle entspricht der Gitterkonstante a des Elementes. Silizium und Germanium haben mit $a_{\text{Si}} = 5,431 \text{ \AA}$ und $a_{\text{Ge}} = 5,658 \text{ \AA}$ Gitterkonstanten, die um 4,18 % voneinander abweichen.

Moderne Epitaxieverfahren, wie z. B. die Molekularstrahlepitaxie (MBE) oder die chemische Abscheidung von Verbindungen aus der Gasphase (CVD), ermöglichen das Wachstum von SiGe-Mischkristallen in beliebigen Mischungsverhältnissen. Der binäre Halbleiter $\text{Si}_{1-x}\text{Ge}_x$ kristallisiert dabei wie die Elementhalbleiter im Diamantgitter, wobei Si- und Ge- Atome statistisch auf den Gitterplätzen verteilt sind. Die Gitterkonstante a_{SiGe} von $\text{Si}_{1-x}\text{Ge}_x$ -Mischkristallen läßt sich in guter Näherung mit dem Vegardschen Gesetz bestimmen, das linear zwischen den Gitterkonstanten der Elementhalbleiter a_{Si} und a_{Ge} interpoliert [3.2]. Die Abweichung der tatsächlichen Gitterkonstante vom Vegardschen Gesetz liegt unter 0,125 %.

$$a_{\text{SiGe}}(x) = a_{\text{Si}} \cdot (1 - x) + a_{\text{Ge}}x \quad (3.1)$$

Die Bandstruktur von $\text{Si}_{1-x}\text{Ge}_x$ -Legierungen kann teilweise durch lineare Interpolation aus den Bandstrukturen von Si und Ge ermittelt werden. Insbesondere Übergänge an kritischen Punkten zeigen eine nahezu lineare Änderung mit dem Ge-Gehalt. Wie bei den Elementhalbleitern weist der Mischkristall für alle Kompositionen einen indirekten Bandübergang auf. Bei der Bestimmung des Bandabstandes ist eine einfache lineare Interpolation nicht möglich. Abbildung 3.1 zeigt den von Weber und Alonso [3.3] mittels Photolumineszenz ermittelten Verlauf des Bandabstandes für unverspanntes $\text{Si}_{1-x}\text{Ge}_x$ bei einer Temperatur von 4,2 K als Funktion des Ge-Gehaltes x . Die Bandlücke variiert nichtlinear zwischen dem Wert des reinen Siliziums $E_{g,\text{Si}}(T = 4,2 \text{ K}) = 1,155 \text{ eV}$ und dem des reinen Germaniums $E_{g,\text{Ge}}(T = 4,2 \text{ K}) = 0,740 \text{ eV}$. Bei $x = 0,85$ geht der Mischkristall vom Si-artigen in den Ge-artigen Zustand über. Dies ist physikalisch darin begründet, daß für $x < 0,85$ das sechsfach entartete Leitungsbandminimum des Siliziums auf der Δ -Linie (Δ -Minimum) das niedrigste Leitungsbandniveau darstellt. Für $x \geq 0,85$ liegt jedoch das achtfach entartete Leitungsbandminimum des Germaniums am L-Punkt (L-Minimum) energetisch tiefer und bestimmt damit den Bandabstand [3.4].

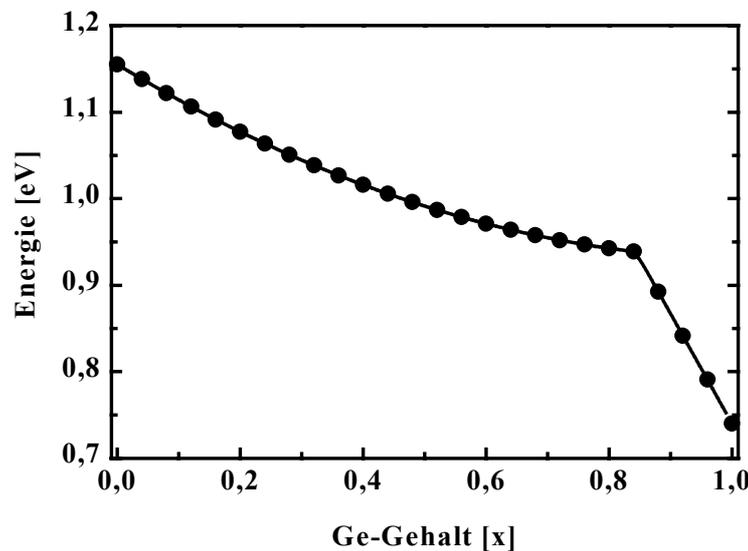


Abbildung 3.1: Verlauf des Bandabstandes für unverspanntes $\text{Si}_{1-x}\text{Ge}_x$ als Funktion des Ge-Gehaltes x bei 4,2 K [3.3]

Analytisch wird der experimentelle Verlauf gut durch die folgenden Gleichungen dargestellt.

$$E_{g,\text{SiGe}_u}(x) = 1,155 - 0,43x + 0,2056x^2 \text{ eV} \quad \text{für } 0 < x < 0,85 \quad (3.2)$$

$$E_{g,\text{SiGe}_u}(x) = 2,01 - 1,27x \text{ eV} \quad \text{für } 0,85 < x < 1 \quad (3.3)$$

Die numerische Ermittlung des Bandabstandes von relaxierten SiGe-Schichten ist ein wichtiges Hilfsmittel in Hinblick auf die Dimensionierung und Konfiguration des Heteroschichtsystems für den SiGe-HFET. Zusammen mit den Energiebandverschiebungen am Heteroübergang, die explizit im folgenden Abschnitt hergeleitet werden, bestimmt die Energielücke des virtuellen SiGe-Substrates maßgeblich das „Confinement“ der Elektronen im Si-Kanal und damit die Ladungsträgerdichte im 2DEG.

3.1.2 Verspannte SiGe-Heterostrukturen

Standardmäßig werden $\text{Si}_{1-x}\text{Ge}_x$ -Heterostrukturen auf Silizium als Substrat aufgewachsen. Die daraus resultierenden Vorteile sind die bewährten chemischen, mechanischen und thermischen Eigenschaften des Siliziums, das breite und kostengünstige Angebot von Siliziumwafern unterschiedlicher Spezifikationen und die Kompatibilität zu konventionellen Bauelementen und ICs auf Si-Basis.

Durch die unterschiedlichen Gitterkonstanten von Si und Ge kommt es beim Wachstum von Mischkristallen oder Heterostrukturen zu einer Gitterfehlpassung η , die maximal 4,18 % betragen kann [3.5].

$$\eta(x) = \frac{a_{\text{SiGe}}(x) - a_{\text{Si}}}{a_{\text{Si}}} \quad (3.4)$$

Je nach aufgewachsener Schichtdicke bestehen für das Material zwei Möglichkeiten auf die Fehlanpassung zu reagieren. Sind die auf dem Substrat abgeschiedenen Schichten dicker als die kritische Schichtdicke d_{krit} , so wird die in der Epitaxieschicht gespeicherte, homogene elastische Energie so groß, daß es energetisch günstiger ist, Versetzungslinien zu bilden. Dadurch wird die Verspannung zunächst teilweise abgebaut. Bei weiter zunehmender Schichtdicke entstehen so viele Versetzungen, daß die Verspannung verschwindet und die Schicht vollständig relaxiert (Abbildung 3.2 a).

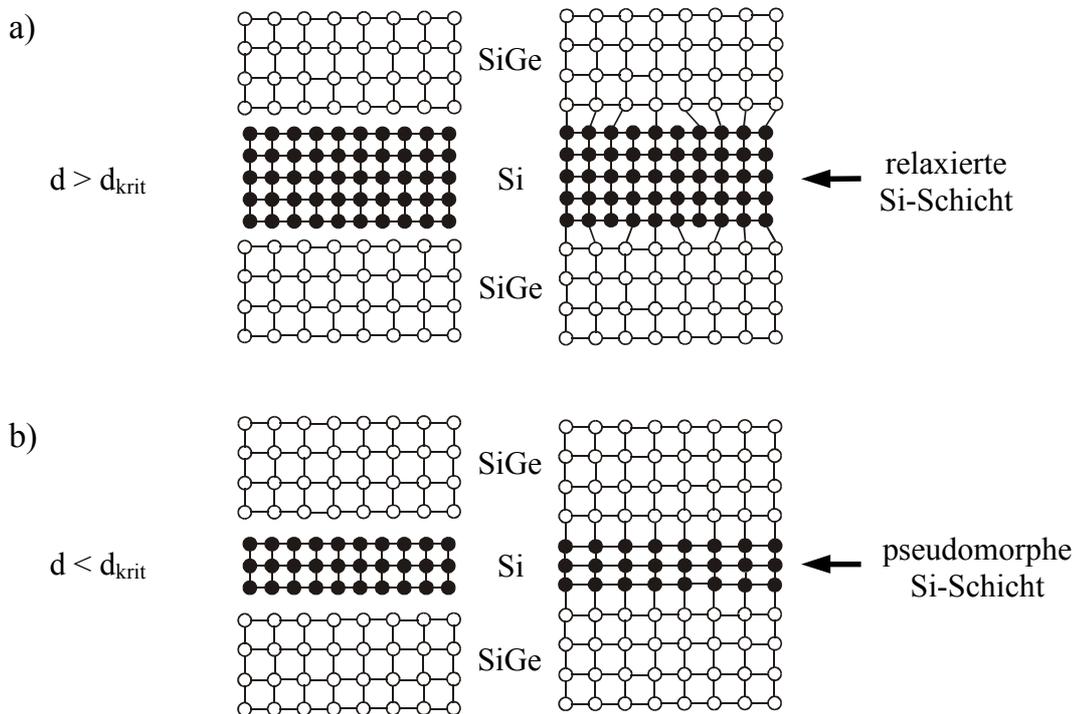


Abbildung 3.2: Schematische Darstellung des Wachstums einer Si-Epitaxieschicht der Dicke d auf relaxiertem SiGe. a) $d > d_{krit}$: die Si-Schicht relaxiert unter Bildung von Versetzungslinien; b) $d < d_{krit}$: das Wachstum der Si-Schicht erfolgt gitterangepaßt

Bei aufgewachsenen Schichten mit einer Dicke d , die geringer ist als die kritische Schichtdicke d_{krit} , nimmt die Epitaxieschicht die Gitterkonstante a_{Sub} des Substrates an. Die aufgewachsene Schicht wird dadurch kristallografisch verspannt und das Wachstum als gitterangepaßt oder pseudomorph bezeichnet. Beim n-Kanal-SiGe-HFET, der im Mittelpunkt dieser Arbeit steht, werden die verbesserten Transporteigenschaften von Elektronen in pseudomorphen Si-Schichten ausgenutzt. Um auf einem Siliziumsubstrat eine verspannte Si-Schicht realisieren zu können, muß erst eine relaxierte $Si_{1-x}Ge_x$ -Pufferschicht, die als virtuelles Substrat fungiert, erzeugt werden. Die darauf gewachsene Epitaxieschicht aus Silizium erfährt, aufgrund der kleineren Gitterkonstante a_{Si} , eine biaxiale Dehnung in der Wachstumsebene und wird gleichzeitig in Wachstumsrichtung komprimiert (Abbildung 3.2 b). Die folgenden Betrachtungen zum SiGe-Heteroschichtsystem konzentrieren sich auf den Fall einer pseudomorphen Si-Epitaxieschicht auf einem relaxierten $Si_{1-x}Ge_x$ -Puffer, da nur diese Konstellation für den n-Kanal-SiGe-Hetero-Feldeffekttransistor von Relevanz ist.

Die kritische Schichtdicke d_{krit} , bis zu der pseudomorphes Schichtwachstum möglich ist, hängt von dem Germaniumgehalt x ab. Nach der von Kasper und Herzog [3.6] aufgestellten mechanischen Gleichgewichtstheorie ist die Stabilitätsgrenze für pseudomorphes Wachstum erreicht, wenn die Ausbildung von Gitterversetzungen energetisch günstiger ist als die weitere Aufnahme elastischer Energie des verspannten Kristalls. Daraus folgt die implizite Formel zur Berechnung der kritischen Schichtdicke in Abhängigkeit vom Ge-Gehalt.

$$d_{\text{krit}} = \frac{1,175 \cdot 10^{-2}}{\eta(x)} \cdot \ln(8,9 \cdot d_{\text{krit}}) \quad (3.5)$$

Die Änderung der Gitterkonstante aufgrund des pseudomorphen Kristallwachstums wirkt sich auch auf die Energiestruktur der Epitaxieschicht aus. Die Energieniveaus verschieben sich, entartete Zustände, wie die Valenzbandmaxima am Γ -Punkt und die Leitungsbandminima, spalten sich auf und daraus resultierend ändert sich auch der Bandabstand der gitterangepaßten Schicht. Die Energieverschiebungen lassen sich mit Hilfe der Deformationpotentialtheorie berechnen [3.7], [3.8]. Sie gibt eine analytische Beziehung zwischen der Bandverschiebung und der Verspannung des Kristalls an. Für kleine Verspannungen gilt das Hooke'sche Gesetz, das über den Tensor der Elastizitätskonstanten c einen linearen Zusammenhang zwischen den Komponenten des Spannungs- und des Dehnungstensors herstellt.

$$\vec{\sigma} = \vec{c} \cdot \vec{\varepsilon} \quad (3.6)$$

Die biaxialen Verspannungen pseudomorpher Schichten können in einen isotropen, hydrostatischen Anteil und einen anisotropen, uniaxialen Anteil aufgespalten werden. Die hydrostatische Komponente steht für eine gleichmäßige Wirkung in alle drei Raumrichtungen ($\sigma_x = \sigma_y = \sigma_z$) und sorgt für eine Verschiebung der Energieschwerpunkte des Leitungs- und Valenzbandes. Der uniaxiale Anteil, der durch die Verringerung der Kristallsymmetrie entsteht, wirkt hingegen nur in negativer Wachstumsrichtung ($\sigma_x = \sigma_y = 0; \sigma_z \neq 0$), was zur Aufspaltung der entarteten Bänder führt [3.9].

Mit Hilfe der Gitterkonstanten $a_{\text{Si}\perp}$ und $a_{\text{Si}\parallel}$ der verspannten Si-Elementarzelle und a_{SiGe} der unverspannten $\text{Si}_{1-x}\text{Ge}_x$ -Schicht lassen sich die Komponenten des Dehnungstensors ε bestimmen. Dabei kennzeichnet der Index \perp Komponenten in Wachstumsrichtung, und der Index \parallel steht für eine Ausrichtung in der Wachstumsebene.

$$\varepsilon_{\parallel} = \left(\frac{a_{\text{Si}\parallel}(\mathbf{x}) - a_{\text{Si}}}{a_{\text{Si}}} \right) \quad \text{mit} \quad a_{\text{Si}\parallel}(\mathbf{x}) = a_{\text{SiGe}}(\mathbf{x}) \quad (3.7)$$

$$\varepsilon_{\perp} = \left(\frac{a_{\text{Si}\perp}(\mathbf{x}) - a_{\text{Si}}}{a_{\text{Si}}} \right) \quad (3.8)$$

$$a_{\text{Si}\perp} = a_{\text{Si}} \left(1 - 2 \cdot \frac{c_{\text{Si}12}}{c_{\text{Si}11}} \cdot \frac{a_{\text{Si}\parallel}(\mathbf{x}) - a_{\text{Si}}}{a_{\text{Si}}} \right) \quad (3.9)$$

Die für die Berechnung erforderlichen Elastizitätskonstanten ergeben sich nach [3.10] zu $c_{\text{Si}11} = 1,675$ Mbar und $c_{\text{Si}12} = 0,65$ Mbar.

Mit den berechneten Komponenten des Dehnungstensors können nun, unter Einbeziehung des hydrostatischen Deformationspotentials a des Valenzbandes und den Deformationspotentialen Ξ_{d} und Ξ_{u} des Leitungsbandes, die Verschiebungen der Bandschwerpunkte berechnet werden.

$$\Delta E_{\text{V}} = a \cdot (2\varepsilon_{\parallel} + \varepsilon_{\perp}) \quad (3.10)$$

$$\Delta E_{\text{L}} = \left(\Xi_{\text{d}} + \frac{1}{3}\Xi_{\text{u}} \right) \cdot (2\varepsilon_{\parallel} + \varepsilon_{\perp}) \quad (3.11)$$

Aus den Gleichungen 3.10 und 3.11 ergibt sich die Änderung des Bandabstandes durch den hydrostatischen Verspannungsanteil.

$$\Delta E_{\text{g,hyd}} = \Delta E_{\text{L}} - \Delta E_{\text{V}} = \left(\Xi_{\text{d}} + \frac{1}{3}\Xi_{\text{u}} - a \right) \cdot (2\varepsilon_{\parallel} + \varepsilon_{\perp}) \quad (3.12)$$

| | $(\Xi_{\text{d}} + 1/3\Xi_{\text{u}} - a)$ [eV] (für Δ -Minima) | $(\Xi_{\text{d}} + 1/3\Xi_{\text{u}} - a)$ [eV] (für L-Minima) | b [eV] | Δ_0 [eV] | Ξ_{u}^{Δ} [eV] |
|----|---|---|----------|-----------------|--------------------------------|
| Si | 1,72 | -3,12 | -2,35 | 0,044 | 9,16 |
| Ge | 1,31 | -2,78 | -2,55 | 0,29 | 9,42 |

Tabelle 3.1: Theoretisch ermittelte Deformationspotentiale für Silizium und Germanium nach [3.11], [3.12]

Das Maximum des Valenzbandes liegt sowohl bei Silizium als auch bei Germanium am Γ -Punkt und ist jeweils sechsfach entartet. Durch die Deformation des pseudomorphen Si-Kristalls und die Spin-Bahn-Kopplung bewirkt der uniaxiale Verspannungsanteil eine Aufspaltung der Si-Valenzbandkante in einen vierfach entarteten (LH = light holes und HH = heavy holes) und einen zweifach entarteten Zustand (SO = split-off) [3.12].

$$\Delta E_V^{\text{LH}} = -\frac{1}{6}\Delta_0 + \frac{1}{4}\delta E_{001} + \frac{1}{2}\left(\Delta_0^2 + \Delta_0\delta E_{001} + \frac{9}{4}\delta E_{001}^2\right)^{1/2} \quad (3.13)$$

$$\Delta E_V^{\text{HH}} = \frac{1}{3}\Delta_0 - \frac{1}{2}\delta E_{001} \quad (3.14)$$

$$\Delta E_V^{\text{SO}} = -\frac{1}{6}\Delta_0 + \frac{1}{4}\delta E_{001} - \frac{1}{2}\left(\Delta_0^2 + \Delta_0\delta E_{001} + \frac{9}{4}\delta E_{001}^2\right)^{1/2} \quad (3.15)$$

$$\text{mit } \delta E_{001} = 2b(\varepsilon_{\perp} - \varepsilon_{\parallel})$$

Für das Leitungsband ergibt sich bei einem Germaniumgehalt unter 85 % eine Aufspaltung des sechsfach entarteten Δ -Minimums in ein zweifach entartetes Δ_2 -Minimum in Wachstumsrichtung und ein vierfach entartetes, energetisch höher liegendes Δ_4 -Minimum mit der Ausrichtung parallel zur Wachstumsebene.

$$\Delta E_L^{\Delta_2} = +\frac{2}{3}\Xi_u^{\Delta}(\varepsilon_{\perp} - \varepsilon_{\parallel}) \quad (3.16)$$

$$\Delta E_L^{\Delta_4} = -\frac{1}{3}\Xi_u^{\Delta}(\varepsilon_{\perp} - \varepsilon_{\parallel}) \quad (3.17)$$

Mit Hilfe der ermittelten hydrostatischen Bandlückenveränderung $\Delta E_{g,\text{hyd}}$ und der berechneten Aufspaltung der Bänder kann nun der indirekte Bandabstand $E_{g\text{Si},s}$ der verspannten Si Epitaxieschicht ermittelt werden.

$$E_{g\text{Si},s} = E_{g\text{Si}} + \Delta E_{g,\text{hyd}} + \Delta E_L^{\Delta_2} - \Delta E_V^{\text{LH}} \quad (3.18)$$

Die Auswirkungen der hydrostatischen und uniaxialen Verspannungsanteile auf die Bandstruktur der pseudomorphen Siliziumschicht sind in Abbildung 3.3 noch einmal schematisch dargestellt.

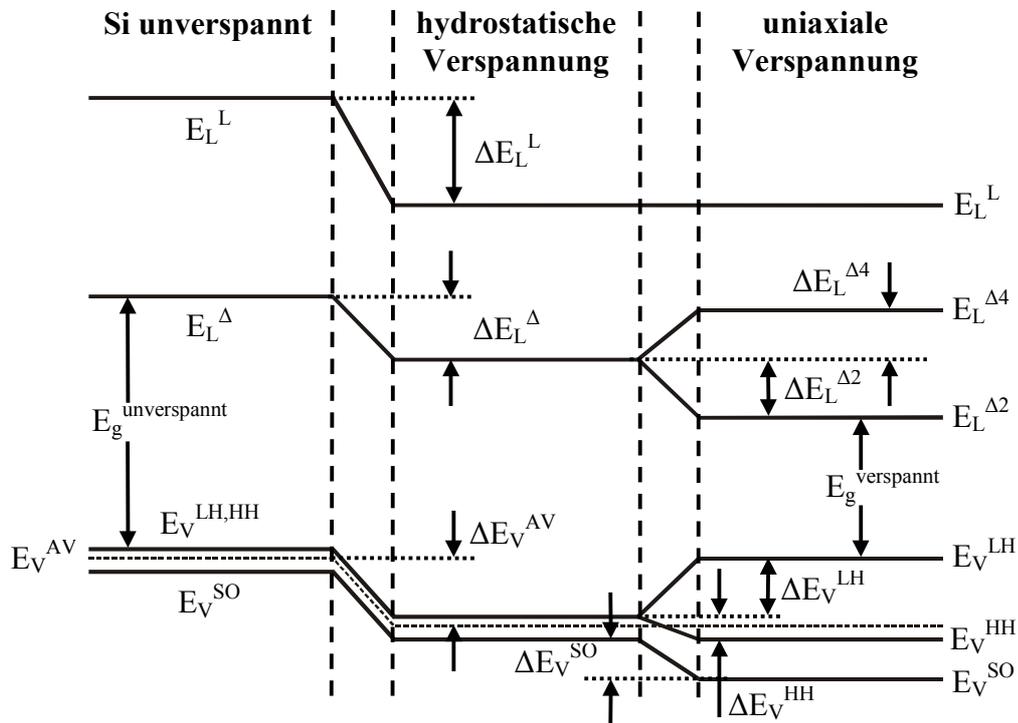


Abbildung 3.3: Schematische Darstellung der hydrostatischen und uniaxialen Bandaufspaltungen für einen $\text{Si}_{0.55}\text{Ge}_{0.45}/\text{Si}$ -Heterübergang

Auf der Basis der theoretischen Arbeiten von Van de Walle und Martin [3.7], [3.12] und Colombo et al. [3.13] präsentiert Schäffler 1997 [3.14] quantitative Verläufe der verschiedenen Leitungs- und Valenzbandzustände in Abhängigkeit vom Germaniumgehalt x des virtuellen $\text{Si}_{1-x}\text{Ge}_x$ -Substrates (Abbildung 3.4). Dabei repräsentieren die roten Linien die Bandniveaus im SiGe-Substrat, während die blauen Linien die Zustände in der verspannten Si-Schicht darstellen. Hier ist deutlich die zunehmende Aufspaltung der Energieniveaus im Leitungs- und Valenzband der Si-Schicht und die daraus resultierende Abnahme der Bandlücke zu erkennen. Außerdem weist der Schnittpunkt zwischen dem Δ_{Sub} - und dem L_{Sub} -Verlauf bei einem Germaniumgehalt von $x \approx 0,8$ darauf hin, daß das SiGe-Substrat vom Si-artigen in den Ge-artigen Zustand übergeht. Dabei verschiebt sich das Leitungsbandminimum vom Δ -Punkt zum L-Punkt und entartet entsprechend der physikalischen Eigenschaften des Germaniums achtfach.

Aus Abbildung 3.4 lassen sich ebenfalls direkt die Werte der Banddiskontinuitäten für das Valenz- und Leitungsband am Heterübergang zwischen SiGe-Substrat und verspannter Si-Schicht bestimmen.

$$\Delta E_V = E_V^{LH} - E_V^{LH,HH}_{Sub} \quad (3.19)$$

$$\Delta E_L = E_L^{\Delta^2} - E_L^{\Delta}_{Sub} \quad \text{für } x < 0,85 \quad (3.20)$$

$$\Delta E_L = E_L^{\Delta^2} - E_L^{\Delta^2}_{Sub} \quad \text{für } x > 0,85 \quad (3.21)$$

Es ist deutlich erkennbar, daß der Leitungsbandsprung stärker mit dem Ge-Gehalt wächst als der Valenzbandsprung und bei $x = 0,5$ etwa dreimal größer ist. Diese Erkenntnis bekräftigt die Anwendung dieser Schichtkonfiguration zur Realisierung von SiGe-HFETs mit einem Elektronenkanal.

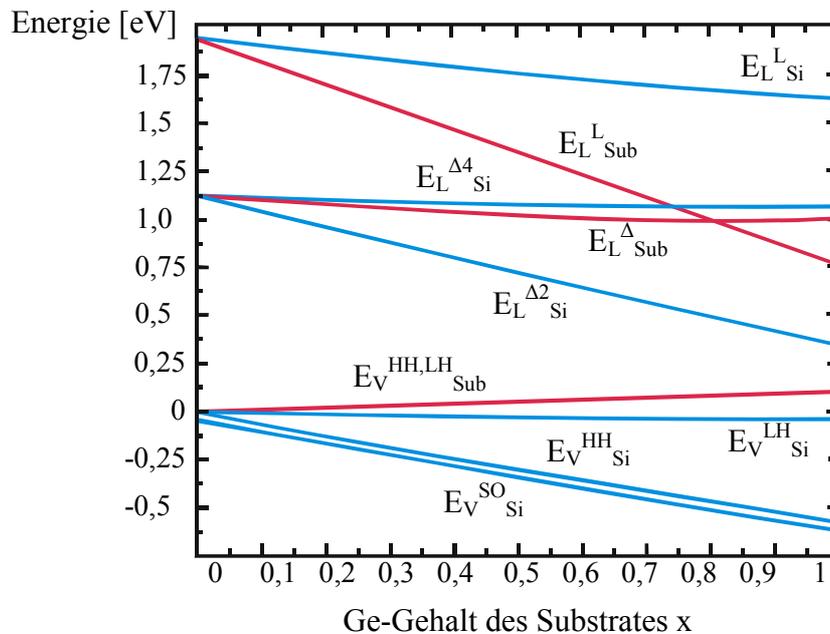


Abbildung 3.4: Leitungs- und Valenzbandzustände für eine pseudomorphe Si-Schicht (blau) auf einem relaxierten $Si_{1-x}Ge_x$ -Substrat (rot)

Eine einfache Interpolationsformel zur Bestimmung der Banddiskontinuität des gemittelten Valenzbandniveaus ΔE_V^{AV} in einem SiGe-Schichtsystem geben Rieger und Vogl in [3.10] an. Dabei gehen sie davon aus, daß sich eine pseudomorphe $Si_{1-x}Ge_x$ -Schicht auf einem relaxierten $Si_{1-y}Ge_y$ -Substrat befindet. Für den hier betrachteten Fall einer verspannten Si-Schicht auf einem virtuellen SiGe-

Puffer muß der Ge-Gehalt x der Epitaxieschicht zu Null gesetzt werden, damit die Formel übertragbar ist.

$$\Delta E_v^{AV}(x, y) = (0,47 - 0,06y) \cdot (x - y) \quad (3.22)$$

Der entstehende Leitungsbandsprung für den hier angenommenen Fall läßt sich durch die folgende analytische Beziehung gut darstellen.

$$\Delta E_L(x) = -0,2x^2 - 0,53x \quad (3.23)$$

Die aus Abbildung 3.4 ermittelten Bandsprünge am Heterübergang zeigen eindeutig, daß die Si/Si_{1-x}Ge_x-Schichtkonstellation für alle Ge-Gehalte x einen Typ II-Heterübergang bildet, bei dem beide Banddiskontinuitäten negativ auftreten und der Bandabstand der pseudomorphen Si-Schicht immer kleiner ist als der des SiGe-Substrates (Abbildung 3.5). Ausgeprägte Typ I-Heterübergänge ergeben sich im SiGe-System für germaniumreiche pseudomorphe Si_{1-x}Ge_x-Schichten mit $x > 0,8$, die auf relaxierten SiGe-Puffern mit einem Germaniumgehalt größer als 60 % gewachsen werden [3.14]. Dabei liegen sowohl das Leitungsbandminimum als auch das Valenzbandmaximum in der verspannten SiGe-Schicht. Diese Schichtanordnungen zeigen einen stark ausgebildeten Potentialtopf im Valenzband und werden deshalb zur Realisierung von SiGe-p-Kanal-Hetero-Feldeffekt-transistoren eingesetzt [3.15], [3.16].

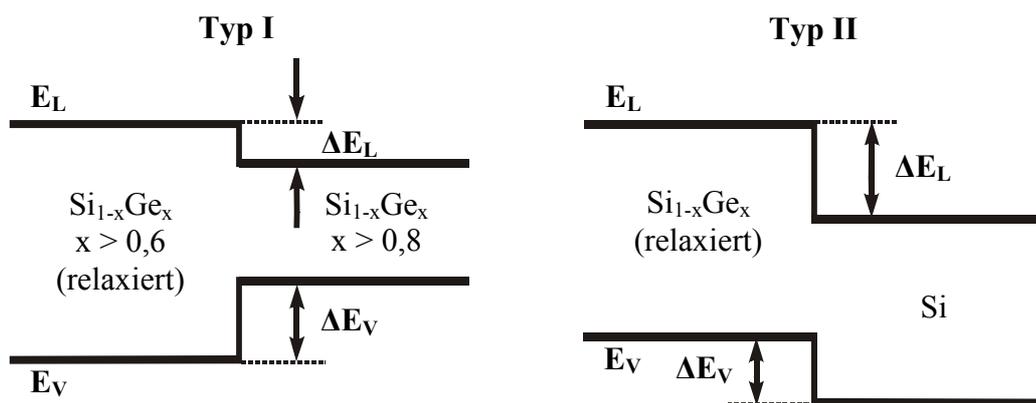


Abbildung 3.5: Typ I- und Typ II-Heterübergang im Si/SiGe-System

3.2 Bauelementstruktur und Funktionsprinzip des SiGe-n-Kanal-Hetero-FETs

Erste Versuche mit modulationsdotierten Heterostrukturen im SiGe-Materialsystem gehen auf People et al. in das Jahr 1984 zurück [3.17], [3.18]. Sie experimentierten mit einem pseudomorphen SiGe-Kanal, der zwischen bordotierten Si-Schichten eingebettet war und Löcherbeweglichkeiten von $3300 \text{ cm}^2/\text{Vs}$ bei 4,2 K lieferte. Ein Jahr später veröffentlichten Jorke und Herzog eine Struktur, bei der zwischen dotierten SiGe-Schichten eingebrachte, verspannte Si-Kanäle eine erhöhte Elektronenmobilität aufwiesen [3.19]. Damit war das Funktionsprinzip und die Realisierbarkeit von hochbeweglichen Elektronengasen in modulationsdotierten Si/SiGe-Heterostrukturen experimentell nachgewiesen. Bereits im darauf folgenden Jahr wendeten Dämbkes et al. [3.20] dieses Prinzip bei der Herstellung des ersten SiGe-n-Kanal-HFETs mit Modulationsdotierung an. Den schematischen Aufbau eines SiGe-Hetero-Feldeffekttransistors und die Schichtstruktur für eine Bauelementversion mit doppelseitig dotiertem Siliziumkanal zeigt Abbildung 3.6.

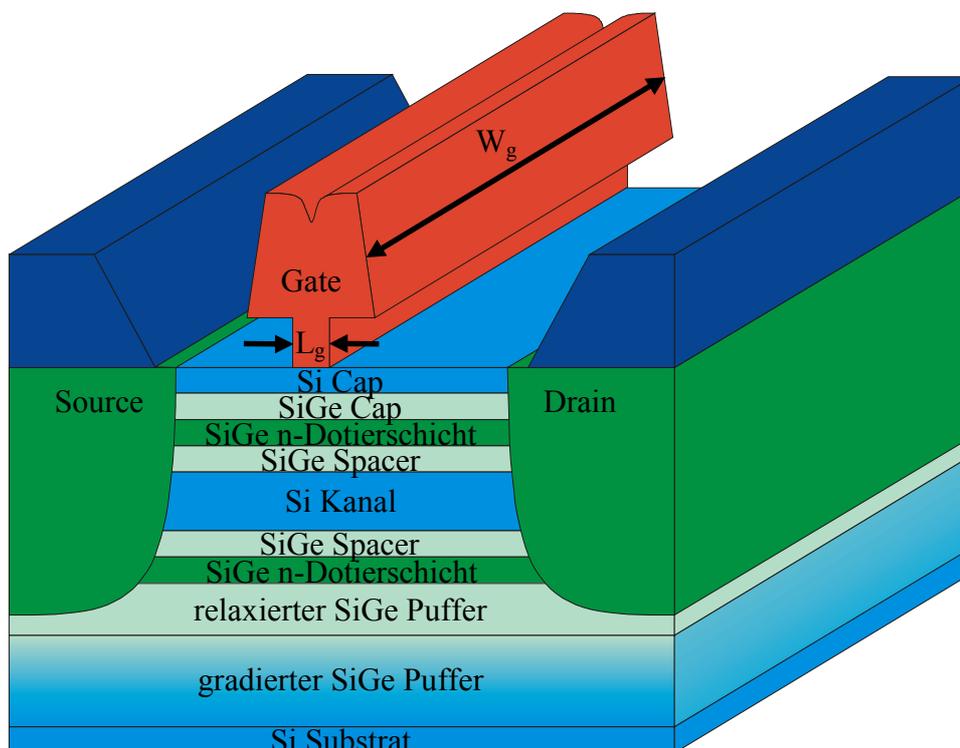


Abbildung 3.6: Schematischer Aufbau eines SiGe-n-Kanal-HFETs mit doppelseitig dotiertem Siliziumkanal

Die Basis für eine modulationsdotierte SiGe-Heterostruktur mit Potentialtopf im Leitungsband stellt immer eine Pufferschicht dar, die den gewünschten Ge-Gehalt für die aktiven Schichten bereitstellt. Meistens gliedert sich der Puffer in einen Gradiententeil und einen Bereich mit konstantem Ge-Gehalt. Der Gradientenpuffer bildet das Interface zum hochohmigen Siliziumsubstrat und steigert seinen Ge-Anteil linear oder parabolisch mit zunehmender Dicke auf den geforderten Endwert. Die nachfolgende Pufferschicht mit konstanter Komposition stellt das eigentliche „virtuelle Substrat“ dar und sollte deshalb vollkommen relaxiert sein und eine möglichst geringe Defektdichte aufweisen. Das Herz der Schichtstruktur bildet der etwa 5 -10 nm dicke, pseudomorphe Si-Kanal, der symmetrisch zwischen den undotierten SiGe-Spacern und den Dotierschichten auf der Vorder- und Rückseite eingebettet ist. Die gesamte Schichtstruktur schließt mit einem SiGe-Cap-Spacer und einer Si-Schicht ab.

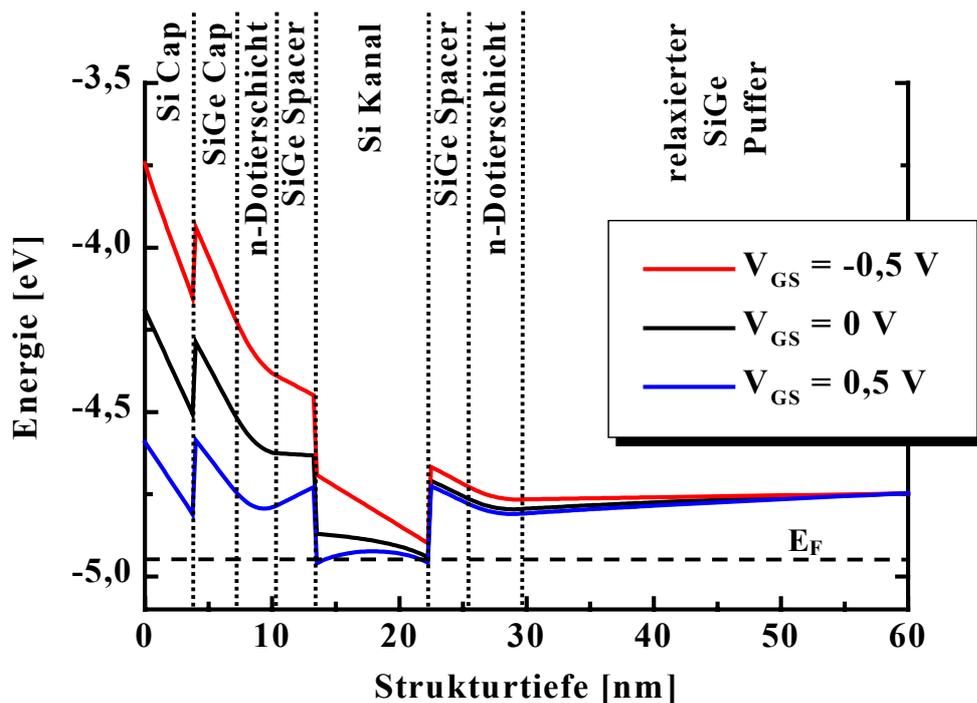


Abbildung 3.7: Charakteristischer Leitungsbandverlauf eines doppelseitig dotierten SiGe-n-Kanal-HFETs für verschiedene Gatepotentiale

Im Banddiagramm liegen sowohl das Leitungsband als auch das Valenzband der Si-Schicht energetisch niedriger als bei den umgebenden SiGe-Schichten. Die Banddiskontinuitäten des sich ergebenden Typ II-Heteroüberganges können mit Hilfe der in Abschnitt 3.1 abgeleiteten Formeln berechnet werden. Die Dotier-

schichten erfahren durch die hohe Donatorkonzentration ebenfalls eine energetische Absenkung Richtung Fermienergie und stellen eine große Anzahl freier Ladungsträger zur Verfügung. Diese Elektronen durchtunneln die dünnen SiGe-Spacer und sammeln sich in dem für sie energetisch günstigeren Potentialtopf des Si-Kanals. Dadurch entsteht ein 2DEG mit hoher Ladungsträgerbeweglichkeit, welches für den Stromfluß im Transistor sorgt. Die Steuerung des Stromflusses erfolgt über einen Metall-Halbleiter-Kontakt. Über die Änderung des Gatepotentials kommt es zu einer Verschiebung der Bandverläufe relativ zum Fermienergie E_F (Abbildung 3.7). Eine Anhebung des Gatepotentials in positive Richtung führt zu einer Absenkung des Leitungs- und Valenzbandes und resultiert in der Erhöhung der Ladungsträgerkonzentration n_s im 2DEG. Bei zunehmend negativem Potential am Gate werden die Bandverläufe angehoben, und die Anzahl der Ladungsträger im Potentialtopf verringert sich oder wird sogar zu Null. Der Zugang der Source/Drain-Anschlüsse zum Si-Kanal wird durch ohmsche Kontakte realisiert, die mittels Ionenimplantation hergestellt werden.

3.3 Eigenschaften des 2DEGs

Durch die Banddiskontinuitäten an den Heteroübergängen zwischen den relaxierten $\text{Si}_{1-x}\text{Ge}_x$ -Schichten und dem pseudomorph integrierten Si-Kanal kommt es im Leitungsband zur Ausbildung eines Potentialtopfes. Liegt die Breite dieses Potentialtopfes in der Größenordnung der de Broglie-Wellenlänge, wird die freie Beweglichkeit der Elektronen in Wachstumsrichtung (z-Richtung) eingeschränkt, und Quantisierungseffekte müssen berücksichtigt werden. Im Leitungsband bilden sich Energiesubbänder (E_0, E_1, \dots), die durch die Elektronen im Potentialtopf besetzt werden. Wenn der Abstand der Subbänder größer ist als die thermische Energie des Kristallgitters $k \cdot T$, quantisiert sich die Bewegung der Ladungsträger in z-Richtung in Abhängigkeit der Bandabstände. Eine uneingeschränkte Bewegungsfreiheit der Elektronen besteht nur noch in der Wachstumsebene, daher die Bezeichnung 2-dimensionales Elektronengas.

3.3.1 Ladungsträgerbeweglichkeit und Streuprozesse

Seit dem experimentellen Nachweis des 2DEGs in verspannten Si/SiGe-Systemen wurden immer höhere Beweglichkeitsrekorde publiziert. Abbildung 3.8 zeigt die Entwicklung der Elektronenbeweglichkeit μ_n in modulationsdotierten SiGe-Heterostrukturen für unterschiedliche Pufferkonzepte im Verlauf der Zeit.

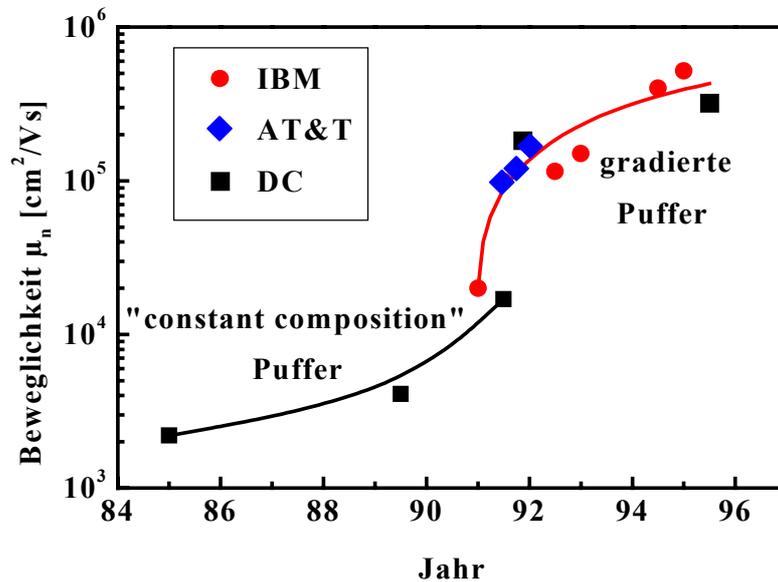


Abbildung 3.8: Entwicklung der Elektronenbeweglichkeit μ_n von n-modulationsdotierten Quantentöpfen (QW) auf unterschiedlichen SiGe-Puffern bei Temperaturen $< 4,2$ K

Durch die Einführung des gradierten Puffers sind Niedertemperaturbeweglichkeiten von bis zu $520000 \text{ cm}^2/\text{Vs}$ bei $0,4$ K möglich geworden [3.21], was nah an die höchsten Werte von GaAs/AlGaAs-Heterostrukturen heranreicht, die bei $1 \cdot 10^6 \text{ cm}^2/\text{Vs}$ liegen [3.22]. Der Rekordwert bei Raumtemperatur beträgt derzeit $2830 \text{ cm}^2/\text{Vs}$ [3.23] und ist damit mehr als 2,5-mal höher als die ermittelten Maximalwerte in konventionellen Si-MOSFETs, die sich im Bereich von $1000 \text{ cm}^2/\text{Vs}$ bewegen [3.24].

Wie läßt sich jedoch diese erhöhte Beweglichkeit der Ladungsträger physikalisch erklären? Zur näheren Analyse des Stromtransportes im pseudomorphen Si-Kanal wird hier auf die Darstellung der Leitungsbandminima als Flächen konstanter Energie zurückgegriffen. Die sich daraus ergebenden sechs Rotationsellipsoide entlang der $\langle 100 \rangle$ -Richtungen der Brillouinzone sind in Abbildung 3.9 für den Fall einer zugverspannten Siliziumschicht dargestellt. Aufgrund der anisotropen Bandstruktur ist die effektive Masse m^* eines Elektrons im 2DEG vom (Quasi-) Impuls \vec{k} abhängig. In Richtung der Längsachse des Ellipsoiden ist dabei die longitudinale Masse m_l wirksam, senkrecht dazu die transversale Masse m_t .

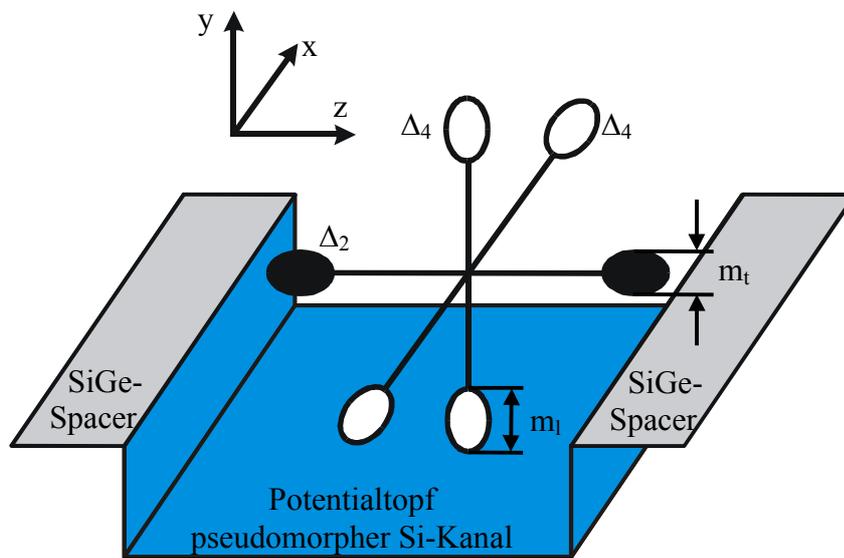


Abbildung 3.9: Ausrichtung der Leitungsbandminima als Flächen gleicher Energie in einem pseudomorphen Si-Kanal

Wie bereits dargestellt, sind die Ladungsträger im 2DEG nur in der xy -Ebene parallel zum Heteroübergang frei beweglich. Außerdem kommt es, wie in Kapitel 3.1 abgeleitet, zu einer Aufspaltung des sechsfach entarteten Leitungsbandminimums in ein zweifach entartetes Δ_2 -Minimum in Wachstumsrichtung und ein vierfach entartetes, energetisch höher liegendes Δ_4 -Minimum in der Wachstumsebene. Aufgrund der höheren Besetzungswahrscheinlichkeit des Δ_2 -Minimums wird für die Ladungsträger im Kanal nur die transversale Masse $m_t = 0,19 m_0$ wirksam. Die daraus resultierende kleinere effektive Masse führt direkt zu einer Erhöhung der Elektronenbeweglichkeit μ_n . Zusätzlich reduziert die Aufspaltung der Leitungsbandminima auch die Wahrscheinlichkeit einer Ladungsträgerstreuung zwischen den Energietälern (Intervalley scattering) [3.25]. Aus Abbildung 3.4 läßt sich ablesen, daß die Energiedifferenz zwischen dem Δ_4 - und dem Δ_2 -Minimum mehr als 200 meV für einen Ge-Gehalt von 30 % beträgt. Bei einer thermischen Energie der Ladungsträger von etwa 25,9 meV und einem nicht zu großen elektrischen Feld in Kanal wird eine Streuung der Elektronen in höhere Minima stark verringert. Weiteren positiven Einfluß auf die Beweglichkeit des 2DEGs hat die räumliche Trennung der Ladungsträger von den Atomrümpfen der ionisierten Donatoren in den Dotierschichten. Dadurch kommt es zu einer drastischen Reduzierung der Coulomb-Wechselwirkung und der damit verbundenen Streuprozesse, deren Anteil über die Dicke der undotierten Spacer eingestellt werden kann.

Den Einfluß der Streuprozesse auf die Beweglichkeit der Ladungsträger beschreibt Gleichung 3.24. Die Zeitkonstante τ steht dabei für die Transportzeit der Ladungsträger zwischen zwei Streuprozessen. Mit Hilfe der „Matthiessen'sche Regel“ (Gleichung 3.25) lassen sich die Relaxationszeiten verschiedener Streuprozesse addieren, wobei der Streumechanismus mit der kleinsten Zeitkonstante den größten Einfluß auf die Beweglichkeit hat.

$$\mu = \frac{q}{m^*} \cdot \tau \quad (3.24)$$

$$\frac{1}{\tau} = \sum_i \frac{1}{\tau_i} \quad (3.25)$$

In Abhängigkeit von der Temperatur wird die Beweglichkeit des 2DEGs von unterschiedlichen Streuprozessen bestimmt. Bei Raumtemperatur und damit dem relevanten Bereich für die meisten Bauelementanwendungen dominiert eindeutig die Phononenstreuung, die aus der Gitterschwingung resultiert. Darüber hinaus beeinflussen die oben beschriebene Coulomb-Wechselwirkung und Hintergrunddotierungen im Kanal die Ladungsträgerbeweglichkeit. Je nach Qualität des verwendeten Puffers und Epitaxieprozesses kommt noch die negative Wirkung von Gitterdefekten in der pseudomorphen Siliziumschicht hinzu. Falls die Donatoratome bei Raumtemperatur nicht vollständig ionisieren, bilden sich parallele Stromkanäle in den Dotierschichten, die eine signifikant reduzierte Beweglichkeit aufweisen und damit die Gesamtbeweglichkeit ebenfalls mindern. Mit abnehmender Temperatur reduzieren sich die Gitterschwingungen und verlieren ihren dominanten Charakter, und als bestimmender Streumechanismus bleibt die Coulomb-Wechselwirkung bestehen. Die parasitären Kanäle „frieren ein“ und haben keinen Anteil mehr am Stromtransport. Statt dessen gewinnt die Hintergrunddotierung des Kanals an Einfluß, ebenso wie Grenzflächenladungen an den Heteroübergängen. Ein weiterer Streubeitrag ergibt sich durch die Grenzflächenrauigkeit des Kanals. Diese resultiert aus dem Abbau der Gitterverspannungen im Puffer und steigt mit zunehmendem Germaniumgehalt an.

3.3.2 Ladungsträgerdichte

Da die Beweglichkeit der Elektronen eng mit der Ladungsträgerdichte im Kanal gekoppelt ist, hat für reale Bauelementanwendungen das Produkt aus μ_n und n_s hohe Aussagekraft über die Qualität der Heterostruktur. Im folgenden soll deshalb hergeleitet werden, wie die Ladungsträgerkonzentration n_s im Kanal analytisch bestimmbar ist.

Die Anzahl der besetzbaren Zustände im Quantentopf der pseudomorphen Si-Schicht und damit die Ladungsträgerdichte des 2DEGs berechnet sich im thermischen Gleichgewicht nach Gleichung 3.26. Dabei wird das Produkt aus der Zustandsdichte der Subbänder D_n^{2D} und der Fermi-Dirac Verteilungsfunktion $F(E)$ über alle Subbänder integriert.

$$n_s = \int_{E_L}^{\infty} D_n^{2D} \cdot F(E) dE \quad \text{mit} \quad (3.26)$$

$$F(E) = \frac{1}{1 + \exp\left(\frac{E - E_F}{kT}\right)} \quad (3.27)$$

Die Zustandsdichte D_n^{2D} für zweidimensionale Systeme ist unabhängig von der Energie und für jedes Subband konstant [3.28].

$$D_n^{2D} = \frac{m^*}{\pi \cdot \hbar^2} \quad \text{für } E \geq E_n \quad (3.28)$$

Hieraus ergibt sich die Elektronenkonzentration $n_{s,n}$ für das Subband mit dem Energieniveau E_n .

$$n_{s,n} = D_n^{2D} \cdot k \cdot T \cdot \ln \left[1 + \exp\left(\frac{E_F - E_n}{k \cdot T}\right) \right] \quad (3.29)$$

Die Subbandenergie E_n für einen dreiecksförmigen Potentialtopf berechnet sich nach Gleichung 3.30.

$$E_n = \left(\frac{\hbar^2}{2 \cdot m^*} \right)^{1/3} \cdot \left[\frac{3}{2} \cdot \frac{q^2 \pi}{\epsilon} \cdot \left(n + \frac{3}{4} \right) \right]^{2/3} \quad (3.30)$$

Die Ladungsträgerdichte $n_s(z)$ am Ort z folgt aus der Summation über die Energieniveaus der Subbänder n .

$$n_s(z) = \sum_n |\psi_n(z)|^2 \cdot n_{s,n} \quad (3.31)$$

Um neben der Elektronenkonzentration im 2DEG Informationen über die Bandverläufe, die Energieniveaus und die Wellenfunktionen der Ladungsträger zu erlangen, ist die selbstkonsistente Lösung der Schrödinger- und Poisson-Gleichung unter den gegebenen Randbedingungen erforderlich. Die sich einstellenden Bandverbiegungen resultieren aus dem elektrostatischen Potential $V(z)$, welches nur von der z -Koordinate abhängt. Da des weiteren nur Zustände im thermischen Gleichgewicht betrachtet werden, reduziert sich die Schrödinger-Gleichung auf den eindimensionalen, stationären Fall.

$$\frac{\hbar^2}{2 \cdot m^*} \cdot \frac{d^2\psi(z)}{dz^2} + [E_n - V(z)] \cdot \psi(z) = 0 \quad (3.32)$$

Die Wellenfunktion $\psi(z)$ stellt die Lösung der Schrödinger-Gleichung dar und beschreibt quantenmechanisch die Ladungsträgerverteilung. Die Eigenwerte E_n der Gleichung ergeben im Fall des betrachteten Potentialtopfes die Energieniveaus der Subbänder.

$$\frac{d^2V(z)}{dz^2} = -\frac{\rho(z)}{\epsilon} \quad \text{mit} \quad (3.33)$$

$$\rho(z) = q \cdot (N_A^-(z) + n_s(z) - N_D^+(z)) \quad (3.34)$$

Die Poisson-Gleichung liefert als Lösung das elektrostatische Potential in Abhängigkeit der auftretenden Raumladungszonen.

Die Ermittlung der Bandverläufe und Ladungsträgerdichten zur Optimierung bestimmter Schichtkonfigurationen erfolgte im Rahmen dieser Arbeit mit einem Simulationsprogramm [3.26], welches den vorgestellten Formalismus anwendet. Nach der Vorgabe der Schichtstruktur und der Randbedingungen in Form von Oberflächenpotential und Temperatur wird eine Startpotentialverteilung $V_0(z)$ ermittelt. Die dazu gehörige Ladungsträgerverteilung und der Bandverlauf werden anschließend unter Verwendung der Fermi-Dirac-Statistik und der Schrödinger-Gleichung bestimmt. Die daraus resultierenden, modifizierten Potentialverhältnisse werden mit dem Ergebnis der vorherigen Iteration verglichen. Dieser

numerische Ablauf wiederholt sich solange, bis eine hinreichend gute Anpassung für $V(z)$, $\psi(z)$ und E_n erzielt wurde. Der Abbruch des Algorithmus kann durch ein vorgegebenes Fehlerkriterium festgelegt werden.

3.4 Ladungsträgersteuerung im Gleichstromfall

Zur Steuerung der Ladungsträgerkonzentration im 2DEG des n-Kanal-HFETs kommt ein Schottky-Gate zum Einsatz, welches auf die Silizium-Cap-Schicht zwischen Source- und Drain-Elektrode aufgebracht wird. Die Wirkung des Gatepotentials auf die Ladungsträger im Kanal läßt sich gut mit dem „Charge-Control“-Modell beschreiben. Dieses Konzept wurde ursprünglich von Johnson und Rose im Jahre 1959 entworfen [3.27] und später von Delagebeaudeuf und Linh für die Anwendung auf modulationsdotierte III/V-HEMTs modifiziert [3.28], [3.29]. Im folgenden soll dieser Modellansatz Verwendung finden, um für den Fall eines vorderseitig dotierten SiGe-HFETs die Ladungsträgerdichte im 2DEG zu bestimmen.

Durch den Schottky-Kontakt an der Halbleiteroberfläche ergibt sich neben der Verarmungszone am Heteroübergang eine zweite Raumladungszone direkt unterhalb des Gates. Wenn die vorderseitige Dotierschicht dünn ist oder eine entsprechend negative Gatespannung angelegt wird, überschneiden sich die beiden Zonen, und der Halbleiterbereich zwischen Gateelektrode und Si-Kanal ist komplett verarmt. In diesem Fall muß die Ladungsträgerdichte n_d , um die sich die Dotierschicht verarmt hat, gleich sein mit der Anzahl an besetzten Zuständen n_s , die sich im Potentialtopf des Kanals befindet.

$$n_d = \frac{\epsilon}{q \cdot d} \cdot \left(V_G - \Phi_B + \frac{\Delta E_L}{q} - \frac{E_F}{q} + \frac{q \cdot N_D \cdot (d - d_i)^2}{2 \cdot \epsilon} \right) = \quad (3.35)$$

$$n_s = \sum_n D_n^{2D} \cdot k \cdot T \cdot \ln \left[1 + \exp \left(\frac{E_F - E_n}{k \cdot T} \right) \right]$$

Dabei ist N_D die Dotierstoffkonzentration der SiGe-Schicht, d die Gesamtdicke aller Schichten zwischen der Gateelektrode und dem Kanal, d_i die Dicke des Spacers zwischen der Dotierschicht und dem Kanal, ΔE_L die Banddiskontinuität des Leitungsbandes, E_F das Fermi-niveau, Φ_B die Schottky-Barriere des Metallgates und V_G das Gatepotential.

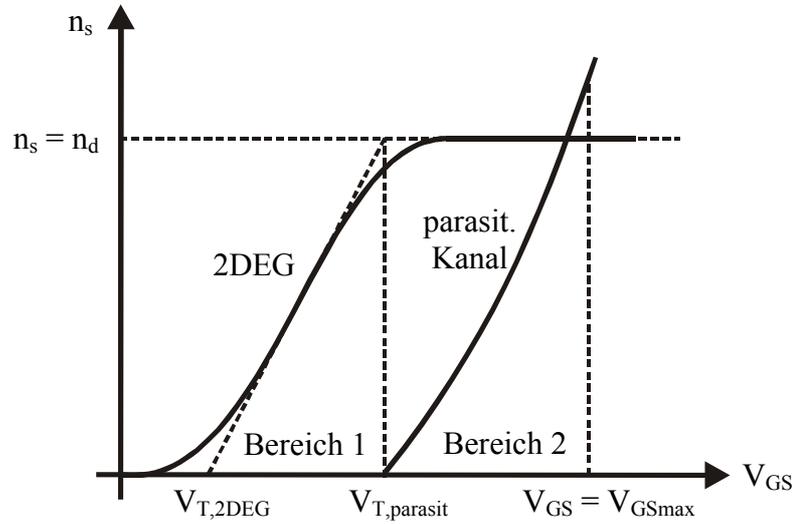


Abbildung 3.10: Qualitative Darstellung der Ladungsträgerkonzentrationen im Potentialtopf (2DEG) und im parasitären Kanal in der Dotierschicht in Abhängigkeit von der Gatespannung V_{GS}

Abbildung 3.10 zeigt den schematischen Verlauf der Ladungsträgerkonzentration im 2DEG und im parasitären Kanal der Dotierschicht in Abhängigkeit von der Gatespannung. Danach sind zwei Bereiche zu unterscheiden.

Bereich 1: Für Gatespannungen $V_{T,2DEG} \leq V_{GS} \leq V_{T,parasit}$ ist die Dotierschicht vollständig verarmt, und nur die Elektronen im Quantentopf tragen zum Stromfluß bei. Die Pinch-off Spannung für das 2DEG $V_{P,2DEG}$ ergibt sich aus Gleichung (3.36), dabei ist d_d die Dicke der Dotierschicht, d_{Cap} beschreibt die vertikale Distanz zwischen der Dotierschicht und dem Schottky-Gate, und Δd erfaßt den Abstand des Ladungsträgerschwerpunktes (2DEG) zum Potentialtopfrand.

$$V_{P,2DEG} = \frac{q \cdot n_d \cdot (d + \Delta d)}{\epsilon} \quad \text{mit} \quad d = d_d + d_{Cap} + d_i \quad (3.36)$$

Bereich 2: Für Gatespannungen $V_{GS} > V_{T,parasit}$ hat zusätzlich zum 2DEG im Potentialtopf der parallele Strompfad in der nicht vollständig verarmten Dotierschicht einen Anteil am Gesamtstrom. Die Pinch-off Spannung $V_{P,parasit}$ für diesen Fall ist:

$$V_{P,parasit} = \frac{q \cdot N_D \cdot (d_d + d_{Cap})^2}{2 \cdot \epsilon} \quad (3.37)$$

Für alle weiteren Betrachtungen werden die Bedingungen in Bereich 1 angenommen. Die Ladungsträgerdichte im 2DEG in Abhängigkeit der Gatespannung ergibt sich zu:

$$n_s = \frac{\varepsilon}{q \cdot (d_d + d_{\text{Cap}} + d_i + \Delta d)} \cdot (V_{\text{GS}} - V_T) \quad (3.38)$$

Dieser Modellansatz basiert auf dem Prinzip eines Plattenkondensators. Dabei werden die verarmten Deckschichten als Dielektrikum angenommen, und der Gatekontakt und das 2DEG bilden jeweils die Elektroden. Die Kapazität des Kondensators, normiert auf die Gatefläche, beträgt damit:

$$C_0 = \frac{\varepsilon_0 \cdot \varepsilon_r}{d_d + d_{\text{Cap}} + d_i + \Delta d} \quad (3.39)$$

V_T wird als Schwellspannung bezeichnet und gibt das Gatepotential an, bei dem gerade kein Stromfluß mehr durch den Kanal erfolgt.

$$V_T = V_{T,2\text{DEG}} = \Phi_B - \frac{(\Delta E_L - E_F)}{q} - V_{P,\text{parasit}} \quad (3.40)$$

Im folgenden soll eine Strom-Spannungs-Beziehung für die Ladungsträgersteuerung im Kanal hergeleitet werden. Das dabei zugrunde gelegte Modell von Das und Roszak [3.30] kombiniert den oben betrachteten „Charge-Control“-Ansatz mit einem Konzept zur Berücksichtigung der Ladungsträgergeschwindigkeitssättigung [3.31], [3.32]. Dabei gelten die folgenden Randbedingungen [3.33]:

- Die Dotierbereiche sind vollständig verarmt. Über das Gate erfolgt nur noch eine Modulation des 2DEGs. Eine Anreicherung von Ladungsträgern in den parallelen Dotierschichten wird nicht berücksichtigt.
- Es gilt Stromkontinuität über den ganzen Verlauf des Kanals, d. h. zusätzliche, parasitäre Stromflüsse über die Schottky-Barriere des Gates oder Sperrströme der pn-Übergänge treten nicht auf. Außerdem werden Generations- und Rekombinationsprozesse vernachlässigt.
- I_D wird von Driftströmen zwischen Source und Drain dominiert. Diffusionsströme sind vernachlässigbar.

- Die auftretenden lateralen Feldstärken sind sehr viel kleiner als die vertikalen Felder in Source-Drain-Richtung $|E_x| \gg |E_y|$ („Gradual Channel“-Näherung)
- Ballistischer Ladungsträgertransport, Velocity Overshoot und andere Kurzkanaleffekte werden nicht berücksichtigt.
- Substratleckströme über nicht ideal isolierende Pufferschichten werden vernachlässigt.
- Der Sourcekontakt liegt auf Massepotential, und es gilt $V_{DS} > 0$ V.

Um den gewünschten Stromfluß zwischen Source- und Drain-Elektrode zu erzeugen, wird zusätzlich zur Steuerspannung V_{GS} nun die Beschleunigungsspannung $V(x)$ in x-Richtung eingeführt. Die Ladungsträgerkonzentration an einem beliebigen Punkt x entlang des Kanals ergibt sich aus Gleichung (3.41).

$$Q_s(x) = q \cdot n_s(x) = C_0 \cdot (V_{GS} - V_T - V(x)) \quad (3.41)$$

Im linearen Anlaufbereich verhält sich die Strom-Spannungs-Beziehung des Transistors wie bei einem Widerstand.

$$j_D = \sigma \cdot E = q \cdot n_s(x) \cdot \mu_n \cdot E(x) \quad (3.42)$$

$$I_D = j_D \cdot W_G = Q_s(x) \cdot v_d(x) \cdot W_G \quad (3.43)$$

Dieser Zusammenhang ist nur gültig, solange eine lineare Beziehung über die Niederfeldbeweglichkeit μ_0 zwischen der Driftgeschwindigkeit v_d und dem elektrischem Feld E besteht. Ab einer kritischen Feldstärke E_{krit} tritt eine Sättigung der Geschwindigkeit ein, die sowohl für unverspanntes als auch für verspanntes Silizium bei $1 \cdot 10^7$ cm/s liegt [3.34], [3.35]. Diese Abhängigkeit der Ladungsträgerdriftgeschwindigkeit vom angelegten elektrischen Feld entlang des Kanals ist von elementarer Bedeutung für die Modellierung der Strom-Spannungs-Charakteristik und wird durch die folgende Beziehung berücksichtigt.

$$v(E) = \frac{\mu_0 \cdot E}{\left(1 + \frac{E}{E_{krit}}\right)} \quad \text{mit} \quad v_{sat} = \mu_0 \cdot E_{krit} \quad (3.44)$$

$$I_D = \frac{W_G \cdot \mu_0 \cdot C_0 \cdot (V_{GS} - V_T - V(x)) \cdot \frac{dV}{dx}}{1 + \frac{\mu_0}{v_{sat}} \cdot \frac{dV}{dx}} \quad (3.45)$$

Aufgrund der angenommenen Stromkontinuität entlang des Kanals kann der Strom zwischen $x = 0$ mit $V(x = 0) = 0$ und $x = L_G$ mit $V(x = L_G) = V_{DS}$ integriert werden, und es ergibt sich die Strom-Spannungs-Beziehung für den linearen Anlaufbereich.

$$I_D = \frac{W_G \cdot \mu_0 \cdot C_0 \cdot \left((V_{GS} - V_T) \cdot V_{DS} - \frac{1}{2} V_{DS}^2 \right)}{L_G + \frac{\mu_0}{v_{sat}} \cdot V_{DS}} \quad (3.46)$$

Differentiation und Nullsetzen von Gleichung (3.46) liefert die Spannung $V_{DS,sat}$, bei der die Sättigung des Drainstromes eintritt.

$$V_{DS,sat} = \sqrt{\left(\frac{v_{sat} \cdot L_G}{\mu_0} \right)^2 + \frac{2 \cdot v_{sat} \cdot L_G}{\mu_0} \cdot (V_{GS} - V_T) - \frac{v_{sat} \cdot L_G}{\mu_0}} \quad (3.47)$$

Aus (3.46) und (3.47) folgt schließlich die Abhängigkeit des Drainstromes von der Gatespannung im Sättigungsbereich.

$$I_{Dsat} = \frac{W_G \cdot L_G \cdot C_0 \cdot v_{sat}^2}{2 \cdot \mu_0} \left(\sqrt{1 + \frac{2 \cdot \mu_0}{v_{sat} \cdot L_G} \cdot (V_{GS} - V_T)} - 1 \right)^2 \quad (3.48)$$

Eine wichtige Größe zur Charakterisierung der Transistoreigenschaften ist die Steilheit g_m . Sie beschreibt die Änderung des Ausgangsstromes I_D mit der Eingangsspannung V_{GS} und gibt damit Auskunft über die Gleichstromverstärkung des Bauelementes. Die Gleichungen zur Bestimmung der Steilheit im Anlaufbereich (3.49) und im Sättigungsbereich (3.50) sind nachfolgend dargestellt.

$$g_m = \left. \frac{dI_D}{dV_{GS}} \right|_{V_{DS}=\text{konst}} = \frac{W_G \cdot \mu_0 \cdot C_0 \cdot V_{DS}}{L_G + \frac{\mu_0}{v_{\text{sat}}} \cdot V_{DS}} \quad (3.49)$$

$$g_m = \left. \frac{dI_D}{dV_{GS}} \right|_{V_{DS}=\text{konst}} = W_G \cdot C_0 \cdot v_{\text{sat}} \left(1 - \frac{1}{\sqrt{1 + \frac{2 \cdot \mu_0 \cdot q \cdot n_s}{L_G \cdot C_0 \cdot v_{\text{sat}}}}} \right) \quad (3.50)$$

Das bisherige Modell betrachtet allein den intrinsischen Transistor. Der Stromfluß I_D durch die Widerstände der Source/Drain-Kontakte und die sich anschließenden Bahngebiete im ungesteuerten Kanalgebiet sorgt jedoch für ohmsche Spannungsabfälle, die nicht vernachlässigt werden können. Sie reduzieren die effektiv am intrinsischen Transistor anliegenden Spannungen $V_{DS,\text{intr}}$ und $V_{GS,\text{intr}}$ zu:

$$V_{DS,\text{intr}} = V_{DS,\text{extr}} - (R_s + R_d) \cdot I_D \quad (3.51)$$

$$V_{GS,\text{intr}} = V_{GS,\text{extr}} - R_s \cdot I_D \quad (3.52)$$

Durch das Einsetzen dieser Ausdrücke in die obigen Stromformeln (3.46) und (3.48) ergeben sich komplexe, quadratische Gleichungen für I_D , die am besten mit numerischen Verfahren gelöst werden. Aus der modifizierten Formel für den Drainstrom I_D im Sättigungsbereich folgt die extrinsische Steilheit $g_{m,\text{extr}}$ des Transistors mit $a = (W_G \cdot C_0 \cdot v_{\text{sat}})$ [3.32].

$$\frac{1}{g_{m,\text{extr}}} = R_s + \frac{1}{W_G \cdot C_0 \cdot v_{\text{sat}}} + \frac{1}{\sqrt{\frac{2 \cdot W_G \cdot \mu_0 \cdot C_0 \cdot I_{D,\text{sat}} \cdot (1 + a \cdot R_s)}{L_G}}} \quad (3.53)$$

Die bisher abgeleitete Strom-Spannungs-Beziehung geht davon aus, daß der Sättigungsstrom $I_{D,\text{sat}}$ ab einem bestimmten Wert $V_{DS,\text{sat}}$ konstant ist. Bei realen Transistoren ist jedoch auch im Sättigungsbereich ein Anstieg des Stromes aufgrund der Kanallängenmodulation zu erkennen. Dabei wandert der Abschnürbereich

des 2DEGs mit zunehmender Drainspannung in Richtung Source und verringert dadurch die effektive Kanallänge. Ein Maß für die Stromänderung im Sättigungsbereich ist der Ausgangsleitwert g_{ds} , der sich auf die effektive Gatelänge $L_{G,eff}$ bezieht.

$$g_{ds} = \left. \frac{dI_D}{dV_{DS}} \right|_{V_{GS}=\text{konst}} = I_{Dsat} \sqrt{\frac{1}{1 + \frac{2 \cdot \mu_0 \cdot (V_{GS,intr} - V_T)}{L_{G,eff} \cdot v_{sat}}}} \cdot \left(\frac{-1}{L_{G,eff}} \cdot \frac{dL_{G,eff}}{dV_{DS}} \right) \quad (3.54)$$

Eine explizite Darstellung von g_{DS} in Abhängigkeit von V_{DS} und V_{GS} ist aufgrund des zweidimensionalen Stromverlaufes und der komplizierten Ladungsträgerverteilung im Drainbereich nicht möglich.

3.5 Kleinsignalersatzschaltbild

Der spezielle Heterostrukturaufbau des SiGe-HFETs und die daraus resultierenden physikalischen Vorteile für die Beweglichkeit der Ladungsträger prädestinieren dieses Bauelement für den Einsatz in Hochfrequenzschaltungen. Dabei wird dem Gleichstromarbeitspunkt des Transistors ein hochfrequentes Signals überlagert. Für kleine Amplituden der Eingangsgröße kann die nichtlineare Kennlinie im betrachteten Punkt durch ihre Tangente angenähert werden (Kleinsignalbetrieb). Da die Tangente eine lineare Beziehung zwischen Strom und Spannung herstellt, ist das Bauelementverhalten im Arbeitspunkt linearisiert. Diese Voraussetzung gestattet die Modellierung der Transistorcharakteristik durch ein Ersatzschaltbild, welches durch Knoten- und Maschenregeln beschrieben werden kann. In Abbildung 3.11 ist ein allgemein gültiges Kleinsignalersatzschaltbild für einen HFET dargestellt [3.36], [3.37]. Zur Unterscheidung zwischen Kleinsignal- und Gleichstromgrößen werden die Komponenten durch Kleinbuchstaben indiziert.

Das Kleinsignalersatzschaltbild ist ein wichtiges Hilfsmittel, um die elektrischen Eigenschaften des Transistors zu optimieren. Mit Hilfe von Gleichstrom- und Hochfrequenzmessungen können die Größen der Ersatzschaltbildelemente für einen bestimmten Arbeitspunkt extrahiert werden. Jedes Element hat dabei einen physikalischen Bezug zum gegebenen Transistoraufbau und den Bedingungen, unter denen das Bauelement betrieben wird. Dadurch ist es möglich Erkenntnisse über die optimale Schichtstruktur, die Dimensionierung des Transistors oder

notwendige Technologieminifikationen zu erlangen. Die physikalische Zuordnung der Ersatzschaltbildelemente zu der Transistorstruktur ist aus Abbildung 3.12 ersichtlich.

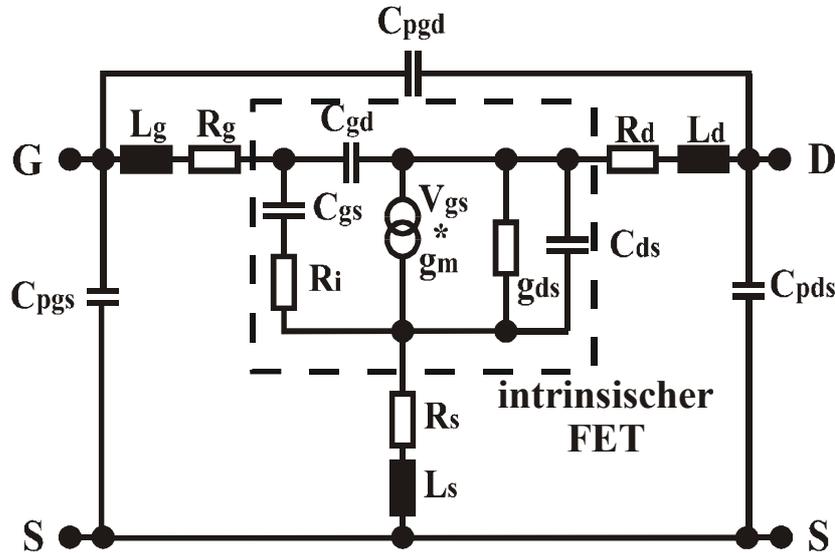


Abbildung 3.11: Kleinsignalersatzschaltbild eines HFET [3.36], [3.37]

Die Elemente des Ersatzschaltbildes haben folgende Bedeutung.

- Die spannungsgesteuerte Stromquelle zwischen Source und Drain wird durch die Kleinsignalsteilheit g_m und die Zeitkonstante τ charakterisiert und beschreibt die Änderung des Kanalstromes mit der Eingangsspannung V_{gs} . Dieser intrinsische Verstärkungsmechanismus erfolgt nicht verzögerungsfrei, sondern wird bestimmt durch die Zeit τ , die zur Neuverteilung der Ladung als Reaktion auf eine Gatespannungsänderung benötigt wird.
- Der Ausgangsleitwert g_{ds} ist ein Maß für die Änderung des Kanalstromes in Abhängigkeit von der Ausgangsspannung zwischen Source und Drain.
- Die Kapazitäten C_{gs} und C_{gd} sind wie folgt definiert:

$$C_{gs} = \left. \frac{\partial Q_s}{\partial V_{gs}} \right|_{V_{ds}=\text{konst}} \quad C_{gd} = \left. \frac{\partial Q_s}{\partial V_{gd}} \right|_{V_{gs}=\text{konst}}$$

Sie repräsentieren die Modulation der Ladung in der Raumladungszone unterhalb des Gates durch die entsprechende Spannungsänderung. C_{gd} führt bei

einem Verstärkerbetrieb des Transistors in Sourceschaltung zu einer kapazitiven Rückkopplung zwischen Ein- und Ausgang und beeinflusst dadurch die Strom- und Leistungsverstärkung des Bauelementes bei hohen Frequenzen.

- C_{ds} berücksichtigt geometrische kapazitive Effekte zwischen Source- und Drainelektrode und kann als arbeitspunktunabhängig betrachtet werden.
- Der Widerstand R_i ist Bestandteil eines verteilten RC-Gliedes, bestehend aus R_i und C_{gs} , welches die Kopplung zwischen leitendem Kanal und Gateelektrode modelliert. Die Einführung von R_i erleichtert die Anpassung der Ersatzschaltparameter an die gemessenen Hochfrequenzwerte, führt jedoch zu einem komplexeren Ersatzschaltbild.
- R_s und R_d beschreiben die ohmschen Kontaktwiderstände der Source- und Draingebiete und die anteiligen Bahnwiderstände des Siliziumkanals zwischen ohmschen Kontakt und 2DEG. Durch R_g wird der Metallwiderstand des Schottky-Gates berücksichtigt.

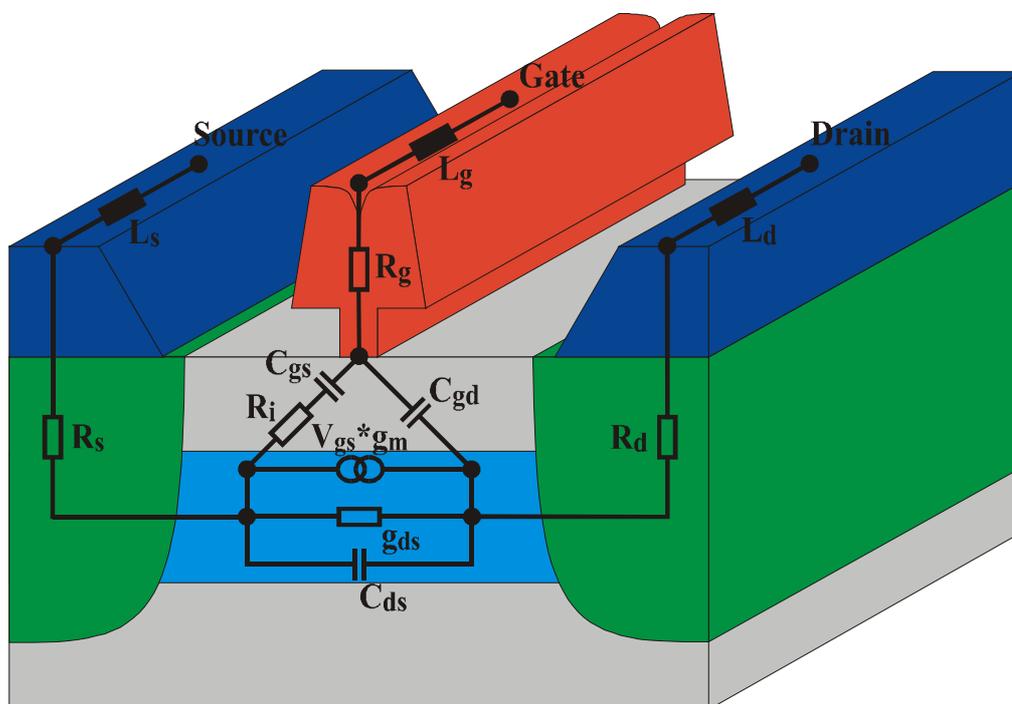


Abbildung 3.12: Physikalische Zuordnung der Ersatzschaltbildelemente zur Bauelementstruktur

Die Kenntnis der Ersatzschaltbildparameter ermöglicht die Abschätzung der Hochfrequenzeigenschaften des Transistors. Zur Charakterisierung werden Grenzfrequenzen verwendet, die Aufschluß über die Strom- und Leistungsverstärkung des Bauelementes bei hohen Frequenzen geben. Eine wichtige Größe, die direkt aus dem Ersatzschaltbild berechnet werden kann, ist die Transitfrequenz f_T . Sie gibt an, wann die Stromverstärkung \underline{h}_{21} des Transistors in Source-schaltung bei ausgangsseitigem Kurzschluß den Wert 1 annimmt. Für den intrinsischen Transistor gilt unter der Voraussetzung, daß $R_i \ll 1/\omega \cdot C_{gs}$ ist:

$$|\underline{h}_{21}| = \left. \frac{I_d}{I_g} \right|_{V_{ds}=0} \cong \frac{|V_{gs} \cdot g_m|}{|j \cdot \omega \cdot (C_{gs} + C_{gd}) \cdot V_{gs}|} = 1 \quad (3.55)$$

$$f_T = \frac{g_m}{2 \cdot \pi \cdot (C_{gs} + C_{gd})} \quad (3.56)$$

Bei physikalischer Abschätzung über die Laufzeit der Ladungsträger von Source nach Drain folgt:

$$f_T \approx \frac{1}{2 \cdot \pi \cdot \tau_T} = \frac{1}{2 \cdot \pi} \cdot \frac{L_G}{v_{sat}} \quad (3.57)$$

Unter Einbeziehung der parasitären Widerstände ergibt sich Gleichung (3.58), die eine Näherung für die extrinsische Grenzfrequenz der Stromverstärkung bereitstellt [3.38].

$$f_{T,extr} = \frac{g_m}{2 \cdot \pi \cdot \left[(C_{gs} + C_{gd}) \cdot \left(1 + \frac{R_s + R_d}{R_{ds}} \right) + C_{gd} \cdot g_m \cdot (R_s + R_d) \right]} \quad (3.58)$$

Als Maß für die Leistungsverstärkung des Transistors gilt die Grenzfrequenz f_{max} . Sie bestimmt den Übergang vom aktiven zum passiven Bauelement und damit die Grenze, bis zu der eine unilaterale Leistungsverstärkung vorhanden ist.

$$f_{\max} = \frac{f_T}{\sqrt{4 \cdot g_{ds} \cdot (R_i + R_g + R_s) + 2 \cdot \frac{C_{gd}}{C_{gs}} \cdot \left(\frac{C_{gd}}{C_{gs}} + g_m \cdot (R_i + R_s) \right)}} \quad (3.59)$$

Auf die Bestimmung der Grenzfrequenzen aus den gemessenen S-Parametern der Hochfrequenzcharakterisierung wird in Kapitel 7 noch detailliert eingegangen.

Die dargestellten Beziehungen zwischen den extrinsischen Grenzfrequenzen und den Elementen des Ersatzschaltbildes ermöglichen, mit Kenntnis der physikalischen Bedeutung der Elemente, die gezielte Optimierung des Transistors für Hochfrequenzanwendungen. Dabei ergeben sich zwei Optimierungsansätze, die teilweise miteinander gekoppelt sind. Zum einen kann durch die Dimensionierung der Schichtkonfiguration eine vertikale Optimierung der Transistorstruktur erfolgen, die primär mit Hilfe der Epitaxie durchführbar ist. Zum anderen spielt die Auswahl des Technologiekonzeptes und die Gestaltung des Layouts eine entscheidende Rolle bei der lateralen Optimierung des Bauelementes. Im folgenden Kapitel wird deshalb zuerst die Konfiguration des Schichtsystems diskutiert und anschließend in den weiteren Kapiteln die prozess technologische Realisierung des Transistors eingehender betrachtet.

Kapitel 4

Dimensionierung und Konfiguration des SiGe-Heteroschichtsystems

Im folgenden soll die Dimensionierung und Auslegung des SiGe-Heterostruktursystems optimiert werden. Ziel ist es dabei, eine Schichtkonfiguration zu finden, die dem Transistor eine möglichst ideale und leistungsfähige Strom-Spannungs-Charakteristik verleiht und gleichzeitig das Potential als Hochfrequenzbauelement unterstreicht. An dieser Stelle sei darauf hingewiesen, daß alle Schichtsysteme, die im Rahmen dieser Arbeit für Analysen und Bauelementprozessierungen zur Verfügung standen, mit Hilfe der Molekularstrahlepitaxie (MBE) hergestellt wurden. Zur Charakterisierung der typischen Schichteigenschaften kamen diverse Analyseverfahren zur Anwendung. Die Morphologie und die strukturellen Beschaffenheiten der Schichten wurden mit der Transmissions-Elektronen-Mikroskopie (TEM), der Sekundärionen-Massenspektroskopie (SIMS) und der Röntgendiffraktometrie (XRD) untersucht. Um die elektrischen Eigenschaften der Strukturen klassifizieren zu können, kamen integrale und magnetfeldabhängige Hallmessungen, konventionelle Schichtwiderstandsmessungen und spezielle Extraktionsmethoden aus der Bauelementcharakterisierung zum Einsatz.

4.1 Pufferschichten

Der SiGe-Puffer ist die Grundlage für die aktiven Schichten, die ihrerseits die elektrischen Eigenschaften des Bauelementes bestimmen. Die Aufgabe des Puffers ist, es den erforderlichen Ge-Gehalt als eine Art virtuelles Substrat bereitzustellen und sowohl in Morphologie als auch in den strukturellen Eigenschaften die optimale Basis für die nachfolgende Schichtkonfiguration zu bieten. Die Realisierung von SiGe-n-Kanal-HFETs erfordert einen Ge-Gehalt im Substrat von bis zu 50 %, um einen entsprechend ausgeprägten Quantentopf im Leitungs-

band umsetzen zu können. Das epitaktische Wachstum von SiGe-Schichten auf einem Silizium-Standardwafer führt bei überkritischen Schichtdicken zur Relaxation der Gitterfehlpassung durch Bildung von Defektstrukturen im Kristall. Kommt es dabei zur Entstehung von senkrecht aufsteigenden Versetzungslinien („Threading dislocations“), die ihren Ursprung im Puffer haben und sich bis zum Kanal der aktiven Schicht fortpflanzen, hat dies einen reduzierenden Einfluß auf die Beweglichkeit der Ladungsträger. „Threading dislocations“ resultieren meistens aus engmaschigen Defektsegmenten am Übergang zwischen Si-Substrat und SiGe-Epitaxieschicht. Deshalb zeichnet sich ein gutes Pufferkonzept durch ein vergrabenes Defektstrukturnetzwerk aus, das mit zunehmender Pufferdicke eine wachsende Segmentgröße aufweist und vornehmlich durch lange, parallel zum Interface verlaufende Versetzungslinien beschrieben wird. Zwar begünstigen höhere Epitaxietemperaturen die Ausbildung von langen Defekten, aber gleichzeitig fördern sie auch das dreidimensionale Wachstum, was zu einer starken Oberflächenrauigkeit führt. Bestimmte Störungen der regelmäßigen Gitterstruktur resultieren in Oberflächenstufen und führen dadurch zu einem anisotropen Schichtwachstum. Dieses spiegelt sich in einem typischen Oberflächenmuster wider, was in Form einer regelmäßigen Kreuzschraffierung, dem sogenannten „Cross-hatch“ auftritt.

Als Voraussetzungen für eine leistungsfähige aktive Schichtfolge sollte der SiGe-Puffer folgende Qualitätsmerkmale aufweisen:

- präzise Annäherung an den geforderten Ge-Gehalt
- hohen Relaxationsgrad ($> 90\%$)
- geringe Dichte an „Threading Dislocations“ ($< 1 \cdot 10^5 \text{ cm}^{-2}$)
- minimale Oberflächenrauigkeit

4.1.1 Klassische Pufferkonzepte

Der einfachste Weg, einen SiGe-Puffer zu erzeugen, ist das Wachstum einer überkritischen SiGe-Schicht mit dem gewünschten Ge-Gehalt x direkt auf ein Si-Substrat. Dieser sogenannte „constant-composition buffer“ (CCB) kann, je nach Ge-Gehalt, mit einer Dicke von nur 500 nm bei einem Relaxationsgrad von über 90 % realisiert werden [4.1]. Untersuchungen von Pufferquerschnitten mit der Transmissions-Elektronen-Mikroskopie (XTEM) zeigen jedoch mit 10^9 - 10^{11} cm^{-2} eine hohe Dichte an aufsteigenden Versetzungslinien [4.2]. Die Erhöhung der Epitaxietemperatur zur Reduzierung dieser Defekte ist nur im kleinen

Rahmen möglich, da das Einsetzen eines dreidimensionalen Schichtwachstums die Oberflächenqualität weiter verschlechtern würde.

Mit der Einführung des gradierten Pufferkonzeptes konnte die Kristallqualität in nahezu allen Aspekten gegenüber dem Puffer mit konstanter Komposition verbessert werden [4.3]. Der Gesamtpuffer teilt sich dabei in einen Bereich mit einem ansteigenden Ge-Gehalt und einen Teil mit einer konstanten Ge-Konzentration. Das Grundprinzip dieses Konzeptes ist die lineare oder stufenweise Anhebung des Ge-Gehaltes über einen weiten Wachstumsbereich, bei gleichzeitiger Absenkung der Epitaxietemperatur (Abbildung 4.1).

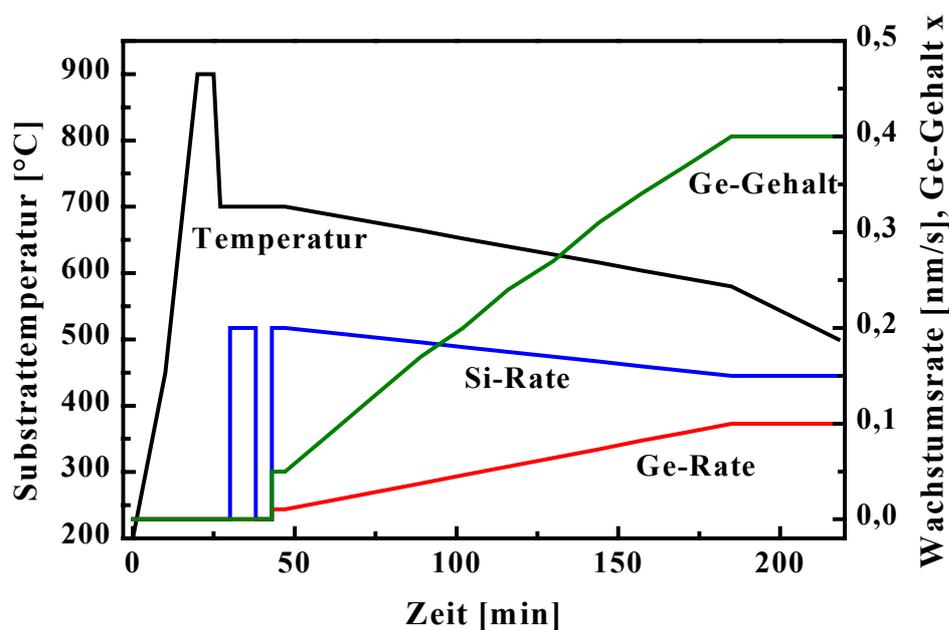


Abbildung 4.1: Zeitlicher Verlauf der Wachstumsparameter für einen gradierten SiGe-Puffer mit 40 % Ge-Gehalt [4.4]

Der Epitaxieprozeß beginnt mit einer Oxiddesorption bei 900 °C und dem Wachstum einer dünnen Nukleationsschicht aus Silizium. Die SiGe-Abscheidung startet üblicherweise mit einem Ge-Gehalt von 5 % und einer Substrattemperatur von 690 °C - 740 °C. In Abhängigkeit der Epitaxiezeit erhöht sich die Ge-Rate linear, was in einem Anstieg des Ge-Gehaltes von 10 % - 40 % pro μm Schichtwachstum resultiert. Die Substrattemperatur wird gleichzeitig um 2 - 3 °C/% verringert, um dreidimensionales Wachstum zu verhindern. Nach dem Erreichen des Ge-Endwertes beginnt die Abscheidung der etwa 500 nm dicken Schicht mit

konstantem Ge-Gehalt. Dabei wird die Substrattemperatur weiter reduziert und erreicht schließlich 450 °C - 500 °C.

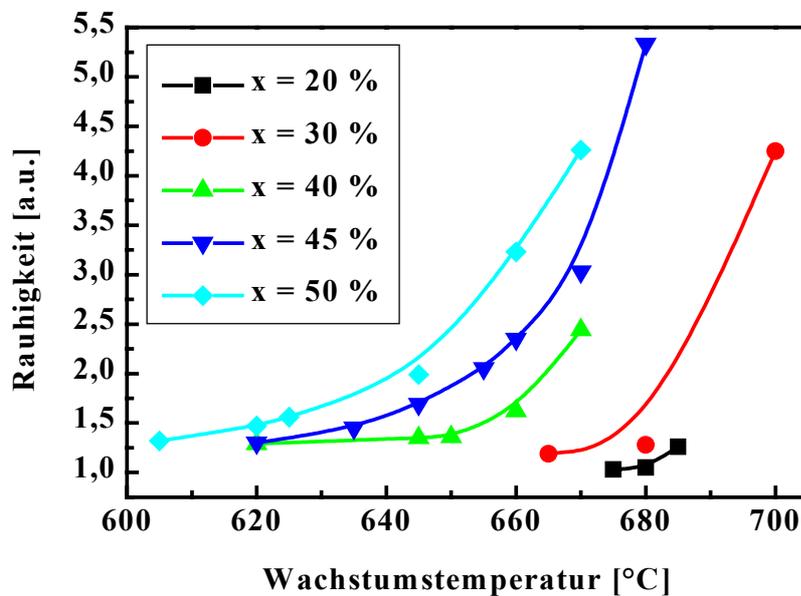


Abbildung 4.2: Abhängigkeit der Oberflächenrauhigkeit des Puffers von der mittleren Wachstumstemperatur und vom Ge-Gehalt

Mit Hilfe von AFM (Atomic Force Microscopy) und optischer Mikroskopie kann nachgewiesen werden, daß eine feine, regelmäßige und verzerrungsfreie Anordnung der Kreuzschraffierung an der Halbleiteroberfläche immer auch ein Indikator für eine gute Kristallqualität ist. Die Analyse der Oberflächenmorphologie verschiedener SiGe-Puffer hat ergeben, daß eine Abhängigkeit von der Steigerungsrate des Ge-Gehaltes besteht. Im Bereich von 10 % - 30 %/ μm bei entsprechender Adaption der Wachstumstemperatur zeigt die Pufferoberfläche ein feines, regelmäßig angeordnetes Cross-hatch Muster. Über 30 %/ μm erhöht sich die Rauigkeit der Oberfläche, und ein zunehmend dreidimensionales Wachstum wird sichtbar. XTEM-Analysen an gradierten Pufferschichten haben ergeben, daß die Dichte der aufsteigenden Versetzungslinien ebenfalls von der Steilheit der Ge-Rampe abhängt und bei 10 %/ μm auf $2 - 5 \cdot 10^5 \text{ cm}^{-2}$ absinkt [4.5]. Ein weiterer Parameter, der Einfluß auf die Oberflächenmorphologie hat, ist die Substrattemperatur beim Wachstum des gradierten Puffers. Abbildung 4.2 zeigt die Oberflächenrauhigkeit von Puffern unterschiedlichen Ge-Gehalts, die alle mit einer Steigerung des Ge-Gehaltes von 20 %/ μm realisiert wurden, in Abhängig-

keit von der mittleren Wachstumstemperatur. Um eine minimale Oberflächenrauigkeit der Schichtstruktur zu erzielen, muß die Epitaxietemperatur mit steigendem Ge-Endgehalt reduziert werden. Das abrupte Ansteigen der Rauigkeit beim 30 % und 50 % Puffer deutet auf das Einsetzen des dreidimensionalen Wachstums hin und beschreibt dadurch die maximale Epitaxietemperatur.

Den strukturellen Aufbau und die auftretenden Kristalldefekte eines linear gradierten Puffers mit einem Ge-Gehalt von 45 % zeigt die querschnittliche TEM-Aufnahme in Abbildung 4.3. Die zur Relaxation führenden Gitterversetzungen befinden sich ausschließlich im gradierten Teil des Puffers und beeinflussen nicht die Kristallqualität des Schichtabschnittes mit konstantem Ge-Gehalt. Die Oberfläche des Puffers weist eine glatte Morphologie auf, ohne aufsteigende Versetzungslinien oder andere Defekte. Der Relaxationsgrad eines gradierten Puffers zeigt zwar eine gewisse Abhängigkeit von der Ge-Konzentration, liegt aber bei guter Qualität in jedem Fall im Bereich über 90 %.

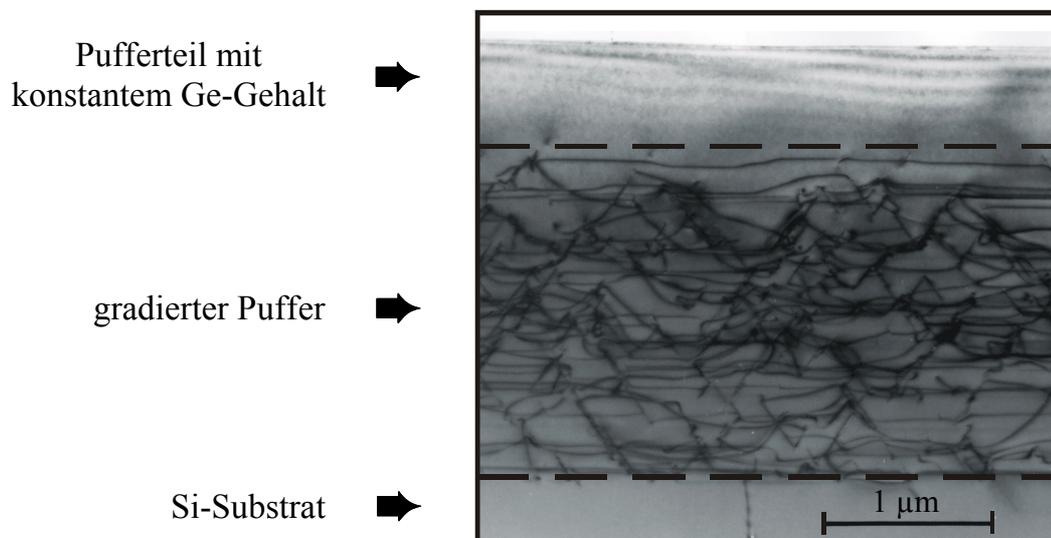


Abbildung 4.3: XTEM-Aufnahme eines linear gradierten Puffers mit $x = 45\%$

4.1.2 Alternative Pufferkonzepte

Der große Nachteil des gradierten Puffers besteht in der langen Wachstumsdauer. Bei einer typischen Dicke des Gesamtpuffers von 2,1 µm und einer mittleren Epitaxiedauer von 3 - 4 Stunden im MBE-System ist das Konzept aus betriebs-

wirtschaftlichen Aspekten nur schwer in einen Produktionsablauf mit hohem Durchsatz zu integrieren. Aus diesem Grund werden große Anstrengungen unternommen, alternative Konzepte zu entwickeln, die qualitativ die Güte eines gradierten Puffers ermöglichen und gleichzeitig eine geringe Pufferdicke aufweisen. Die Basis vieler dieser Strategien ist die bewußte Lokalisierung von Gitterdefekten in einer dünnen, definierten Zone im Bereich der Si/SiGe-Grenzschicht. In dieser Zone werden die auftretenden Gitterspannungen in Form von Kristallversetzungen aufgenommen und die entstehenden Threading Dislocations in ihrer weiteren Ausbreitung eingeschränkt. Die Realisierung dieser Defektschicht kann z. B. durch das Wachstum einer pseudomorphen SiGe-Schicht bei tiefen Epitaxietemperaturen erfolgen [4.6] oder durch nachträglichen Beschuß mit Ionen. Das eigentliche virtuelle Substrat wird anschließend nahezu versetzungsfrei auf die Defektschicht aufgewachsen, so daß sich eine gesamte Dicke im Bereich von 300 - 500 nm ergibt. Ein weiterer Ansatz erzeugt die Defektzone im Si-Substrat unterhalb der pseudomorphen SiGe-Schicht mit Hilfe einer hochdosierten Wasserstoffimplantation [4.7]. Der nachfolgende Tempersschritt bei 800 °C führt zu einer nahezu kompletten Relaxation des Puffers [4.8]. Das Abführen der Gitterversetzungen nach unten in Richtung Substrat ist auch das Prinzip des SOI-Puffers [4.9] Dabei wird das virtuelle SiGe-Substrat auf einer dünnen Si-Schicht aufgewachsen, die ihrerseits auf Oxid gebettet ist. Die Si-Schicht hat dabei eine relaxierende Wirkung, weil sie sich plastisch an die Gitterkonstante der SiGe-Schicht anpaßt. Darüber hinaus hat der Einsatz von Surfactant-Materialien, insbesondere von Antimon, gute Ergebnisse bei der Herstellung von SiGe- und sogar reinen Ge-Puffern erzielt [4.10], [4.11]. Sie zeichnen sich durch gute Kristallqualität und einen hohen Relaxationsgrad aus.

Abschließend soll darauf hingewiesen werden, daß sich alle beschriebenen alternativen Pufferkonzepte noch im Entwicklungsstadium befinden und noch nicht die technologische Reife und Qualität des gradierten Puffers erreicht haben.

4.2 Dimensionierung und Auslegung der aktiven Schichten

Das Konfigurieren und Abstimmen der aktiven Schichten auf den gewünschten Anwendungsbereich des Bauelementes ist ein Optimierungsproblem mit mehreren Variablen, die sich teilweise gegenseitig beeinflussen. Die Eigenschaften des Transistors können von der Epitaxieseite über den Ge-Gehalt, die Kanaldicke,

die Anordnung und Dotierstoffkonzentration der Dotierschichten, die Spacerdicke und die Auslegung der Cap-Schichten manipuliert werden. Neben einem gut abgestimmten Gleichstromverhalten sollen sich die hier betrachteten SiGe-Hetero-FETs vor allem durch eine leistungsfähige Hochfrequenzcharakteristik für Anwendungen über 10 GHz auszeichnen. Aus diesem Grund sind die Grenzfrequenzen f_T und f_{max} wichtige Indikatoren für die Tauglichkeit der Transistoren für diesen Einsatzbereich. Die Grenzfrequenzen werden nach Gleichung 3.58 und 3.59 nachhaltig durch die Steilheit g_m , die Kapazitäten C_{gs} und C_{gd} und die Widerstände R_s und R_d bestimmt. Die Abhängigkeiten dieser Größen von den einstellbaren Schichtparametern sind in Abbildung 4.4 schematisch wiedergegeben. Dabei stehen die zu optimierenden Zielgrößen im Mittelpunkt der Darstellung und sind über die Zwischengrößen μ_o , n_s und C_0 mit den Schichtparametern verbunden.

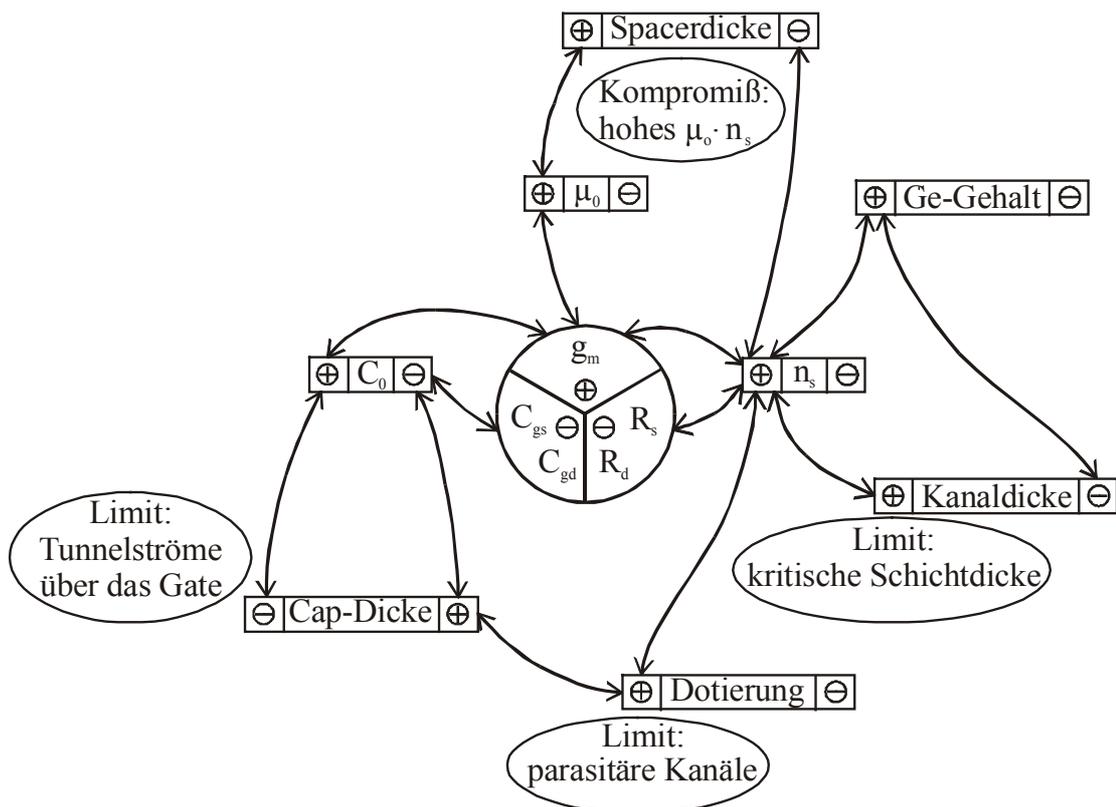


Abbildung 4.4: Schematische Darstellung der Abhängigkeiten zwischen den physikalischen Ersatzschaltbildelementen g_m , C_{gs} , C_{gd} , R_s und R_d und den Schichtparametern Ge-Gehalt, Kanaldicke, Dotierkonzentration, Spacerdicke und Cap-Schichtdicke

4.2.1 Einfluß des Ge-Gehaltes auf die Schichtparameter

Eine guter Ansatzpunkt für die Konfiguration der aktiven Schichten ist die Wahl des Ge-Gehaltes x . Dieser wird schon beim Wachstum des Puffers berücksichtigt und spiegelt sich im Ge-Endwert des Pufferteils mit konstanter Komposition wider. Während in der Anfangszeit des modulationsdotierten SiGe-HFETs Ge-Gehalte von 25 % - 30 % anvisiert wurden [4.12], [4.13], [4.14], erfolgte im Zuge dieser Schichtoptimierung eine permanente Steigerung von x , so daß heute 40 % - 50 % Germanium als Standard gilt. Damit zeigt sich ein deutlicher Trend zu Schichtkonfigurationen, die auf große Ladungsträgerdichten im Kanal ausgelegt sind und dadurch hohe Steilheiten und niedrige Schichtwiderstände anstreben. Die Abhängigkeit der Beweglichkeit μ_0 und der Ladungsträgerdichte n_s vom Ge-Gehalt zeigt Abbildung 4.5. Anhand realer modulationsdotierter Schichten mit gleicher Konfiguration, aber unterschiedlichem x -Anteil wird hier dargestellt, daß die zunehmende Banddiskontinuität im Leitungsband ein besseres "Confinement" der Ladungsträger im Kanal verursacht. Durch die Zunahme der Streuwahrscheinlichkeit zwischen den Ladungsträgern im 2DEG verringert sich jedoch die Niederfeldbeweglichkeit. Das für die Steilheit und den Strom im linearen Bereich relevante Produkt aus $\mu_0 \cdot n_s$ steigt allerdings für höhere Ge-Gehalte weiter an und rechtfertigt damit diesen Entwicklungstrend (Abbildung 4.6).

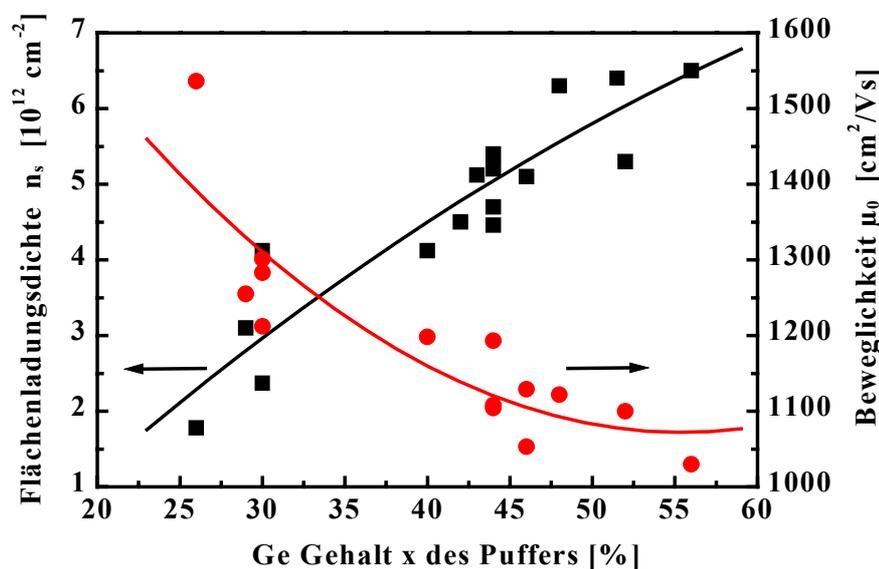


Abbildung 4.5: Abhängigkeit der Ladungsträgerdichte n_s und der Beweglichkeit μ_0 von dem Ge-Gehalt x des Puffers, bei sonst identischer Schichtkonfiguration

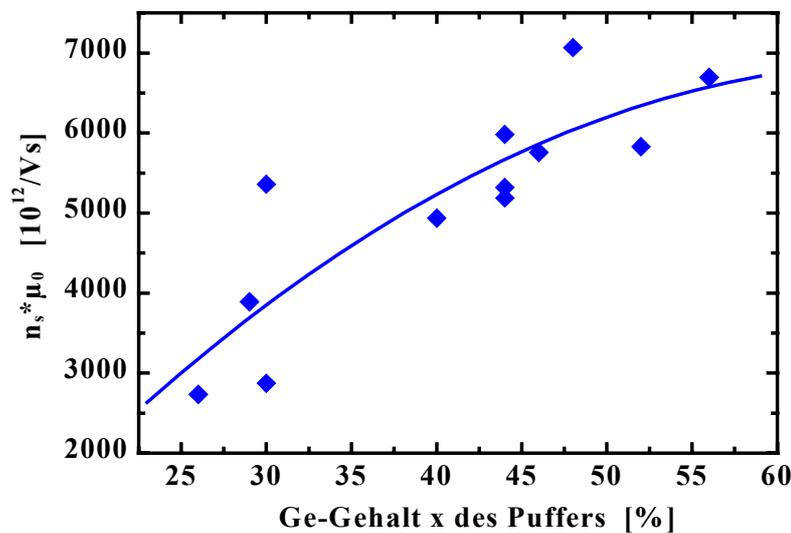


Abbildung 4.6: Entwicklung des Produktes $n_s \cdot \mu_0$ mit zunehmendem Ge-Anteil

4.2.2 Die kritische Schichtdicke des Kanals

Der limitierende Faktor für den zunehmenden Ge-Gehalt ist die abnehmende kritische Schichtdicke des Kanals. Dabei ist eine unterkritische Kanaldicke unbedingt einzuhalten, da der Übergang vom pseudomorphen Kristallzustand in den teilrelaxierten Bereich von der Ausbildung von Gitterversetzungen begleitet wird, die einen drastischen Abfall der Beweglichkeit mit sich bringen. In Abbildung 4.7 ist die Abhängigkeit der kritischen Schichtdicke d_{krit} vom Ge-Gehalt kompressiv verspannter SiGe-Schichten auf Si-Substrat dargestellt.

Im stabilen Epitaxiebereich wird die Gitterfehlanspannung vollständig in elastische Kristallverspannungen umgewandelt, und es erfolgt pseudomorphes Schichtwachstum. Die Stabilitätsgrenze ergibt sich aus der mechanischen Gleichgewichtstheorie nach Matthews et al. [4.15] und bildet die Trennungslinie zwischen dem stabilen und dem metastabilen Schichtbereich. Oberhalb dieser Grenze schließt sich eine Ungleichgewichtsphase an, in der die Bildung von Gitterversetzungen energetisch günstiger ist als der weitere Aufbau von Verspannungen. Die Generation und Ausbreitung dieser Versetzungen ist in Verbindung mit der Wachstumstemperatur stark von kinetischen Limitierungen und Energiebarrieren für die Nukleation von Defekten abhängig. Dadurch bietet sich die Möglichkeit, mit Niedertemperaturepitaxieverfahren den Bereich pseudo-

morphen Schichtwachstums über den Gleichgewichtsgrenzwert hinaus zu erweitern. Experimentelle Daten liefern die Wachstumsversuche von Bean et al. [4.16], die bei 550 °C mit der Molekularstrahlepitaxie pseudomorphe SiGe-Schichten auf Si-Substrat, weit oberhalb der Stabilitätsgrenze erzeugt haben. Im Relaxationsbereich wird der Abbau der Kristallverspannungen dominiert von der Generation und Verzweigung von Gitterversetzungen. Dieser Vorgang kann durch die Angabe des Relaxationsgrades oder die verbleibende Verspannung in der Wachstumsebene charakterisiert werden.

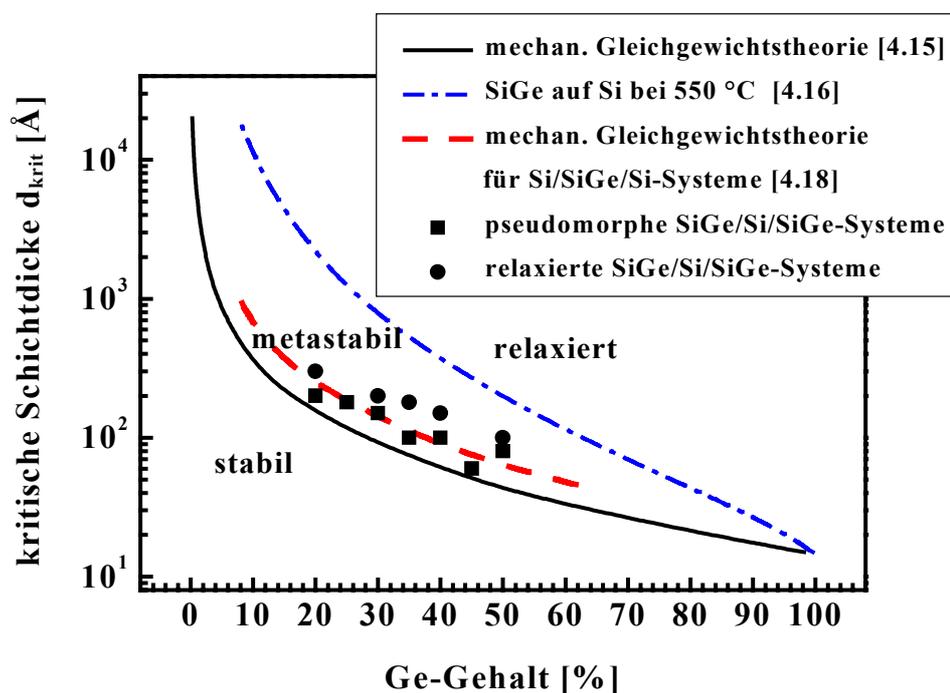


Abbildung 4.7: Verlauf der kritischen Schichtdicke d_{krit} für epitaktische $\text{Si}_{1-x}\text{Ge}_x$ -Schichten auf Si-Substrat und symmetrisch verspannte Si-Kanäle in Abhängigkeit des Ge-Gehaltes x

Die Übertragbarkeit der dargestellten Ergebnisse auf SiGe-Heterostrukturen mit eingebetteten, zugverspannten Si-Kanälen untersuchte Glück [4.17]. Er bestätigte dabei eine von Jain et al. [4.18] aufgestellte Gleichgewichtstheorie für vergrabene SiGe-Kanäle im relaxierten Silizium. Die Analyse der im Rahmen dieser Arbeit verwendeten modulationsdotierten SiGe-Heterostrukturen mit pseudomorphem Si-Kanal untermauern die experimentellen Resultate von Glück. Die untersuchten Schichtstrukturen ohne ein Anzeichen für Relaxation sind dabei als Quadrate in Abbildung 4.7 dargestellt. SiGe-Heterosysteme, bei denen Untersuchun-

gen mit der Hall-Beweglichkeitsmethode deutliche Merkmale für eine Relaxation zeigten, treten als gefüllte Kreise auf. Der Übergang vom stabilen in den teilrelaxierten Bereich liegt, im Gegensatz zu den experimentellen Werten von Bean et al. für verspannte SiGe-Schichten, bei pseudomorphen Si-Kanälen nur wenig oberhalb der Gleichgewichtstheorie. Die Gründe dafür liegen in der höheren Defektdichte des epitaktischen SiGe-Puffers im Gegensatz zum gezogenen Si-Kristall des Substrates. Diese Defekte fördern die Ausbreitung von Gitterversetzungen und damit eine früher einsetzende Relaxation. Für das Wachstum von pseudomorphen Si-Kanälen sollte deshalb unbedingt die Limitierung durch die Gleichgewichtstheorie eingehalten werden.

4.2.3 Konfiguration der Dotierschichten

Nach der Bestimmung des Ge-Gehaltes und der daraus resultierenden maximalen Kanaldicke ist die Anordnung und Auslegung der Dotierschichten von entscheidender Bedeutung für die Charakteristik des Transistors. Die Dotierschichten versorgen den Kanal mit Ladungsträgern, die dann das 2DEG bilden. Als Grundregel für die Optimierung sollte im Arbeitspunkt des Bauelementes eine maximale Ladungsträgerdichte im Kanal erzeugt werden, bei vollständiger Lokalisierung der freien Elektronen aus den Dotierschichten im Potentialtopf. Dieses Ziel erfordert eine optimierte Abstimmung der Dotierschichtkonfiguration mit der Auslegung der Spacer und Cap-Schichten und ist deshalb nur in Kombination lösbar.

Prinzipiell sollte zuerst geklärt werden, welche Dotierschichtanordnung den Bauelementanforderungen am besten genügt. Neben der klassischen, aus der III/V-Technologie bekannten, vorderseitig modulationsdotierten Struktur bieten sich außerdem die Möglichkeiten einer doppelseitigen Dotierschichtanordnung und einer Variation mit Kanaldotierung an (Abbildung 4.8).

Abbildung 4.9 vergleicht die konkreten Schichtanordnungen der drei Dotierungskonzepte aus Abbildung 4.8 mit Hilfe der Simulation hinsichtlich der Ladungsträgerdichten im 2DEG und in den parasitären Kanälen.

Die einseitige Dotierschichtanordnung auf der Vorderseite bietet, bei gegebener Potentialtopftiefe, einen begrenzten Spielraum, die Ladungsträgerdichte im Kanal zu erhöhen. Mit steigender Dotierstoffkonzentration werden nicht mehr alle freien Ladungsträger in den Kanal transferiert, und es kommt schon im ange-

strebten Arbeitspunkt zur Ausbildung eines parasitären Parallelkanals mit 3D-Leitfähigkeit. Dieser schirmt die Ladungsträger im Potentialtopf ab und reduziert dadurch die Steuerwirkung auf das 2DEG und damit die Steilheit. Zwar läßt sich die Steilheit durch eine Reduzierung der Cap-Schichtdicke erhöhen, dies steigert jedoch auch die Wahrscheinlichkeit von Tunnelströmen über das Schottky-Gate.

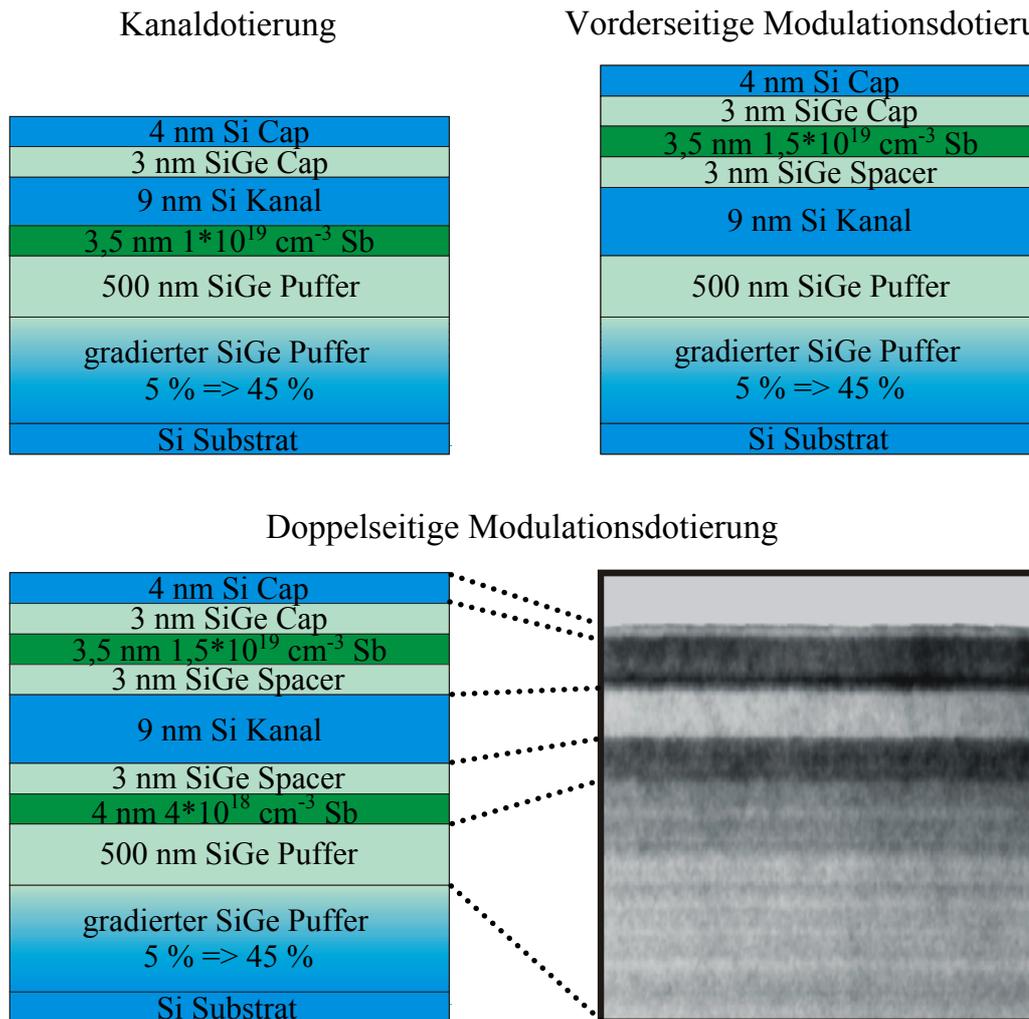


Abbildung 4.8: Schematische Darstellung der Variationsmöglichkeiten zur Anordnung der Dotierschichten beim Si/SiGe-n-Kanal-HFET

Durch die Einführung einer zweiten Dotierschicht, die ebenfalls getrennt durch einen Spacer auf der Rückseite des Kanal angeordnet ist, lassen sich gesteigerte Ladungsträgerdichten im Kanal erzielen. Die Abstimmung und das Verhältnis der Dotierstoffkonzentrationen von Vorder- und Rückseitendotierschicht haben

dabei entscheidenden Einfluß auf die Ladungsträgersteuerung und -dichte im Kanal. Eine zu hohe Dotierung auf der Vorderseite führt zum Einsetzen des parasitären MESFETs und zu einer verminderten Ladungsträgerkontrolle, eine überdosierte Rückseitendotierung erzeugt einen nicht abschnürbaren Parallelkanal, der Leckströme nach sich zieht. Durch die Rückseitendotierung verschiebt sich ebenfalls der Bandverlauf des Kanals. Das Absenken der hinteren Bandkante läßt aus der Dreiecksform einen rechteckigen Potentialtopf entstehen und steigert dadurch die Besetzungswahrscheinlichkeit der Energiezustände. Dies zieht eine zusätzliche Erhöhung der Ladungsträgerdichte im Kanal nach sich.

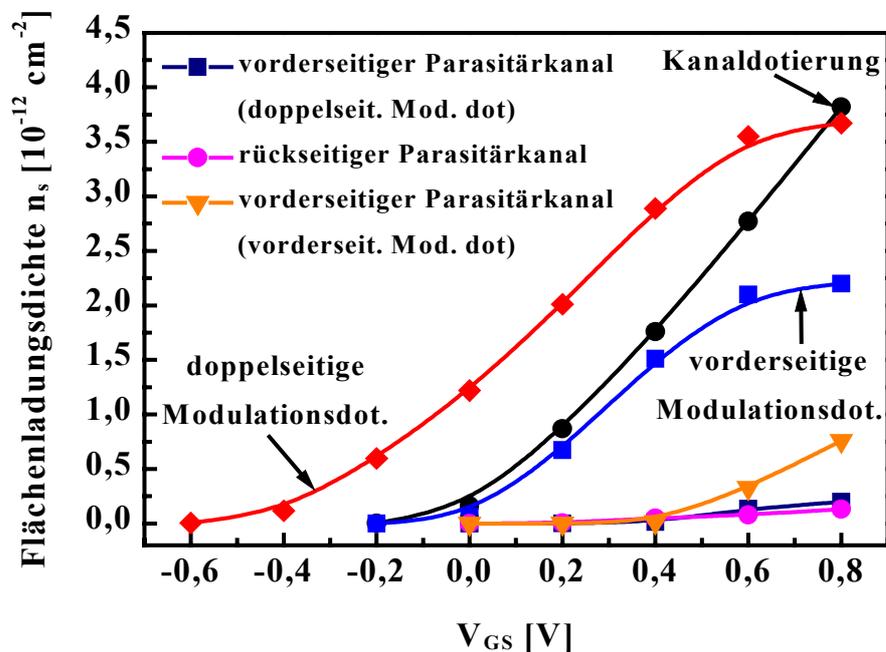


Abbildung 4.9: Simulation der erzielbaren Ladungsträgerdichten im 2DEG von Si/Si_{0,55}Ge_{0,45}-HFETs für die drei Dotierschichtanordnungen und -konfigurationen aus Abbildung 4.8. Struktur 1: Kanaldotierung (schwarz); Struktur 2: vorderseitige Modulationsdotierung (blau); Struktur 3: doppelseitige Modulationsdotierung (rot)

Eine Abweichung vom Prinzip des modulationsdotierten HFETs stellt das Kanaldotierungskonzept dar. Es ermöglicht die Realisierung sehr hoher Ladungsträgerdichten, jedoch auf Kosten der Beweglichkeit. Abbildung 4.9 zeigt im Vergleich zur Vorder- und Doppelseitendotierung keine parasitären Kanäle und eine lineare Steuercharakteristik der Ladungsträger. Dieser Ansatz bietet sich insbesondere für Leistungsbaulemente an, die hohe Ströme und Verstärkungen an-

streben. Für rauscharme Hochfrequenzbauelemente sind hingegen die Konzepte der Vorder- und Doppelseitendotierung vorzuziehen.

Die Dimensionierung der Dotierschichten wird maßgeblich von den MBE-spezifischen Dotierverfahren LTE und DSI bestimmt. Beide Verfahren nutzen eine Antimonvorbelegung auf der Kristalloberfläche als Basis für die Dotierung. Die angestrebte Dotierstoffkonzentration in den Schichten ist sowohl von der Höhe der Vorbelegung ($2,0 - 2,4 \cdot 10^{14} \text{ cm}^{-3}$) als auch vom Antimonfluß während des Wachstums ($1 - 8 \cdot 10^{14} \text{ cm}^{-3}$) abhängig. Beim Low Temperature Epitaxy Verfahren wird ausgenutzt, daß der Antimoneinbau während des Wachstums temperaturabhängig ist. Unterhalb von $300 \text{ }^\circ\text{C}$ können Konzentrationen von über $2 \cdot 10^{19} \text{ cm}^{-3}$ erzielt werden, dabei nimmt jedoch die Kristallqualität stark ab, und es bilden sich viele Gitterversetzungen. Oberhalb von $350 \text{ }^\circ\text{C}$ fällt die Antimonkonzentration rapide unter $1 \cdot 10^{16} \text{ cm}^{-3}$, so daß Wachstumstemperaturen zwischen $350 \text{ }^\circ\text{C}$ und $300 \text{ }^\circ\text{C}$ angestrebt werden.

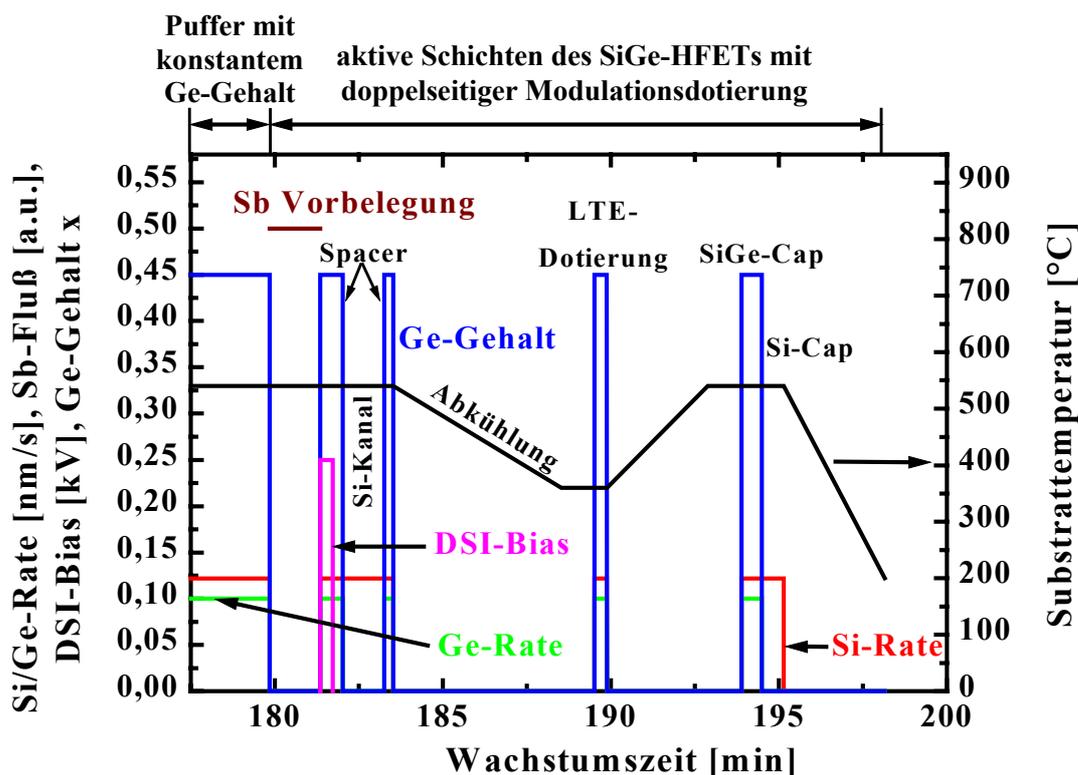


Abbildung 4.10: Verlauf des Ge-Gehaltes, der Substrattemperatur, der Si- und Ge-Rate und der Antimonvorbelegung beim Wachstum einer aktiven Schichtstruktur mit doppelseitiger Dotierung [4.19]

Das DSI-Verfahren (Doping by Secondary Ion Implantation) nutzt den Anteil von etwa 2 % Siliziumionen, der neben den Siliziumatomen in der Effusionszelle der MBE-Anlage entsteht. Durch das Anlegen einer negativen Spannung an das Substrat werden diese Ionen beschleunigt und bauen beim Auftreffen auf die Kristalloberfläche die vorbelegten Antimonionen in die Epitaxieschicht mit ein. Die Anzahl der implantierten Antimonionen ist von der Substratspannung abhängig und erreicht ihren Maximalwert von $2 \cdot 10^{19} \text{ cm}^{-3}$ bei etwa -350 V. Das DSI-Verfahren wird standardmäßig für die Rückseitendotierung eingesetzt, weil es in Hinblick auf die Epitaxie des nachfolgenden Kanals die bessere Kristallqualität liefert. In Abbildung 4.10 ist der zeitliche Verlauf der Si- und Ge-Rate, des Ge-Gehaltes, der Antimonvorbelegung und der Substrattemperatur beim Wachstum einer aktiven Schichtstruktur mit rückseitiger DSI-Dotierung und vorderseitiger LTE-Dotierung dargestellt.

Beide Verfahren benötigen eine endliche Wachstumszeit, um die gewünschte Dotierstoffkonzentration in die Epitaxieschichten einzubauen. Dadurch ergeben sich Dotierschichtdicken im Bereich von 2,5 - 4 nm. Im Vergleich zur flächenhaften δ -Dotierung bewirkt die integrale Dotierungsmethode dieselbe absolute Bandverbiegung, hat aber einen anderen Bandverlauf zur Folge. Dadurch wird bei extrem dünnen Cap-Schichten die Wahrscheinlichkeit von Tunnelströmen über das Gate erhöht.

Alle im Rahmen dieser Arbeit bearbeiteten Schichtkonfigurationen nutzen das Konzept einer doppelseitigen Dotierschicht. Gegenüber der klassischen Vorderseitendotierung liefert es eine höhere Ladungsträgerdichte im Potentialtopf, bei gleichzeitig effektiverer Modulation der Ladungsträger und einem geringeren Anteil an parasitären Kanälen. Um einen parasitären Rückseitenkanal zu verhindern, beträgt die Dosis der Dotierung dort nicht mehr als $2 \cdot 10^{12} \text{ cm}^{-2}$. Das Verhältnis zwischen Vorder- und Rückseitendotierung liegt dabei je nach Spacerdicke und Abstand zwischen Gate und Kanal bei 3/1 bis 5/1.

4.2.4 Dimensionierung der Spacerdicke

Die Effektivität der Dotierung läßt sich nur in Zusammenhang mit der gewählten Spacerdicke ermitteln. Die undotierten SiGe-Schichten zwischen dem Si-Kanal und den Dotierschichten sorgen für eine räumliche Trennung der Ladungsträger im 2DEG von den ionisierten Dotieratomrümpfen. Da die beweglichkeitshemmende Coulomb-Wechselwirkung mit $1/r^2$ abnimmt, führt eine zunehmende

Spacerdicke zu einem Beweglichkeitsanstieg, bis es zu einem Sättigungsprozess kommt. Die Ladungsträgerdichte im Potentialtopf reduziert sich jedoch, weil aus energetischen Gründen nicht mehr alle freien Elektronen in den Kanal transferiert werden. Experimentelle Analysen an modulationsdotierten Schichtstrukturen unterschiedlicher Spacerdicke ergeben den in Abbildung 4.11 dargestellten Verlauf von Beweglichkeit und Ladungsträgerdichte im Kanal. Die untersuchten Proben weisen alle eine doppelte Anordnung der Dotierschichten auf, mit symmetrischer Spacerdicke auf der Vorder- und Rückseite. Die Dotierkonzentration ist nominell bei allen Strukturen gleich, und der Ge-Gehalt liegt zwischen 46 % und 51 %.

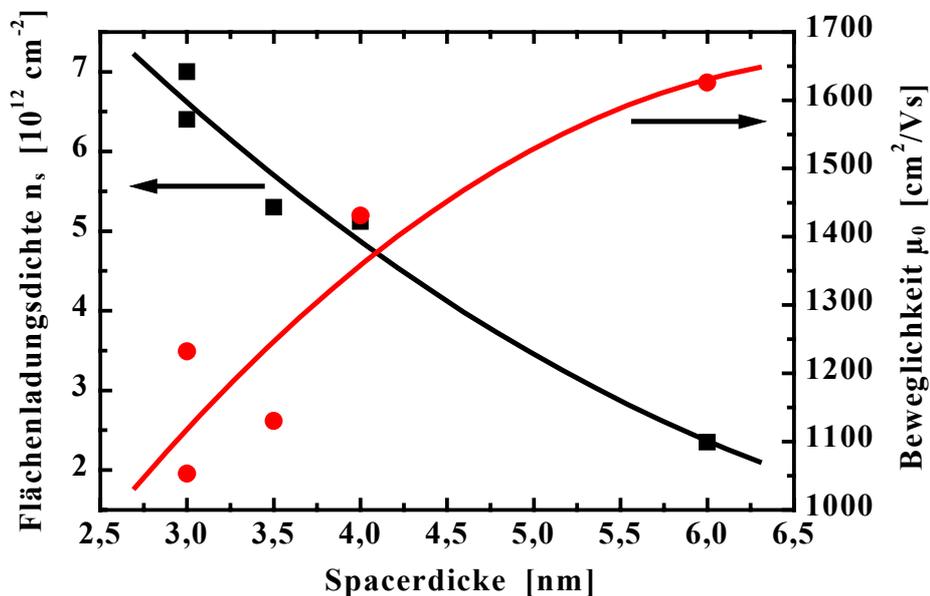


Abbildung 4.11: Abhängigkeit der Ladungsträgerdichte n_s und der Beweglichkeit μ_0 des 2DEGs von der Spacerdicke bei doppelseitig modulationsdotierten Schichtstrukturen

Das für reale Bauelementanwendungen wichtige Produkt aus Ladungsträgerdichte und Beweglichkeit findet für Spacerdicken zwischen 3,5 und 4,5 nm sein Optimum. Durch den Einsatz von Spacern im Bereich um 3 nm kann mehr Gewicht auf eine hohe Ladungsträgerdichte gelegt werden, was gleichzeitig zu einer Verringerung der parasitären Widerstände R_s und R_d führt.

4.2.5 Funktion und Auslegung der Cap-Schichten

Oberhalb der Vorderseitendotierschicht schließt das modulationsdotierte Schichtsystem mit einer undotierten SiGe-Schicht und einem dünnen Si-Cap ab. Das Si-Cap stellt das Interface zum Metallgate dar und bildet mit dem Platin eine Schottky-Diode. Die entstehende Schottky-Barriere von etwa 0,84 eV ermöglicht eine weite Aussteuerung der Gateelektrode um den Arbeitspunkt und hält die Leckströme über das Gate gering. Ein abschließendes SiGe-Cap würde zu einer niedrigeren Barriere führen und eine sehr reaktive Oberfläche mit einer hohen Anzahl an undefinierten Zuständen hinterlassen. Da das Si-Cap wie der Si-Kanal zugverspannt ist, muß seine Dicke unterhalb der kritischen Schichtdicke liegen. Reale Cap-Dicken bewegen sich im Bereich von 3 bis 5 nm, dabei muß berücksichtigt werden, daß durch Reinigungssequenzen und Oxidationsschritte während der Prozessierung ein Abtrag von etwa 1,5 nm erfolgen kann. Da der Abstand zwischen Gateelektrode und dem Schwerpunkt der Ladungsträger im 2DEG über C_0 effektiv die Steuerung des Drainstromes und damit die Steilheit beeinflusst, sollte er so gering wie möglich gehalten werden. Dies läßt sich bei einer bereits aufeinander abgestimmten Dotierschicht-Spacer-Kombination über die Dicke der undotierten SiGe-Schicht vornehmen. Limitiert wird diese Reduzierung durch zunehmende Tunnelströme über das Schottky-Gate. Sie tragen dann zum Drainstrom bei und verringern die Steuerwirkung der Gateelektrode. Weiterhin können sich bei zunehmend positiver Gatespannung Ladungsträger aus der Dotierschicht in einem parasitären Kanal im Si-Cap sammeln und dort einen unkontrollierbaren Stromfluß verursachen. Deshalb haben sich in Abhängigkeit von der Vorderseitendotierung Schichtdicken von 5 - 10 nm bewährt.

Hochdotierte Cap-Schichten zur Kontaktierung des stromführenden Kanals haben sich bei den SiGe-HFETs im Gegensatz zur III/V-Technologie nicht durchgesetzt. Die zur Verbesserung der Kontaktwiderstände nötigen Legierkontakte sind bei den SiGe-Transistoren nicht konkurrenzfähig zu einer Kombination aus Implantation und Silizidierung. Durch dieses Vorgehen kann außerdem die Kompatibilität zur Standard-CMOS-Technologie gewahrt bleiben.

Kapitel 5

Integrationskonzepte des SiGe-Hetero- Feldeffekttransistors

Das Integrationskonzept ist ein wichtiger Bestandteil im Realisierungsprozeß anwendungsoptimierter Halbleiterbauelemente. Es bestimmt neben dem strukturellen Aufbau des Transistors auch die technologische Umsetzung und den Herstellungsprozeß und sollte sich deshalb durch folgende Kriterien auszeichnen:

- optimale Ausnutzung des Leistungspotentials des Bauelementes
- hohe Packungsdichte
- gute Reproduzierbarkeit
- hohe Ausbeute
- Kompatibilität zu Standardtechnologien
- niedrige Kosten

Das Ziel des vorangegangenen Kapitels war die Dimensionierung und Konfiguration des Schichtsystems und damit die Optimierung der vertikalen Bauelementstruktur in Hinblick auf ein hohes Leistungspotential des Transistors. Im folgenden soll nun die konzeptionelle Realisierung des Bauelementes und die laterale Struktur eingehender betrachtet werden. Dies bietet die Möglichkeit, Einfluß auf die Packungsdichte und auf parasitäre Elemente zu nehmen, die eine negative Wirkung auf die Bauelementcharakteristik haben. Wie bei der Auslegung des Schichtsystems stehen dabei hauptsächlich parasitäre Widerstände, Induktivitäten und Kapazitäten im Mittelpunkt, die sich reduzierend auf die Grenzfrequenzen auswirken. Eine wichtige Rolle spielen außerdem layoutspezifische geometrische Abmessungen, wie die Gatelänge L_G , die Gateweite W_G oder die Abstände zwischen den ohmschen Kontakten und der Gateelektrode.

In diesem Kapitel werden zwei unterschiedliche Integrationskonzepte zur Herstellung von SiGe-n-Kanal-Hetero-Feldeffekttransistoren vorgestellt, die beide am Forschungszentrum der DaimlerChrysler AG in Ulm entwickelt wurden und sich stark in ihrem Optimierungsgrad voneinander abgrenzen. Das konventionelle Integrationskonzept ist angelehnt an Realisierungsverfahren, die aus der III/V-HEMT-Technologie stammen und von dort in der Anfangsphase des SiGe-HFETs adaptiert wurden. Im Gegensatz dazu übernimmt das selbstjustierende Konzept Prozeßmethoden aus der CMOS-Technologie, um die laterale Struktur und die Integration des Transistors weiter zu optimieren.

5.1 Das konventionelle Integrationskonzept

5.1.1 Bauelementstruktur und Herstellungsablauf

Das konventionelle Integrationskonzept basiert auf einer mesaisolierten, passivierten Transistorstruktur mit implantierten Source/Drain-Gebieten und einem Metall-T-Gate, das mit Hilfe der e-Beam-Lithografie erzeugt wird. Den technologischen Prozeßablauf und den Querschnitt der Bauelementstruktur zeigt Abbildung 5.1.

Ausgehend von dem spezifischen SiGe-Hetero-Schichtsystem des HFETs beginnt der Herstellungsprozeß mit der Definition der Bauelementmesa. Sie legt die Abmessungen des aktiven Transistorgebietes fest und isoliert die einzelnen Bauelemente eines Chips voneinander, da diese durch die ganzflächige Epitaxie miteinander verbunden sind. Anders als bei der III/V-Technologie hat sich die Separation durch Ionen- bzw. Protonenbeschuß bei SiGe-Bauelementen nicht durchgesetzt. Die residualen Leckströme bewegen sich im μA -Bereich und sind damit inakzeptabel. Zur Isolation der Mesaflanken wird anschließend ganzflächig eine SiO_2 -Schicht als Feldoxid aufgebracht. Um den aktiven Bereich des Transistors wieder freizulegen, wird mit Hilfe einer lithografischen Maske und einem Ätzschritt ein Fenster auf der Mesa geöffnet. Die Kontaktierung des hochbeweglichen 2DEGs im Si-Kanal erfolgt über ohmsche Kontakte, die durch Ionenimplantation in den Kristall eingebracht werden. Die Definition der lateralen geometrischen Abmessungen der Source/Drain-Gebiete geschieht dabei über eine Fotolackmaske. Nach der Implantation muß der eingebrachte Dotierstoff auf regulären Gitterplätzen aktiviert werden und die Rekristallisation der Gitterschäden stattfinden. Dazu wird ein RTA-Prozeß bei 570 - 650 °C durchgeführt.

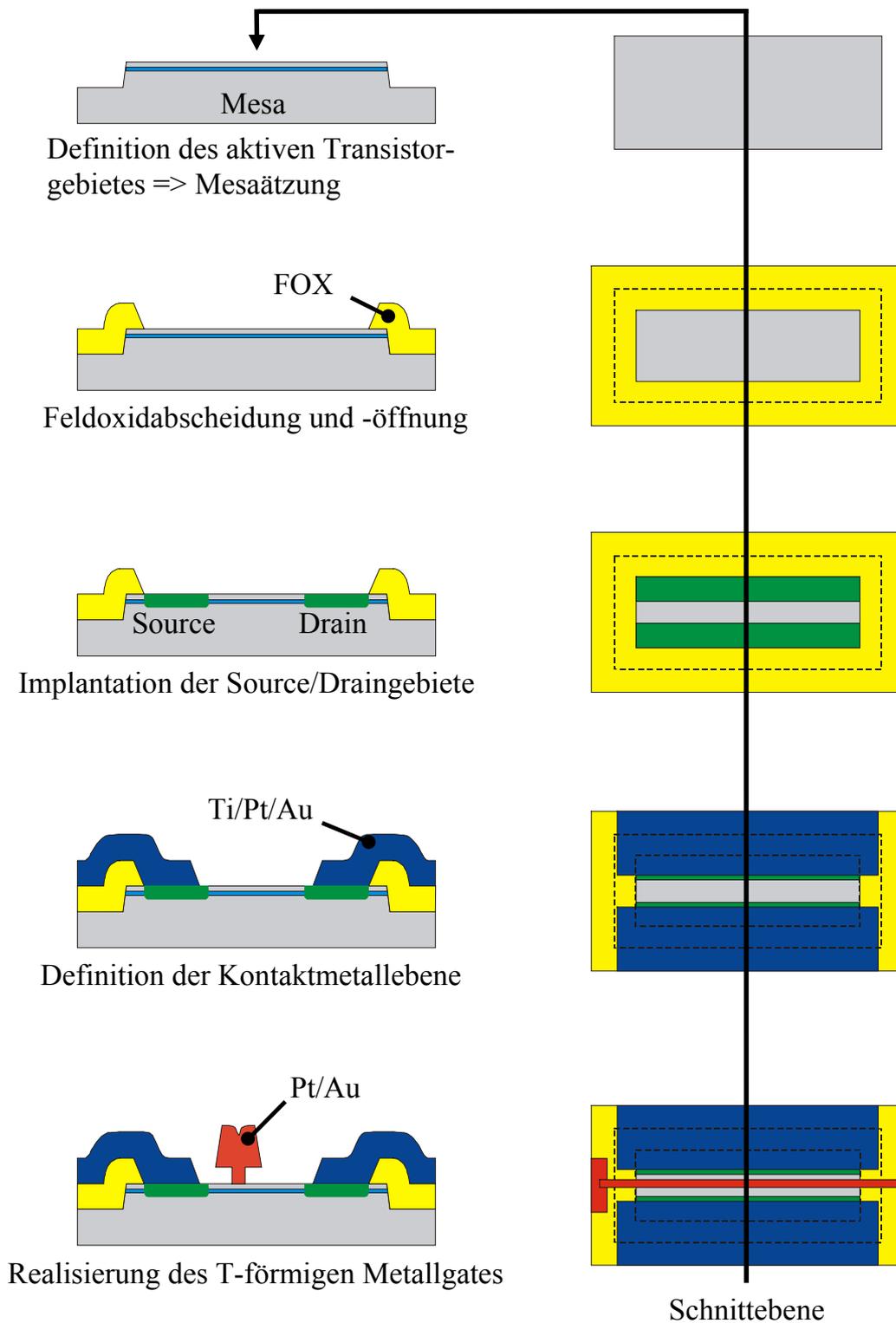


Abbildung 5.1: Prozeßablauf des konventionellen Integrationskonzeptes im Bauelementquerschnitt und als Maskenebenen

Zur Kontaktierung und Charakterisierung der Einzeltransistoren mit Hilfe eines Spitzenmeßplatzes bzw. zur Verschaltung verschiedener aktiver und passiver Bauelemente wird die ohmsche Kontaktmetallebene benötigt. Dabei kommt eine Metallkombination aus Ti/Pt/Au zum Einsatz, die direkt auf die ohmschen Kontakte aufgedampft wird. Die Strukturierung erfolgt mit einer Lithografiemaske im Lift-off Verfahren. Abschließend findet die Definition der Gateelektrode über eine e-Beam-Lithografie mit anschließendem Lift-off statt. Die Position des Gates wird dabei durch Justiermarken festgelegt, die mit der Metallisierungsebene aufgebracht werden. Der eingesetzte Lithografieprozeß nutzt ein Dreilagenslacksystem, um ein T-förmiges Gate zu realisieren. Der Vorteil einer Gatestruktur mit T-förmigem Querschnitt gegenüber dem standardmäßig eingesetzten rechteckigen Gate spiegelt sich insbesondere in den Hochfrequenzeigenschaften wider. Der gewünschte kleine Gatewiderstand R_g wird über die große Querschnittsfläche des breiten Gatekopfes realisiert, und die angestrebte kleine Gatelänge L_G im Bereich von 250 - 100 nm kommt über den schmalen Gatefuß zustande. Als Materialkombination für das Gate kommt Pt/Au zum Einsatz, was einerseits eine hohe Schottkybarriere zur Si-Oberfläche erzeugt und andererseits den Widerstand klein hält.

5.1.2 Modifikation und partielle Optimierung des Konzeptes

Das vorgestellte Integrationskonzept ermöglicht die einfache und schnelle Herstellung von SiGe-HFETs. Dazu werden vier Maskenebenen und ein e-Beam-Lithografieschritt benötigt. Für Schnelltests, die Aufschluß darüber geben sollen, ob die verwendete Konfiguration des Schichtsystems bauelementtauglich ist, kann die teure und zeitaufwendige e-Beam-Lithografie durch eine optische Gate-Realisierung ersetzt werden. Dadurch ist auch eine preisgünstige Potentialabschätzung für neue Bauelementvariationen einfach durchführbar. Um jedoch Anforderungen wie Reproduzierbarkeit, Ausbeute und Reduzierung der parasitären Elemente gerecht zu werden, bedarf es einiger Modifikationen dieses Konzeptes.

Ein großer Schwachpunkt dieser Transistorrealisierung ist die mangelnde laterale Optimierung und die daraus folgende geringe Reproduzierbarkeit. Die optische Justage verschiedener Lithografiemasken zueinander birgt immer viele Unsicherheiten bei der Einhaltung der nominellen Abmessungen und Abstände. Das kann zur Folge haben, daß aufgrund von Fehljustagen eine Verschiebung der Maskenebenen untereinander auftritt, die nachhaltig das Transistorverhalten beeinflusst. Insbesondere der Abstand zwischen den ohmschen Source/Drain-

Kontakten und die relative Position des Gatefusses zu den Implantationsgebieten sollten in Hinblick auf eine reproduzierbare Bauelementcharakteristik exakt einstellbar sein. Für die präzise Justage des Gates werden metallische Justiermarken benötigt. Diese müssen einen ausreichenden Kontrast zum darunterliegenden Siliziumdioxid liefern, damit sie automatisch von der Elektronenstrahlbelichtungsanlage erkannt werden. Bei gutem Kontrast und scharfen, geradlinigen Kanten der Justiermarken erfolgt die Belichtung des Lacksystems mit einer Justiergenauigkeit von ± 50 nm. Um den Prozeßablauf so einfach wie möglich zu halten, werden die quadratischen Justiermarken mit der ohmschen Kontaktmetallebene aufgedampft. Die etwa 250 nm dicke Metallkombination aus Ti/Pt/Au erfüllt dabei die oben genannten Voraussetzungen, um eine exakte Justage zu gewährleisten.

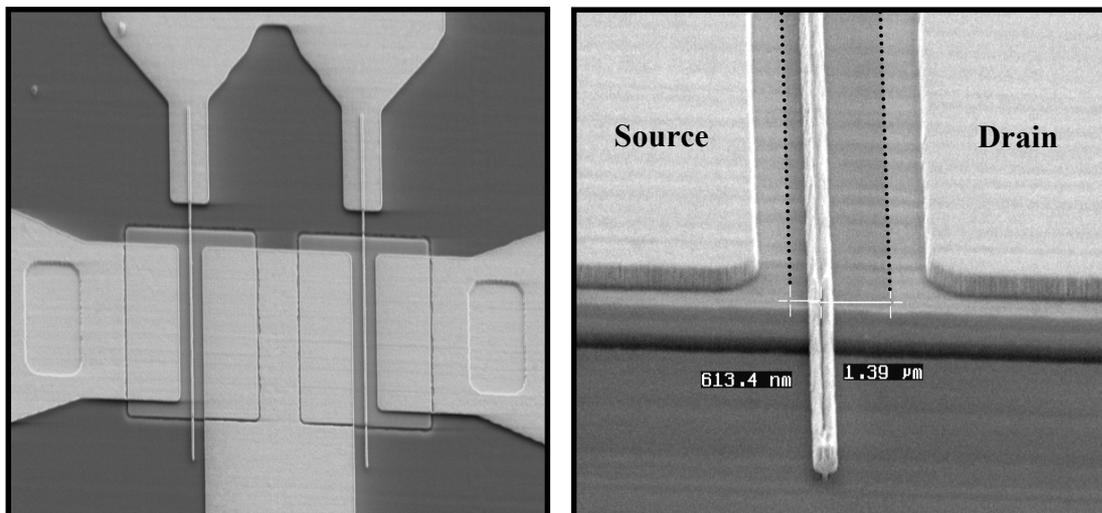


Abbildung 5.2: Foto eines SiGe-HFETs mit π -Gateanordnung für Hochfrequenzanwendungen mit der Teilansicht des Source/Drain-Bereiches

Wie aus Abbildung 5.2 deutlich zu erkennen ist, wird das Gate asymmetrisch zwischen Source und Drain angeordnet und dabei der Abstand zwischen Gatefuß und Implantationsgebiet auf der Sourceseite so klein wie möglich gehalten. Dieses reduziert den Sourcewiderstand R_s , der entsprechend Gleichung 3.53 entscheidenden Einfluß auf die extrinsische Steilheit $g_{m,extr}$ hat und darüber die Grenzfrequenzen f_T und f_{max} bestimmt. Durch eine gegenläufige Fehljustage zwischen der Implantationsebene und der ohmschen Kontaktmetallebene können die in Abbildung 5.3 dargestellten Abweichungen von der nominellen Gateausrichtung eintreten. Bei einer Verschiebung der Metallmaske in Richtung des inneren

Randes des Sourcegebietes wandert die Gateposition zum Draingebiet. Dadurch vergrößert sich der Abstand zwischen Gatefuß und Sourceimplantation und damit auch der Sourcewiderstand, was die Steilheit verringert. Im umgekehrten Fall rückt das Gate näher an das Sourcegebiet heran. Bei einem nominell angestrebten Abstand L_{GS} von 200 - 500 nm und einer Justiergenauigkeit im Kontaktbelichtungsverfahren von etwa 500 nm kommt es dabei mit hoher Wahrscheinlichkeit zu einem Kurzschluß zwischen Gate- und Source-Elektrode.

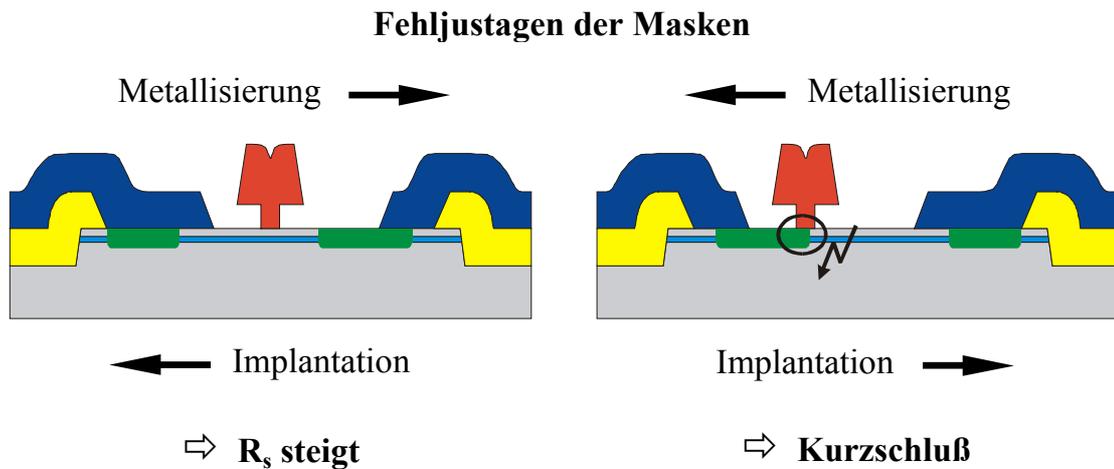


Abbildung 5.3: Einfluß einer gegenläufigen Fehljustage zwischen der Implantationsebene und der ohmschen Kontaktmetallebene auf die Position des Gates

Eine wirksame Maßnahme gegen diese Schwäche im Prozeßablauf ist die Realisierung der e-Beam-Justiermarken zusammen mit der Implantationsebene. Dadurch hängt die reale Abweichung der Gateposition nur noch von der Toleranz der lichteoptischen Implantationsmaske und der Justiergenauigkeit des Elektronenstrahlbelichters ab. Bei der Umsetzung dieser Modifikation wird die Lackmaske nach dem Implantationsschritt zunächst nicht entfernt. Zum Schutz der implantierten Source/Drain-Gebiete werden diese mit einer fotolithografischen Hilfsmaske abgedeckt, die Öffnungen für die Justiermarken aber freigelassen. Aufgrund der Karbonisierung der Lackoberfläche durch den Ionenbeschuß ist die Implantationsmaske resistent gegen einen weiteren Lithografieprozeß. Das Metall für die Justiermarken kann nun in gewünschter Dicke aufgedampft werden und anschließend erfolgt die Strukturierung der Marken durch den gleichzeitigen Lift-off-Prozeß beider Lackmasken. Der Prozeßablauf wird danach, wie oben beschrieben, mit der Aktivierung des implantierten Dotierstoffes fortgeführt. Durch diese Maßnahme konnte die Anzahl der Transistoren, die aufgrund von

Gate-Source-Kurzschlüssen ausfallen, um mehr als 90 % verringert werden. Außerdem stieg die Reproduzierbarkeit der Transistoreigenschaften über den Wafer und innerhalb eines Waferlots. Der Gesamttechnologie gewinnt dadurch an Stabilität gegenüber prozeßbedingten Schwankungen und steigert die Ausbeute an funktionierenden Transistoren.

In Kapitel 3 wurde anhand von Formeln explizit dargestellt, welchen signifikanten Einfluß die extrinsischen Widerstände R_s und R_d auf das Gleichstrom- und Hochfrequenzverhalten des Transistors haben. Obwohl üblicherweise zu konzentrierten Elementen zusammengefaßt, bestehen sie bei näherer Betrachtung des Transistoraufbaus aus mehreren Anteilen.

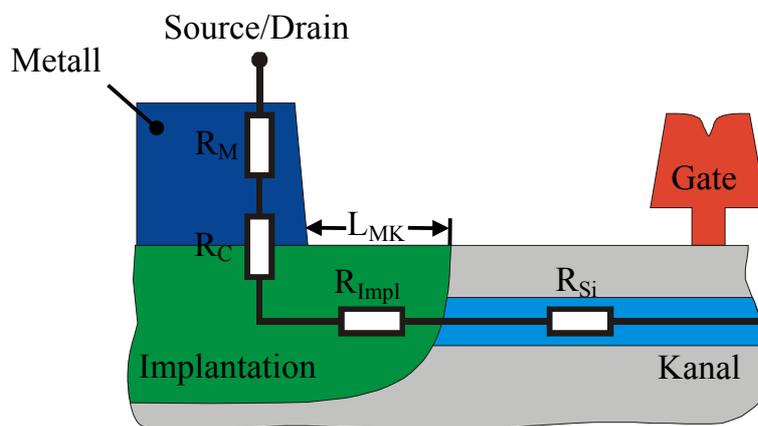


Abbildung 5.4: Schematische Darstellung der einzelnen Komponenten der extrinsischen Widerstände R_s und R_d

Nach Abbildung 5.4 setzt sich der Gesamtwiderstand R_s bzw. R_d aus dem Metallwiderstand R_M der ersten Kontaktierungsebene, dem Kontaktwiderstand R_C zwischen der ohmschen Metallisierung und dem hochdotierten n^+ -Halbleiter, dem Bahnwiderstand R_{Impl} des Implantationsgebietes und dem Widerstand des ungesteuerten Kanalgebietes R_{Si} zusammen.

Die Maßnahmen zur Reduzierung des Schichtwiderstandes für den nicht vom Gate kontrollierten Kanalbereich wurden bereits ausführlich in Kapitel 4 diskutiert. Für die Verringerung der anderen Widerstandsanteile bietet sich der Einsatz eines Silizides an. Während auf die konkrete technologische Realisierung in Kapitel 6 eingegangen wird, soll hier nur die konzeptionelle Integration der Silizi-

dierung in den Prozeßablauf beleuchtet werden. Der einfachste Weg besteht darin, das Silizidierungsmetall direkt nach der Implantation auf die bestehende Fotolackmaske aufzudampfen. Dadurch wird die laterale Ausdehnung der zu silizidierenden Fläche auf das Implantationsgebiet beschränkt. Nach dem Lift-off der gemeinsamen Maske erfolgt dann ein einzelner Tempersschritt. Dadurch läuft der Prozeß der Silizidierung gleichzeitig mit der Aktivierung des Dotierstoffes und der Rekristallisation des Atomgitters ab. Dieses Vorgehen erfordert die Auswahl eines Silizidierungsmetalls, dessen Aktivierungstemperatur im gleichen Bereich liegt wie die Ausheiltemperatur der Implantation. Die zweite Realisierungsvariante nutzt die Einführung einer dünnen Hilfsschicht aus Siliziumdioxid zur Entkopplung dieser physikalischen Abläufe. Dazu muß nach dem Öffnen des Feldoxidfensters und vor der Lithografie der Implantationsmaske eine 20 - 50 nm dicke Oxidschicht auf die Struktur aufgebracht werden. Die nachfolgend erzeugte Implantationsmaske dient mit Hilfe einer naßchemischen Ätzung zur Übertragung der Implantationsfenster in die dünne Oxidschicht. Nach der Implantation wird die Lackmaske entfernt und die Implantation bei optimaler Temperatur aktiviert. Der Silizidierungsprozeß erfolgt danach separat durch Aufdampfen des Metalles und anschließendes Tempern. Dabei bewirkt die bestehende Oxidmaske eine lokale Silizidierung des Metalls in den Implantationsgebieten. Um Kurzschlüsse durch das nicht silizidierte Metall auf dem Oxid zu vermeiden, muß es nach der Temperung selektiv entfernt werden. Der Nachteil dieses Verfahrens liegt in dem zusätzlichen Arbeitsaufwand durch das Aufbringen, Strukturieren und Entfernen der Hilfsschicht. Der Vorteil ist, daß zwei getrennte Tempersschritte stattfinden, die in ihrem Temperaturbereich individuell zum einen auf die Erfordernisse der Implantation und zum anderen auf den gewünschten minimalen Widerstand des Silizidierungsmetalls abgestimmt werden können.

Eine weitere Verringerung der extrinsischen Widerstände und eine Erhöhung der Packungsdichte läßt sich durch eine Modifikation des Layouts erzielen. In Abbildung 5.2 ist zu erkennen, daß der nominelle Abstand zwischen den Implantationsgebieten 2,0 - 2,5 µm beträgt. Für die sichere und reproduzierbare Herstellung eines e-Beam-Gates mit einer maximalen Kopfbreite von 600 nm ist aber ein Abstand von 1 µm vollkommen ausreichend. Dadurch ist es möglich, den Widerstandsanteil des ungesteuerten Kanals um bis zu 60 % zu senken.

Eine effektive Maßnahme, die Steilheit des Transistors zu erhöhen und darüber auf das Hochfrequenzverhalten Einfluß zu nehmen, ist die Reduzierung der Gatelänge. Durch die gezielte Abstimmung der Prozeßparameter bei der e-Beam-

Lithografie konnte im Rahmen dieser Arbeit die Gatelänge für das konventionelle Transistorkonzept von 500 nm auf 100 nm reduziert werden. Die Minimierung der Gatelänge bei gleichzeitiger Umsetzung eines niedrigen Gatewiderstandes stellt jedoch erhöhte Anforderungen an die technologische Realisierung. Da sich die Forderung nach einer möglichst großen Querschnittsfläche nicht direkt mit einer permanent sinkenden Gatelänge vereinbaren läßt, kommt das T-Gate Konzept zum Einsatz. Doch die mechanische Stabilität eines schweren, großflächigen Gatekopfes auf einem sehr schmalen Sockel setzt auch diesem Konzept physikalische Grenzen. Da der Trend nach ständig kürzeren Gatelängen auch in Hinblick auf die Packungsdichte dominierend ist, weisen die Gates im statischen Grenzbereich einen steigenden Gatewiderstand auf. Außerdem erschwert die mechanische Sensibilität der Gatekonstruktion die nachfolgende Prozessierung, da aufgebrauchte Passivierungsschichten oder zusätzliche Metallisierungsebenen ein Abknicken oder Umkippen des Gates hervorrufen können. Ein weiterer Schwachpunkt bei der Herstellung von T-Gates mit der e-Beam-Lithografie ist die proportional zur Gatelänge sinkende Höhe des Gatefusses. Dadurch rückt der breite Gatekopf näher an die Halbleiteroberfläche heran, und die kapazitive Kopplung zwischen den überlappenden Gateflanken und der Oberfläche bzw. den freien Ladungen des Kanals steigt. Diese Überlappkapazitäten C_{gso} und C_{gdo} sind Teilkapazitäten von C_{gs} und C_{gd} und haben aus diesem Grund direkten Einfluß auf das Hochfrequenzverhalten des Transistors. Die optimale Gatekonstruktion eines Hochfrequenztransistors sollte deshalb die individuelle Dimensionierung der Gatefußlänge und -höhe ermöglichen und gleichzeitig, ohne eine Gefährdung der mechanischen Stabilität, die Reduzierung des Gatewiderstandes zulassen.

Mit den beschriebenen Optimierungsansätzen für das konventionelle Transistorkonzept konnten neben der Reproduzierbarkeit und der Ausbeute auch das elektrische Leistungspotential der HFETs gesteigert werden. Trotz größerer Gatelänge und vergleichbarer Schichtstruktur erhöhten sich die von Glück [5.1] erzielten Werte der Grenzfrequenzen für f_T und f_{max} um 30 - 50 % [5.2].

5.2 Das selbstjustierende Integrationskonzept

Das hier vorgestellte, komplett neu entwickelte, HFET-Konzept nutzt das Prinzip einer voll-selbstjustierten Source/Gate/Drain-Anordnung, um eine optimierte Lateralstruktur für das Bauelement zu erzeugen [5.3], [5.4]. Hinzu kommt eine

flexible Realisierung der Gateelektrode, die eine Schottky-Gate, MOS- oder pn-Junction-Variante des Transistors ermöglicht. Der folgende Absatz konzentriert sich zunächst auf die relevante Schottky-Gate Version des HFETs. Die Prozeßmodifikationen zur Umsetzung von MOS- und pn-Junction-Transistoren werden später eingehender betrachtet. Abbildung 5.5 zeigt den schematischen Querschnitt und den Prozeßablauf zur Herstellung des Bauelementes mit Schottky-Gate.

Wie beim konventionellen Integrationskonzept startet der Prozeßablauf auf der Basis des spezifischen SiGe-Hetero-Schichtsystems mit der Definition der Bauelementmesa. Anschließend kommt eine selbstjustierende Methode zur Strukturierung des Feldoxides zur Anwendung, um eine möglichst planare Bauelementtopografie zu erhalten. Dabei überlappt die Oxidschicht nicht mehr die Mesakanten, sondern schließt bündig mit ihnen ab. Als nächstes wird eine Dummy- oder Replacement-Gate-Struktur erzeugt, die eine selbstjustierende Anordnung der Source/Drain-Gebiete ermöglicht. Dazu dient eine ganzflächig aufgebrachte Polyimidschicht, die sich präzise und reproduzierbar mit einer Metallmaske strukturieren läßt. Die mit dem Kontaktbelichtungsverfahren definierten Gatestege haben eine minimale Breite von 400 - 500 nm und legen den späteren Abstand der Source/Drain-Bereiche fest.

Um eine niederohmige Kontaktierung der Implantationsgebiete zu erzielen, kommt ein selbstjustierender Silizidierungsprozeß zum Einsatz. Dafür wird zuerst das gewünschte Silizidierungsmetall ganzflächig aufgedampft und anschließend mit einem Temperschritt lokal an der Siliziumoberfläche in ein hochleitendes Silizid umgewandelt. Auf dem maskierenden Dummy-Gate aus Polyimid und dem Feldoxid bleibt das Metall in seiner Ursprungsform erhalten und kann nach chemisch selektiv zum Silizid wieder entfernt werden. Dieser Ätzprozeß schrumpft gleichzeitig das Gate in seinen Abmessungen, so daß je nach Ätzdauer eine laterale Trennung zwischen dem silizidierten Bereich und der nachfolgenden Implantation eingestellt werden kann. Dadurch läßt sich gewährleisten, daß die Implantationsgebiete stets eine größere laterale Ausdehnung haben als das Silizid und damit die Kontaktierung des Kanals immer ohmschen Charakter hat. Nach der Implantation der Source/Drain-Gebiete, bei der das Dummy-Gate und das Feldoxid als Maske dienen, wird der Polyimidsteg einer erneuten Schrumpfung ausgesetzt. Dabei erfolgt mit Hilfe einer trockenchemischen Rückätzung die Definition der endgültigen Gatelänge, die im Bereich von 75 - 150 nm liegt.

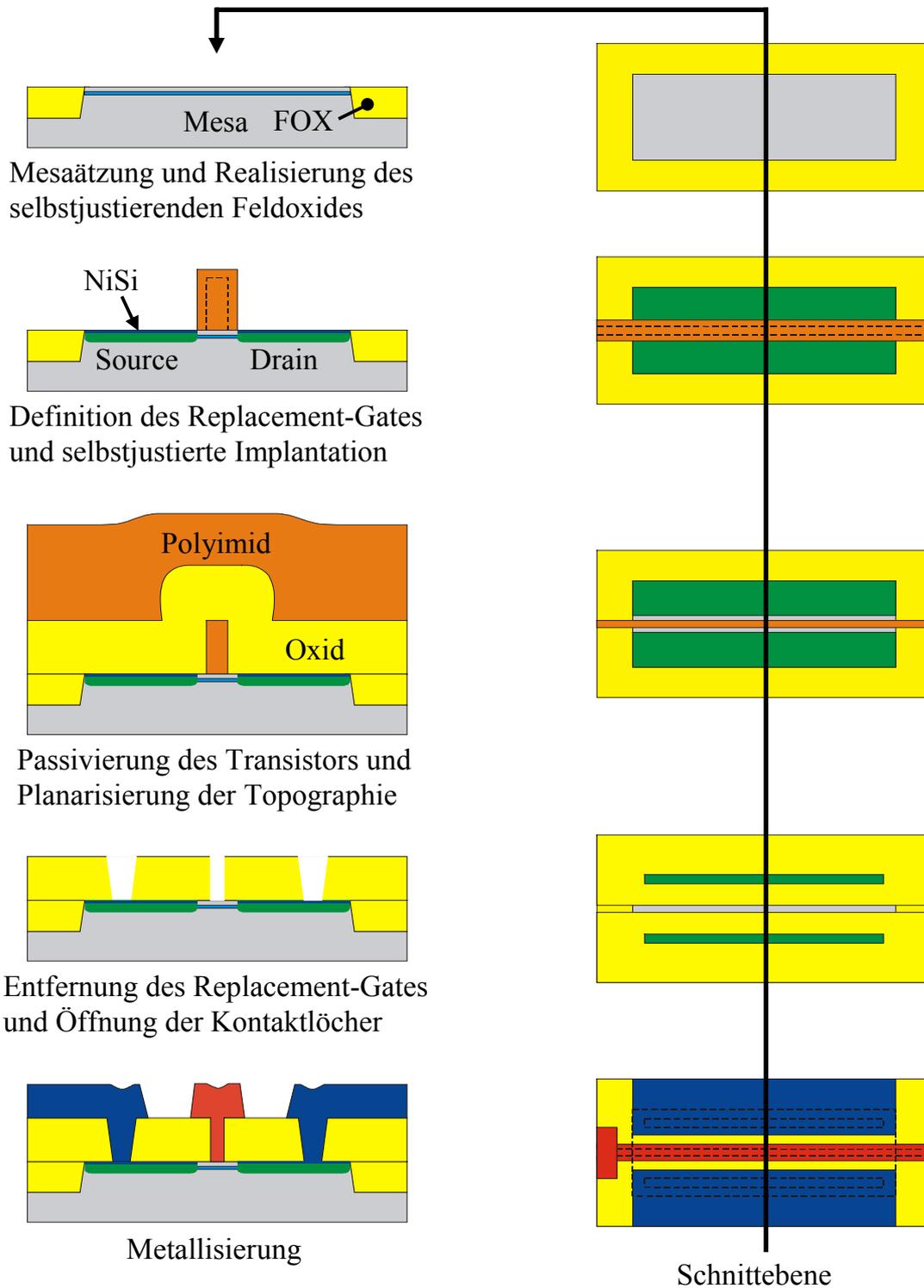


Abbildung 5.5: Prozeßablauf des selbstjustierenden Integrationskonzeptes im Bauelementquerschnitt und als Maskenebenen

Zur Isolation der aktiven Bauelementoberfläche wird anschließend eine Siliziumdioxidschicht abgeschieden, deren Dicke im Bereich der Dummy-Gate-Höhe liegt. Der nachfolgende Planarisierungsschritt sorgt für eine ebene Transistorstruktur, bevor durch ein gleichmäßiges Rückätzen der Oberfläche bis zur Obergrenze des Polyimidsteges die Gatestruktur wieder freigelegt wird. Die Wahl von Polyimid als Material für das Dummy-Gate gewährleistet die schadungsfreie und vollständige Entfernung der Hilfskonstruktion. Vor dem Auffüllen des entstandenen Gategrabens mit dem eigentlichen Gatematerial erfolgt noch die Aktivierung der Kontaktimplantation. Als prädestiniertes Metall für das Schottky-Gate kommt wieder Pt zum Einsatz. Zum einen bildet es die höchste Barriere zu Silizium, zum anderen zeigt es gute Eigenschaften, um die schmalen, aber hohen Gategräben reproduzierbar aufzufüllen. Materialkombinationen aus PtSi oder PtTi sind ebenfalls denkbar, da sie Vorteile bei der thermischen Stabilität und bei Langzeittests zeigen. Nachdem die Kontaktlöcher für die Source/Drain-Gebiete entstanden sind, kann die erste Metallisierungsebene, die den Gatekopf und die Zuleitungen bildet, aufgebracht werden. Die Strukturierung des Kontaktmetalls, das wahlweise aus Al oder Au bestehen kann, geschieht anschließend mit Hilfe einer fotolithografischen Maske und einem Ätzprozeß. Das Aufbringen weiterer Metallisierungsebenen für Schaltungsanwendungen oder die Integration passiver Bauelemente ist ohne besondere Präparation der Transistorstruktur möglich.

Durch einfache Modifikation des Prozeßablaufes kann das Bauelement auch als MOS-Transistor ausgeführt werden. Dazu muß nach der Realisierung des selbstjustierenden Feldoxides und vor dem Aufbringen der Polyimidschicht für das Dummy-Gate die Herstellung eines hochwertigen, thermischen Gateoxides oder eines high- k Dielektrikums erfolgen. Der weitere Prozeßablauf ist identisch mit der bereits beschriebenen Schottky-Gate Version. Nach dem schadungsfreien Entfernen der Dummy-Struktur kann die eigentliche Gateelektrode aus Polysilizium, Metall oder einer Kombination aus beiden auf das freigelegte Gateoxid abgeschieden werden. Durch die entkoppelte Herstellung von Gatefuß und Gatekopf ist eine Ausführung des MOS-Transistors mit T-Gate-Struktur problemlos möglich. Einer Anwendung als SiGe-Hochfrequenz-MOS-Bauelement steht deshalb nichts im Weg.

Da die reduzierte Spannungsfestigkeit von zunehmend dünner werdenden Gateoxiden und insbesondere die Barrierenhöhe des Schottky-Gates den Arbeitsbereich der Transistoren stark einschränkt, sind für Leistungsanwendungen pn-Junction-FETs von Interesse. Die Durchbruchsspannung dieser Transistoren hängt neben dem spezifischen Layout in erster Linie von der Spannungsfestigkeit

des steuernden pn-Überganges ab. Diese Bauelementvariation basiert ebenfalls auf der Standardschichtstruktur des modulationsdotierten SiGe-HFETs, schließt jedoch mit einer zusätzlichen, etwa 15 nm dicken, hochdotierten p⁺-Schicht ab. Sie bildet mit den darunterliegenden, undotierten Si- und SiGe-Schichten einen pn-Übergang, dessen Raumladungszone sich vorwiegend in Richtung Potentialtopf ausdehnt und damit den Stromfluß im Kanal steuert. Während des Prozeßablaufes muß die p⁺-Schicht außerhalb des Gatefusses entfernt werden, um Kurzschlüsse zwischen der Gateelektrode und den ohmschen Kontaktgebieten zu vermeiden. Der technologisch sinnvollste Zeitpunkt dafür ist nach der Schrumpfung des Dummy-Gates auf die endgültige Gatelänge. Dabei dient der provisorische Gatestege als Maske für die Ätzung und überträgt die effektive Gatelänge auf den pn-Übergang. Der weitere Herstellungsprozeß ist dann wieder identisch mit der Originalversion für den Schottky-Gate gesteuerten Transistor. Wie bei der MOS-Variante bietet sich hier auch wieder die Möglichkeit einer T-Gate-Konstruktion aus Metall an, um die verbesserte Spannungsfestigkeit auch bei hohen Arbeitsfrequenzen verfügbar zu machen.

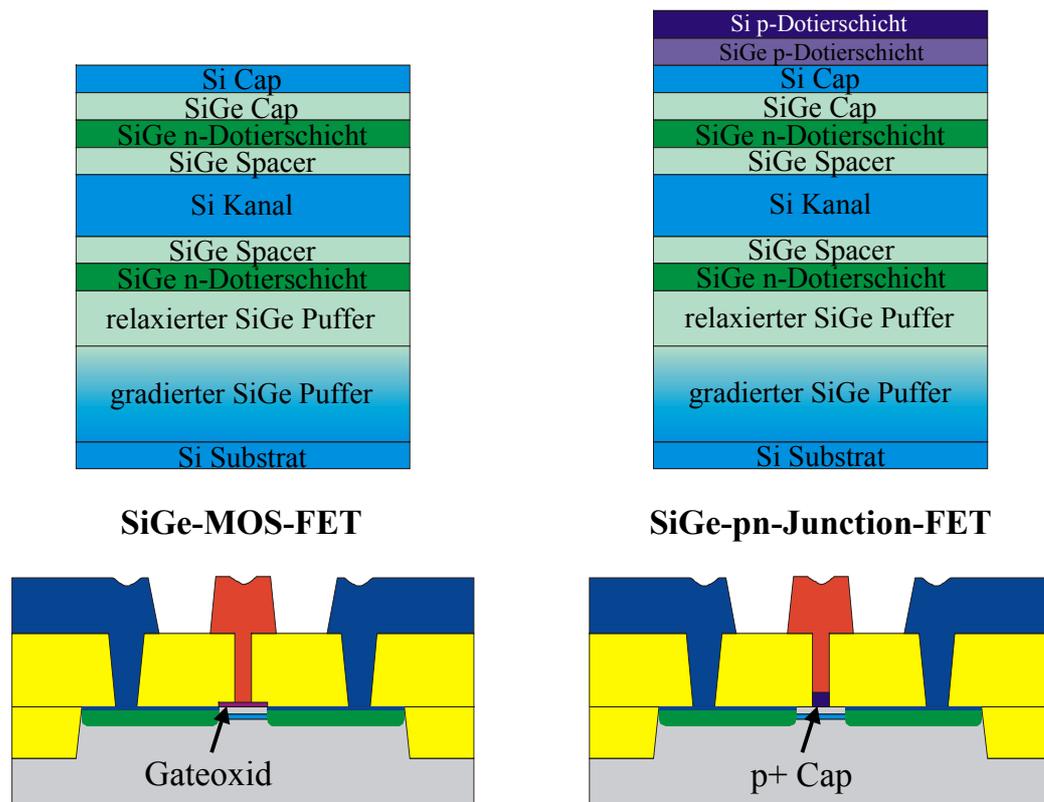
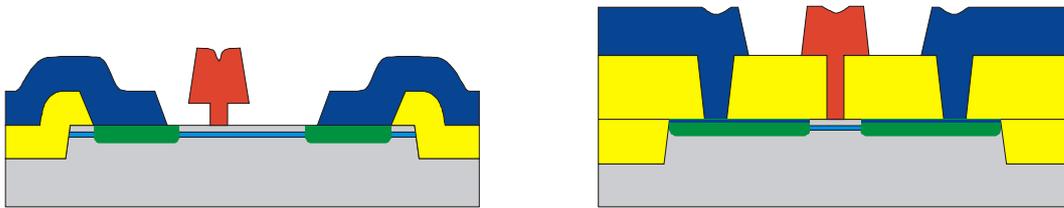


Abbildung 5.6: Schematischer Querschnitt einer MOS- und einer pn-Junction-FET Variante mit dem selbstjustierenden Integrationskonzept

5.3 Vergleich der SiGe-HFET Technologievarianten

Die Forderung nach immer kürzeren Schaltzeiten, steigenden Grenzfrequenzen und höheren Packungsdichten läßt auch den Miniaturisierungsgrad von Hochfrequenzbauelementen permanent ansteigen. Den Strukturverkleinerungen sind jedoch durch die herkömmlichen Lithografieverfahren Grenzen gesetzt. Deshalb soll durch den Einsatz von intelligenten Integrationsprozessen mit einem hohen Anteil an selbstjustierenden Verfahren den hohen Leistungsansprüchen nachgekommen werden. Das oben dargestellte, neue Integrationskonzept realisiert mit Hilfe einer Dummy-Gate-Konstruktion die hochoptimierte Lateralstruktur eines Transistors für Hochfrequenzanwendungen. Durch die Anordnung der Source/Drain-Gebiete in einem selbstjustierenden Prozeß minimieren sich die Zuleitungsstrecken zum 2DEG und damit die extrinsischen Widerstandsanteile des ungesteuerten Kanalbereiches. Außerdem eröffnet die Dummy-Gate Konstruktion die strukturellen Randbedingungen für die Durchführung eines Salizidierungsprozesses, der zusätzlich zur Reduzierung der parasitären Widerstände beiträgt. Die gewählte Integrationsstrategie stellt dabei das Gate in den Mittelpunkt des Konzeptes und richtet alle weiteren Komponenten der aktiven Struktur danach aus.

Das Vorgehen beim konventionellen Herstellungsprozeß ist grundlegend anders. Hier werden zuerst Position und Abmessungen der Source/Drain-Gebiete nach den starren Vorgaben des Layouts festgelegt und im letzten Schritt das Gate mehr oder weniger exakt im vorgegebenen Bereich plaziert. Dieses Prinzip birgt jedoch durch die manuelle Justage der Maskenebenen viele Ungenauigkeiten in sich und ist in seiner Flexibilität stark durch die Randbedingungen des Layouts eingeschränkt. Beim selbstjustierenden Konzept wird mit der Breite der Dummy-Gate-Struktur am Anfang eine layoutgebundene Größe vorgegeben. Alle weiteren Abmessungen sind unabhängig voneinander individuell einstellbar. Die reale Gatelänge zum Beispiel richtet sich nach dem Schrumpfungsgrad des Polyimidsteges. Unabhängig davon kann die Höhe des Gatefusses und damit ein wichtiger Faktor für die Kapazitätsanteile C_{gs0} und C_{gdo} über die Rückätzung der Isolationsschicht reguliert werden. Die Höhe und Breite des Gatekopfes richtet sich schließlich nach der aufgetragenen Dicke der ersten Metallisierungsebene und den geometrischen Abmessungen der Strukturierungsmaske. Dabei sind alle Lithografieschritte lichtoptisch durchführbar, was den Einsatz der teureren und zeitintensiven Elektronenstrahlithografie vermeidet. Die damit realisierten Gatelängen liegen mit bis zu 75 nm, aufgrund des gut abgestimmten Schrumpfungsprozesses, unter den erzielten Werten für die e-Beam-Lithografie.



| konventioneller Prozeß | Kriterien | selbstjustierender Prozeß |
|-------------------------------|--------------------------------------|---|
| Mesaätzung | Separation der Bauelemente | Mesaätzung |
| überlappend | FOX | selbstjustierend |
| abhängig vom Maskensatz | laterale Optimierung | selbstjustierend zum Gate |
| e-Beam-Lithografie | Gate Definition | optische Lithografie |
| ↓ | Grad d. Planarisierung | ↑ |
| nur außerhalb der Mesa | Passivierung des Transistors | ganzflächig |
| ↑ | beanspruchte Fläche | ↓ |
| ↓ | Prozeßkomplexität | ↑ |
| ↓ | Prozeßflexibilität | Ausführung als MOS- und pn-Junction-FET möglich |
| schwierig | Integrationsfähigkeit in Schaltungen | problemlos |

Abbildung 5.7: Funktionaler und technologischer Vergleich der beiden Integrationskonzepte (↑ = hoch, ↓ = niedrig)

Die selbstjustierende Gatekonstruktion zeichnet sich durch eine erhöhte mechanische Stabilität aus, da der Gatefuß, im Gegensatz zum freistehenden e-Beam-Gate, in einer isolierenden Oxidschicht eingebettet ist. Deshalb kann die Größe des Gatekopfes auch unabhängig von der Gatelänge dimensioniert werden. Die den Gatefuß umschließende SiO_2 -Schicht sorgt jedoch nicht nur für die Standfestigkeit des Gates, sondern hat in erster Linie eine passivierende und planarisierende

rende Aufgabe für den Transistoraufbau. Zusammen mit dem selbstjustierenden Feldoxid beinhaltet das Integrationskonzept damit zwei Prozeßmodule, die eine voll-isolierte und planarisierte Bauelementstruktur erzeugen sollen. Das schafft die Basis für eine problemlose Integration des Einzeltransistors in eine komplexe Schaltung, die meistens über mehrere Metallisierungsebenen realisiert wird.

Das selbstjustierende Feldoxid hat noch einen vorteilhaften Nebeneffekt, der sich positiv auf die elektrischen Eigenschaften des Transistors auswirkt. Durch den ebenen Abschluß der Oxidschicht mit der oberen Mesakante lassen sich die Leckströme des HFETs im abgeschnürten Zustand drastisch verringern. Bei nicht-selbstjustierenden Techniken überlappt das Feldoxid entweder die Mesakante oder endet bereits außerhalb der Mesa. Im überlappenden Randbereich hat das Gate durch die mehrere 100 nm dicke Oxidschicht aber keine Steuerwirkung mehr auf die Ladungsträger im Kanal, so daß der Stromfluß bei „normally-on“-Transistoren nicht vollständig unterdrückt werden kann. Falls das Feldoxid außerhalb der Mesa endet, läuft das Gate über die nicht isolierten Mesaflanken und hat darüber Kontakt zum hochleitenden Siliziumkanal, was ebenfalls zu einem nicht kontrollierbaren Stromfluß führt.

Ein weiterer Vorteil des selbstjustierenden Verfahrens besteht in dem Potential, drei verschiedene Bauelementvarianten mit nur geringfügigen Änderungen im Prozeßablauf zu realisieren. Eine Modifikation des Maskenlayouts ist dafür nicht erforderlich, was die multifunktionale Anwendbarkeit des Prozesses unterstreicht.

Zusammengefaßt läßt sich festhalten, daß das konventionelle Integrationskonzept prädestiniert ist für die einfache und schnelle Realisierung von SiGe-HFETs. Die Prozeßführung ist unkompliziert und die Anforderungen an die technologische Ausstattung, die Elektronenstrahlithografie ausgenommen, sind gering. Jedoch sind der Optimierung des Prozesses Grenzen gesetzt, da er sehr layoutgebunden ist. Deshalb ist das ideale Einsatzgebiet die Herstellung von Einzelbauelementen und das Schnelltestverfahren, um das Leistungspotential neuer SiGe-Puffer und aktiver Schichtsysteme zu ermitteln. Das selbstjustierende Konzept erlaubt die flächen- und anwendungsoptimierte Integration von Hochfrequenztransistoren in drei verschiedenen Bauelementvarianten. Die hohe Anzahl an Freiheitsgraden bei der Realisierung der einzelnen Prozeßmodule ermöglicht die gezielte Anpassung der Transistorstruktur an die gestellten Leistungsanforderungen. Durch den Einsatz von selbstjustierenden Verfahren ergibt sich ein lateraloptimierter Bauele-

mentaufbau mit Gatelängen im sub-100 nm Bereich. Dabei verzichtet der Technologieprozeß auf die teure und zeitintensive Elektronenstrahlithografie, die für jeden Herstellungsablauf einen Engpaß darstellt. Die komplett passivierte und planarisierte Bauelementstruktur prädestiniert das Konzept für die Herstellung von anwendungsspezifischen Einzelbauelementen und die Integration in Schaltungen.

Kapitel 6

Technologische Realisierung der Herstellungskonzepte

Das folgende Kapitel beschreibt die im Rahmen dieser Arbeit durchgeführte Technologieentwicklung der wichtigsten Herstellungsschritte und Prozeßmodule, um die in Kapitel 5 vorgestellten Integrationskonzepte für SiGe-HFETs realisieren zu können. Da mit fortschreitender Miniaturisierung der Bauelementabmessungen auch der technologische Aufwand wächst, konzentriert sich dieses Kapitel vorwiegend auf Prozeßmodule, die zur Optimierung der lateralen Transistorstruktur beitragen und die Integrationsfähigkeit des Bauelementes verbessern. Um die Zuordnung zu den Technologiekonzepten zu vereinfachen, zeigt Abbildung 6.1, welche Integrationskomplexe von der konventionellen und der selbstjustierenden Herstellungsvariante in ihrer Entwicklung betrachtet werden.

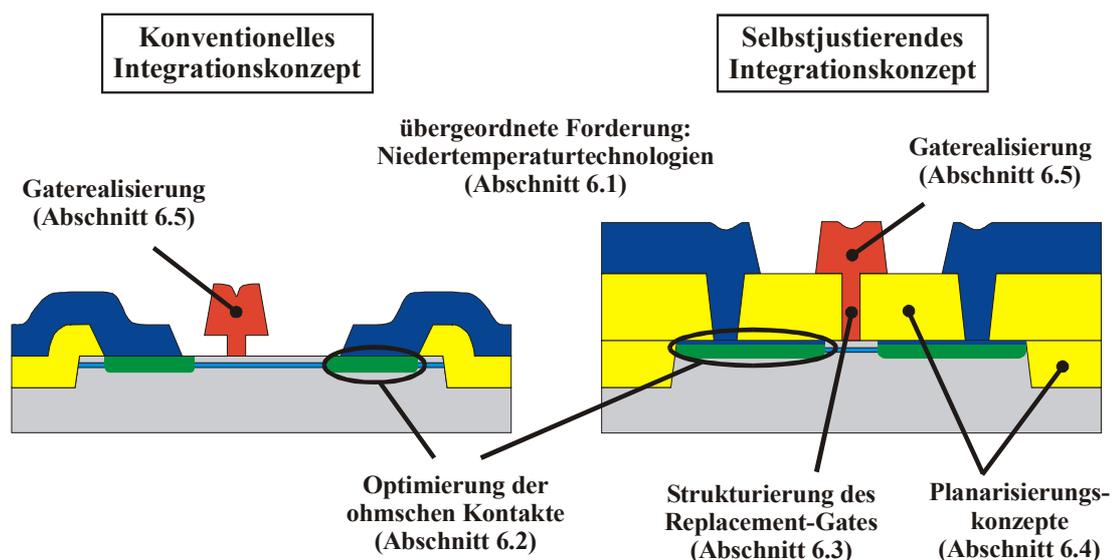


Abbildung 6.1: Übersicht der Prozeßmodule, deren Entwicklung in diesem Kapitel beschrieben wird

Dabei gibt es teilweise verschiedene Lösungsansätze, um den Leistungsanforderungen des späteren Anwendungsgebietes zu entsprechen. Ein Vergleich der technologischen Realisierungsmöglichkeiten hinsichtlich Komplexität, Reproduzierbarkeit und Ausbeute soll darüber Aufschluß geben, welcher Ansatz unter den gegebenen Randbedingungen Vorteile in den jeweiligen Bereichen bringt.

Zuerst soll als Basis für alle Integrationsmodule der Rahmen des Temperaturbudgets bei der Technologieentwicklung abgesteckt werden. Dabei steht die Konservierung der charakteristischen elektrischen Eigenschaften der SiGe-Heterostruktur durch den Einsatz von Niedertemperaturprozessen im Vordergrund. Die niederohmige Kontaktierung des Kanals trägt maßgeblich zur Optimierung des Transistorverhaltens bei und ist für beide Integrationskonzepte von entscheidender Bedeutung. Hier werden Variationen in den Implantationsparametern, unterschiedliche Aktivierungstemperaturen und der Einsatz und die Realisierung von Siliziden eingehender betrachtet. Anschließend ist die Strukturierung des Dummy-Gates für die selbstjustierende Technologie Gegenstand der Diskussion. Die Parameter des eingesetzten Ätzverfahrens haben dabei entscheidenden Einfluß auf die Gatelänge, die Form des Gatefusses und die spätere Auffüllung der Gategräben mit Metall. Für die Umsetzung selbstjustierender Strukturen und die Bereitstellung planarer, gut integrierbarer Transistortopografien sind Planarisierungsverfahren unverzichtbar. Deshalb sollen mit dem selbstjustierenden Feldoxid und der Planarisierung der Oxidpassivierung zwei Prozeßmodule vorgestellt werden, die mit unterschiedlichen Konzepten und ohne kostspielige technologische Ausstattung den gestellten Anforderungen nachkommen. Abschließend steht die komplett unterschiedliche Realisierung des T-Gates im Mittelpunkt der Betrachtung. Dabei wird zum einen die einstufige Lösung mit Hilfe der e-Beam-Lithografie diskutiert und zum anderen die unabhängige Herstellung von Gatefuß und Gatekopf für den selbstjustierenden Prozeß beschrieben.

6.1 Niedertemperaturtechnologien für SiGe-HFETs

Die modulationsdotierte SiGe-Heterostruktur des HFETs zeichnet sich durch einen pseudomorphen, metastabilen Si-Kanal, hochdosierte, lokal begrenzte Antimon-Dotierungen und steile Ge-Gradienten aus. Diese spezifische Auslegung des Schichtsystems sorgt für eine große Anzahl von hochbeweglichen Ladungsträgern im Kanal und den damit verbundenen elektrischen Qualitäten für das Bauelement. Andererseits zeigen solche technisch hochentwickelten Systeme, die

zusätzlich sehr abrupte Übergänge in den Materialeigenschaften aufweisen, immer auch eine erhöhte Sensibilität gegenüber technologischen Einflüssen. Um die Wirkung typischer Prozessschritte der Bauelementfertigung auf die charakteristischen Schichtparameter herauszufinden, bietet sich die Möglichkeit, anhand des Schichtwiderstandes, der Ladungsträgerdichte und der Beweglichkeit Qualitätsänderungen an der Heterostruktur nachzuweisen und zu quantifizieren. Das folgende Experiment legt den Prozessablauf des konventionellen Integrationskonzeptes zugrunde und betrachtet naßchemische, trockenchemische und thermische Technologiesequenzen, die auf das Schichtsystem bis zur Fertigstellung des Transistors einwirken. Das Testobjekt ist eine typische modulationsdotierte Si-Ge-Heterostruktur mit einem Germaniumgehalt x von 45 %. Die Schichtfolge zeichnet sich durch einen 9 nm dicken Si-Kanal aus, der zwischen 3,5 und 4 nm breiten, undotierten SiGe-Spacern eingebettet ist. Der Kanal wird durch eine Vorder- und Rückseitenantimondotierung mit Ladungsträgern versorgt, und die gesamte Dicke der Cap-Schichten oberhalb der vorderen Dotierschicht beträgt 10 nm, wobei 4 nm auf die oberste Si-Schicht abfallen. Die Untersuchungen finden an mehreren Proben des Wafers statt, die alle identische charakteristische Schichtparameter zeigen. Alle Proben durchlaufen die gleichen Prozessschritte, nur die thermische Aktivierung der implantierten Ladungsträger im RTA-System läuft für jede Probe bei einer anderen Temperatur ab. Nach jedem technologischen Bearbeitungsschritt erfolgt die Charakterisierung der Schichteigenschaften, und die Änderungen zum vorherigen Zustand werden erfaßt. In Abbildung 6.2 sind die Ergebnisse der Analyse für den Schichtwiderstand und die Ladungsträgerbeweglichkeit in Abhängigkeit von dem durchgeführten Prozessschritt dargestellt.

Üblicherweise startet der Herstellungsprozeß mit einer naßchemischen Reinigung des Wafers in einem Gemisch aus Schwefelsäure und Wasserstoffperoxid. Bei einem Mischungsverhältnis von 5:1 werden für 2 - 3 Minuten Verunreinigungen und beim MBE-Wachstum aufgeschwemmte Antimonreste an der Waferoberfläche entfernt. Nach der Strukturierung der Mesa, bei der das aktive Transistorgebiet durch eine Fotolackmaske geschützt ist, findet ein weiterer Reinigungsschritt in einem Sauerstoffplasma bei 200 W Leistung statt. Dabei wird die Oberfläche für den nachfolgenden Feldoxidprozeß gesäubert und vorbereitet. Für die Passivierung der Bauelemente mit siliziumbasierten Oxidschichten steht ausschließlich eine PECVD-Anlage zur Verfügung. Die plasmaunterstützte Abscheidung des SiO_2 erfolgt bei einer Hochfrequenzleistung von 20 W und einer Elektrodentemperatur von 400 °C. Der Feldoxidprozeß beinhaltet außerdem die naßchemische Entfernung des Oxides in gepufferter Flußsäure. Dabei wird die Ober-

fläche der Heterostruktur je nach Überätzung 15 - 30 s einer 12,5 %igen Flußsäure ausgesetzt. Die Auswirkungen der Implantation auf das Schichtsystem werden durch eine Temperaturbehandlung berücksichtigt. Mit Hilfe von Kurzzeittemperprozessen in der RTA-Anlage erfolgt dabei die Restaurierung des Kristallgitters und Aktivierung des Dotierstoffes. Die Dauer der Temperung beträgt jeweils 30 s, und die Temperatur staffelt sich zwischen 550 °C und 700 °C.

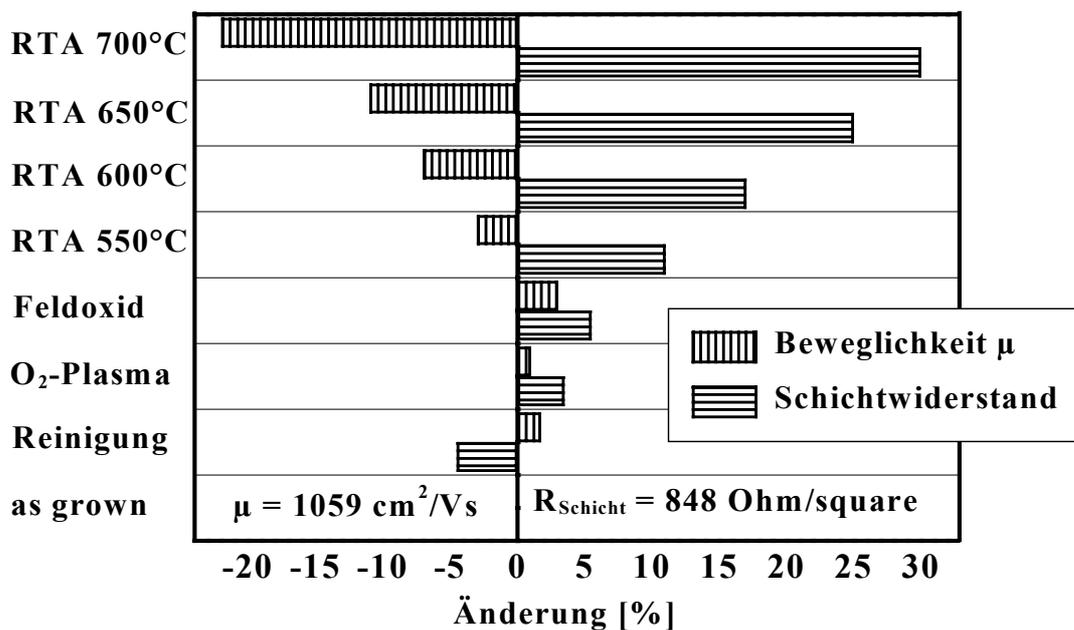


Abbildung 6.2: Änderung des Schichtwiderstandes und der Beweglichkeit bei 300 K in Abhängigkeit von dem durchgeführten Prozessschritt

Die Ergebnisse in Abbildung 6.2 zeigen deutlich, daß naß- und trockenchemische Reinigungsschritte und die Feldoxidprozessierung, wie sie im Rahmen des Herstellungsverfahrens des Transistors auftreten, zu einer tolerierbaren Veränderung des Schichtwiderstandes und der Beweglichkeit von maximal 5 % führen. Das Temperaturbudget erreicht bei diesen Technologiesequenzen sein Maximum von 400 °C bei der Deposition des Oxides. Da für modulationsdotierte SiGe-Heterostrukturen bereits unterhalb von 550 °C eine anfängliche Ausdiffusion des Antimons in den Dotierschichten nachgewiesen werden konnte [6.1], ist hier schon mit dem Auslaufen des Dotierprofils in die anliegenden SiGe-Schichten zu rechnen. Zwar scheint die Kanalbeweglichkeit davon noch unbeeinflusst zu sein, jedoch reduziert sich die Anzahl der Ladungsträger, die zum 2DEG beitragen

und damit der Schichtwiderstand der Struktur. Eine signifikante Reduzierung der Beweglichkeit und eine gleichzeitige Erhöhung des Schichtwiderstandes treten eindeutig mit steigender Aktivierungstemperatur während der Temperung auf. Schon bei einer Temperaturbelastung von 600 °C für 30 s liegt der Schichtwiderstand mehr als 20 % über dem Referenzwert der unbehandelten Schicht, und die Beweglichkeit nimmt im Vergleich zum Prozeßmodul für die Feldoxidstrukturierung um 7 % ab. Die bei 700 °C festgestellten Einbußen in der Beweglichkeit von ca. 20 % sind nicht mehr zu akzeptieren, so daß die maximale Prozeßtemperatur deutlich niedriger liegen muß.

Um die Eigenschaften des 2DEGs isoliert von der Reststruktur charakterisieren zu können, werden die Änderungen der Ladungsträgerdichte n_s und der Beweglichkeit μ_0 bei einer Temperatur von 77 K ermittelt (Abbildung 6.3).

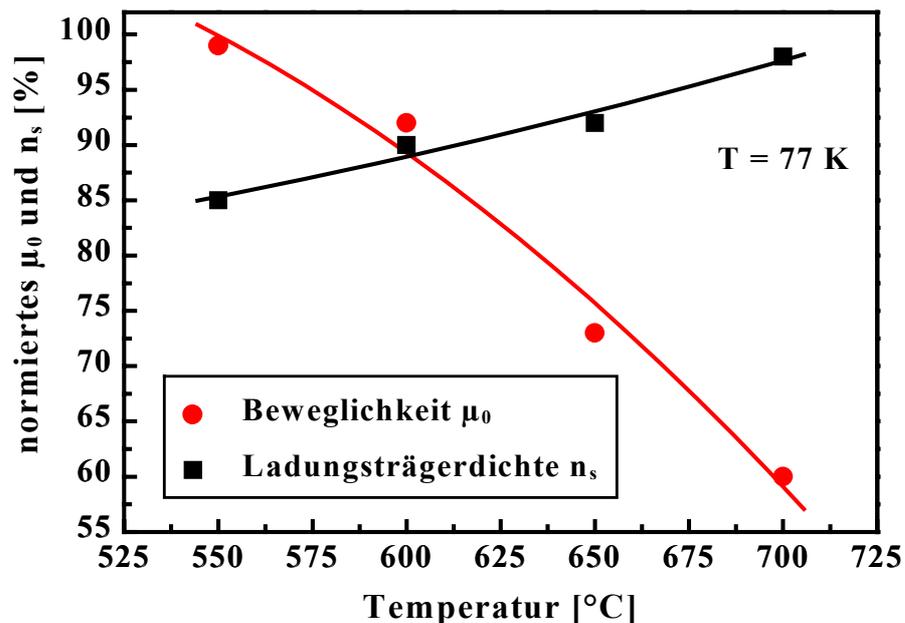


Abbildung 6.3: Auswirkungen von Kurzzeittemperversuchen mit einer Dauer von 30 s auf die Ladungsträgerdichte und die Beweglichkeit von modulationsdotierten SiGe-Heterostrukturen bei 77 K

Die Vorteile der Kryoanalyse bestehen darin, daß die Donatoren in den Dotierschichten nicht mehr ionisiert sind und deshalb keine parasitären Kanäle auftreten, die sich mindernd auf die Beweglichkeit des 2DEGs auswirken. Außerdem kann die Beweglichkeit im Kanal unabhängig von der Phononenstreuung und der

Hintergrunddotierung betrachtet werden, die beide einen maßgeblichen Anteil bei Raumtemperatur haben. Die Qualitätsminderung des 2DEGs durch technologische Prozessschritte tritt dadurch, gefiltert von parasitären Effekten, verstärkt in den Vordergrund. Im Gegensatz zur vorherigen Versuchsreihe steht diesmal nur der Einfluß der Kurzzeitemperschritte im Mittelpunkt, da die Auswirkungen der übrigen Prozeßabläufe vernachlässigbar sind.

Die Meßergebnisse in Abbildung 6.3 bestätigen die Erkenntnisse aus der Analyse des Gesamtprozesses. Bis zu einer Temperaturbelastung von 550 °C bleibt die Beweglichkeit des Kanals nahezu unbeeinflußt. Bei höheren Temperaturen kommt es jedoch zu einer signifikanten Verringerung der Beweglichkeit. Gleichzeitig steigt die bei 550 °C schon geminderte Ladungsträgerdichte mit zunehmender Temperaturbelastung kontinuierlich an. Setzt man die tolerierbare Obergrenze der Beweglichkeitsverringerng bei maximal 10 % an, so darf das Temperaturbudget bei der Prozessierung 600 °C nicht überschreiten.

Bei der theoretischen Analyse der beobachteten Effekte müssen drei physikalische Mechanismen berücksichtigt werden:

- Ausdiffusion des Antimons in den Dotierschichten
- Interdiffusion von Ge und Si
- Relaxation des metastabilen Kanals

Die bereits bei einer Temperaturbelastung von 550 °C um 15 % reduzierte Ladungsträgerdichte im Kanal deutet auf eine substantielle Ausdiffusion des Antimons in den Dotierschichten hin. Dabei kommt es zu einem Auslaufen des ursprünglich steilen Dotierstoffprofils in die umliegenden SiGe-Schichten. Die Verringerung des Antimonpeaks führt zu einem abgeschwächten Modulationseffekt und einer geminderten Ladungsträgerdichte im 2DEG. Mit zunehmender Temperaturbelastung diffundieren immer mehr Dotierstoffatome in Richtung Si-Kanal und sorgen für eine signifikante Beweglichkeitsminderung durch erhöhte Coulomb-Streuung. Gleichzeitig steigt dadurch aber auch der Anteil der Ladungsträger, die bis in den Kanal vordringen, so daß die Dichte des 2DEGs zunimmt.

Um den Temperaturbereich für eine Interdiffusion von Si und Ge im Grenzgebiet des Potentialtopfes abzuschätzen, wird auf Untersuchungen von Klauk et al. zurückgegriffen [6.2]. Sie zeigten anhand von Experimenten mit undotierten SiGe-

HFET-Strukturen, daß es erst bei Langzeittemperaturbelastungen von 900 °C zu starken Einbußen der Hallbeweglichkeit des verspannten Si-Kanals kommt. Gestützt werden diese Ergebnisse durch SIMS- und RBS-Analysen [6.1], [6.3], [6.4] und Untersuchungen mit der Röntgendiffraktometrie [6.5], die den Einsatz der Interdiffusion an verspannten SiGe-Heterostrukturen im Bereich von 850 °C bis 950 °C belegen. Der Diffusionsvorgang beginnt dabei an der Grenzschicht zwischen Si-Kanal und SiGe-Spacern und führt zu einer Änderung der Potentialtopfform. Durch ein Zusammenrücken der abgeflachten Quantentopfflanken steigt die Interface-Streuung, und die Beweglichkeit sinkt drastisch. Mit zunehmender Temperatureinwirkung geht der Si-Kanal dadurch in einen gleichmäßig dotierten SiGe-Mischkristall über, in dem eine dreidimensionale Leitfähigkeit der Ladungsträger vorherrscht.

Die Degradation der Schichtstruktur durch Relaxationseffekte läuft meistens begleitend zur Interdiffusion ab [6.3], [6.5]. Dabei führt die Ausbildung und Verzweigung von Kristalldefekten und aufsteigenden Versetzungslinien zu einer Verringerung der Schichtverspannung. Ursprung der Gitterstörungen ist häufig der nicht vollständig relaxierte Puffer. Da der pseudomorphe Si-Kanal oftmals im metastabilen Bereich gewachsen wird, ist er bei entsprechender Temperaturbelastung besonders anfällig für diese Relaxationsmechanismen. Eine Störung der regelmäßigen Gitterstruktur im Kanal hat wiederum eine drastische Minderung der Beweglichkeit zur Folge.

Frühere Untersuchungen [6.6] an modulationsdotierten SiGe-Heterostrukturen mit geringerem Ge-Gehalt kommen zu dem Ergebnis, daß auch Temperaturbudgets über 600 °C noch ohne wesentliche Leistungseinbußen zulässig sind. Diese Ergebnisse lassen vermuten, daß die SiGe-Komposition auf die Temperaturstabilität des Schichtsystems Einfluß hat.

Deshalb sind weiterführende Experimente gemacht worden, die von vergleichbaren modulationsdotierten Heterostrukturen mit unterschiedlichen Ge-Gehalten ausgehen. Die gemessenen Beweglichkeitsverläufe bei 77 K (Abbildung 6.4) beweisen, daß mit zunehmendem Ge-Anteil des Schichtsystems die Sensibilität gegenüber Hochtemperaturschritten steigt und es früher zu einer Minderung der Qualität kommt. Während bei einem x-Anteil von 30 % Temperprozesse von 650 °C noch vertretbar sind, muß die maximale Temperatur bei 46 % schon unter 600 °C liegen. Dies läßt sich primär auf eine verstärkte Ausdiffusion von Antimon in den Dotierschichten zurückführen. Durch eine Erhöhung des Ge-Gehaltes

in dem SiGe-Schichtsystem verringert sich die Aktivierungsenergie für die Antimondiffusion, und der Diffusionskoeffizient steigt [6.7]. Deshalb kommt es schon bei niedrigeren Temperaturbelastungen zu einer intensiveren Coulomb-Streuung an den ausdiffundierten Dotieratomen. Gleichzeitig zeigt sich für höhere Temperaturbereiche eine Zunahme der Ge-Interdiffusion mit steigendem Ge-Anteil. Dies begründet sich einerseits durch den größeren Ge-Gradienten in der Struktur, andererseits zeigt auch die Diffusionskonstante für Ge eine proportionale Erhöhung mit der Ge-Konzentration [6.4]. Aus diesen Gründen muß auch in Hinblick auf den anhaltenden Trend zu germaniumreicheren Heterosystemen das Temperaturbudget in der Prozeßführung überdacht werden. Insbesondere die Technologieparameter bei der Aktivierung und Rekristallisierung nach der Implantation und beim Silizidierungsprozeß müssen sich an die jeweiligen Randbedingungen anpassen. Das hat wiederum Einfluß auf die Auswahl des Dotierstoffes, die Implantationsdosis und -energie und natürlich auch auf das zu verwendende Silizidierungsmetall. Der sich daraus ergebende Optimierungsprozeß für die Realisierung der ohmschen Kontakte wird unter den hier gesteckten Randbedingungen in Abschnitt 6.2 ausführlich diskutiert.

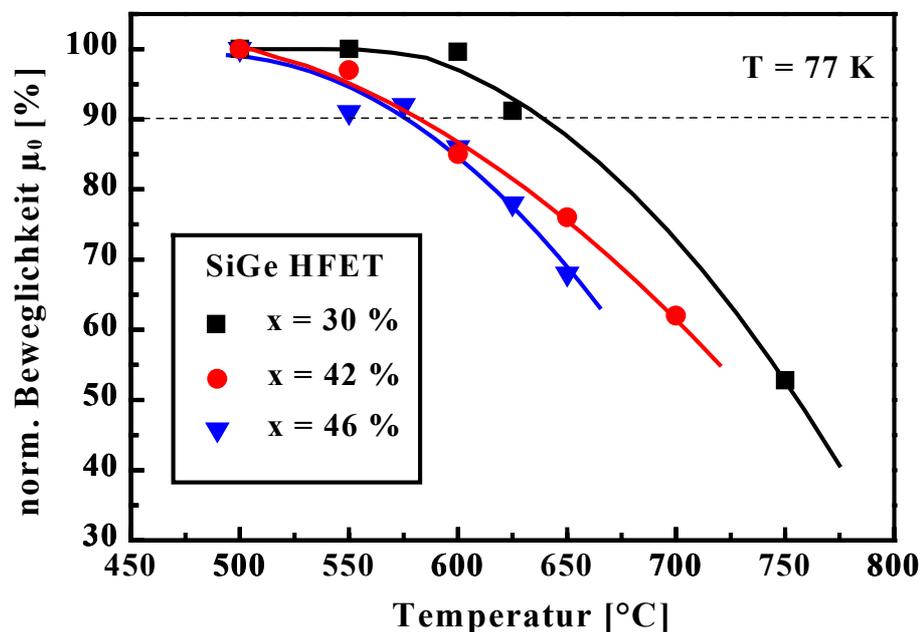


Abbildung 6.4: Abhängigkeit der Beweglichkeit μ_0 modulationsdotierter HFET-Strukturen vom Temperaturbudget der Kurzzeittemperatur (30 s) mit dem Germaniumgehalt x als Parameter

6.2 Optimierung der ohmschen Kontakte

Die große Bedeutung optimierter ohmscher Kontakte für das Leistungspotential des SiGe-HFETs ist explizit an den in Kapitel 3 präsentierten analytischen Ausdrücken für die Steilheit und die Grenzfrequenzen ablesbar. Wie bereits in Kapitel 5 schematisch dargestellt, zerlegen sich Source- und Drainwiderstand in verschiedene Teilkomponenten, die getrennt voneinander betrachtet werden können (Abbildung 5.4). Nach Gleichung 6.1 setzt sich der Gesamtwiderstand R_s bzw. R_d aus dem Metallwiderstand R_M der ersten Kontaktierungsebene, dem Kontaktwiderstand R_C zwischen der ohmschen Metallisierung und dem hochdotierten n+-Halbleiter, dem Bahnwiderstand R_{Impl} des Implantationsgebietes und dem Widerstand des ungesteuerten Kanalgebietes R_{Si} zusammen.

$$R_{s/d} = R_M + R_C + R_{Impl} + R_{Si} \quad (6.1)$$

Bei der Betrachtung von Einzelbauelementen ist der Widerstandsanteil R_M bei angemessener Schichtdicke der Metallisierung gegenüber den restlichen Komponenten zu vernachlässigen. Da die Optimierung des Kanalwiderstandes bereits in Kapitel 4 ausführlich diskutiert wurde, stehen nun R_{Impl} und R_C im Mittelpunkt des Interesses. Der Bahnwiderstand des ohmschen Kontaktes R_{Impl} wird nach Gleichung 6.2 von dem Schichtwiderstand ρ_S des Implantationsgebietes bestimmt. Dabei berücksichtigt L_{MK} den Abstand zwischen dem Rand des Kontaktmetalles und dem Anfang des Kanalgebietes, und W_G stellt die Weite der Kontaktfläche dar.

$$R_{Impl} = \rho_S \cdot \frac{L_{MK}}{W_G} \quad (6.2)$$

Der Kontaktwiderstand R_C ergibt sich aus dem spezifischen Kontaktwiderstand ρ_C und der effektiven Fläche des Kontaktgebietes. Da die maximale Potentialdifferenz zwischen Metallkontakt und Halbleiter am Rand des Metalles auftritt und innerhalb des Metalles exponentiell abfällt, fließt der Strom nur über einen schmalen Randbereich zwischen den beiden Materialien. Bei eindimensionaler Betrachtung ist die Breite dieses Gebietes durch die Transferlänge L_T gegeben. Ist nun die Breite des Kontaktgebietes größer als die Transferlänge, so wird bei der Berechnung des Kontaktwiderstandes nur die effektive Fläche ($L_T \cdot W_G$) berücksichtigt, durch die der Strom fließt.

$$R_C = \frac{\rho_C}{L_C \cdot W_G} \quad \text{für } L_C \leq L_T \quad (6.3)$$

$$R_C = \frac{\rho_C}{L_T \cdot W_G} \quad \text{für } L_C > L_T \quad \text{mit } L_T = \sqrt{\frac{\rho_C}{\rho_S}} \quad (6.4)$$

Die folgenden Abschnitte beschäftigen sich mit der Reduzierung des Schichtwiderstandes für die Source/Drain-Gebiete und der Optimierung des Kontaktwiderstandes. Die Diskussion beginnt mit einer Variation der Implantationsparameter, um eine ideale Kontaktierung des Si-Kanals im heterogenen Schichtsystem zu gewährleisten. Im Anschluß daran erfolgt die Umsetzung der in Abschnitt 6.1 festgelegten thermischen Randbedingungen bei der Aktivierung des Dotierstoffes und der Rekristallisierung des Halbleiters nach der Implantation. Abschließend werden Verfahren und Möglichkeiten erarbeitet, um durch den Einsatz von Siliziden eine weitere Minderung der spezifischen Kontakteigenschaften zu erzielen.

6.2.1 Niederohmige Kontaktierung durch Ionenimplantation

In der Silizium-Technologie hat sich die Ionenimplantation zur Erzeugung niederohmiger Kontaktgebiete durchgesetzt. Darüber hinaus stellt sie bei der CMOS-Integration ein wirksames Mittel dar, um präzise abgestimmte, in zwei Dimensionen optimierte Dotierungsprofile zu realisieren, die den wachsenden Einfluß der Kurzkanaleffekte eindämmen sollen. Im Gegensatz dazu haben sich bei den III/V-Halbleitern für viele Anwendungen legierte ohmsche Kontakte behauptet, die der implantierten Version in der Si-Technologie qualitativ nicht nachstehen. Die Anfangszeit des SiGe-HFETs zeigte eine starke Prägung durch Ideen aus der III/V-Technologie, so daß beide Varianten untersucht und verglichen worden sind. Aufgrund von Schicht- und Kontaktwiderständen, die bis zu einer Größenordnung niedriger lagen, konnte sich die Ionenimplantation auch bei den SiGe-Halbleitern durchsetzen.

Die Anforderungen an die Source/Drain-Gebiete des Transistors sind eine widerstandsoptimierte, ohmsche Kontaktierung des Kanals und ein verlustarmer Übergang vom Implantationsgebiet zur metallischen Kontaktierungsebene. Um die Implantationsparameter möglichst exakt an die gegebene Schichtkonfiguration anpassen zu können, ist die Form des Dotierprofiles von großer Bedeutung. Die

LSS-Theorie [6.8] liefert mit Gleichung (6.5) eine erste Abschätzung für die Reichweitenverteilung des eingebrachten Dotierstoffes in Abhängigkeit vom Ort x und von der implantierten Dosis N_0 [6.9].

$$N(x) = \frac{N_0}{\sqrt{2 \cdot \pi} \cdot \Delta R_p} \cdot \exp\left[-\frac{(x - R_p)^2}{2 \cdot \Delta R_p^2}\right] \quad (6.5)$$

Die mittlere projizierte Reichweite R_p und die Standardabweichung ΔR_p sind für die gebräuchlichsten Halbleitermaterialien, Isolatoren und Photolacke in Tabellenwerken in Abhängigkeit der Implantationsenergie E aufgelistet [6.10], [6.11]. Die variierende heterogene Zusammensetzung des Schichtsystems beim SiGe-HFET erfordert eine Korrektur der Reichweitenparameter. Aus der Literatur sind jedoch keine Untersuchungen bekannt, die eine Änderung der Dotierstoffverteilung mit dem Ge-Anteil berücksichtigen. Deshalb sind SiGe-Schichten mit unterschiedlichen Ge-Konzentrationen bei gleicher Energie und Dosis mit Phosphor implantiert und anschließend mit SIMS analysiert worden [6.12].

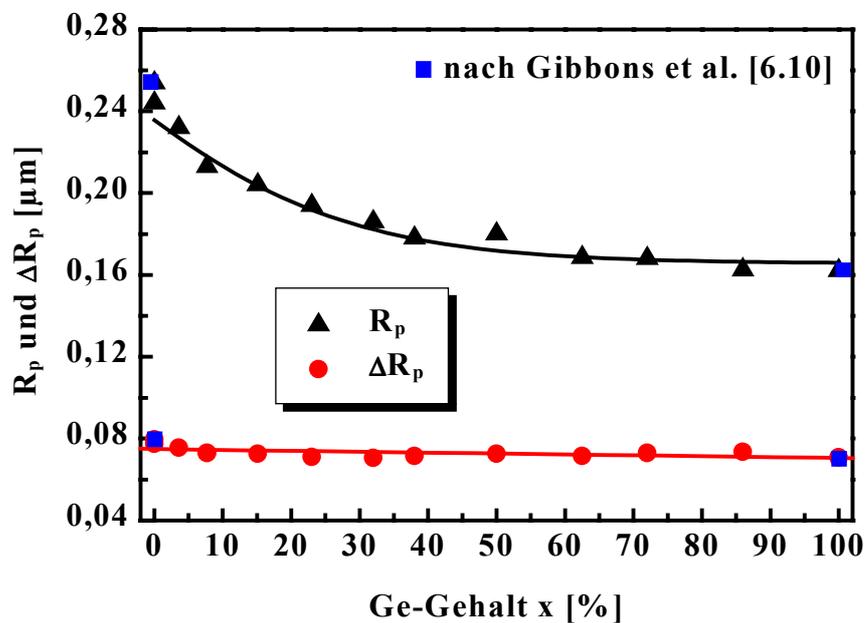


Abbildung 6.5: Experimentell ermittelte Werte für die Reichweitenparameter R_p und ΔR_p in Abhängigkeit des Ge-Gehaltes der untersuchten SiGe Schicht. Die Dotierung der Schichten erfolgte mit Phosphor bei einer Energie von 200 keV und einer Dosis von $1 \cdot 10^{15} \text{ cm}^{-2}$

Die MBE-gewachsenen Proben zeigen eine Variation des Ge-Gehaltes der 500 nm dicken SiGe-Schicht zwischen 0 und 86 %. Nach der Implantation mit einer Energie von 200 keV und einer Dosis von $1 \cdot 10^{15} \text{ cm}^{-2}$ erfolgt die Extraktion der Reichweitenparameter durch die Anpassung von Gleichung (6.5) an das gegebene SIMS-Profil. Die extrahierten Werte für R_p und ΔR_p sind in Abbildung 6.5 in Relation zum Ge-Gehalt der SiGe-Schicht dargestellt. Für zunehmende Ge-Anteile bis ca. 50 % fällt R_p stark ab und geht dann in eine lineare Abhängigkeit über. ΔR_p zeigt über den ganzen Untersuchungsbereich eine schwache Variation mit der Ge-Konzentration. Die zusätzlich aufgeführten Referenzwerte für 0 % und 100 % Germanium nach Gibbson et al. ergänzen die Kurvenverläufe in den Extrema und verifizieren eine Meßungenaugigkeit der SIMS-Analyse von unter 10 %.

Um die Übertragbarkeit der gewonnenen Reichweitenparameter auf reale Heteroschichtsysteme zu verifizieren, ist eine Transistorstruktur mit einem Grundgehalt von 35 % Ge einem typischen Implantationsprozeß für die Source/Drain-Gebiete mit 50 keV Energie und einer Phosphordosis von $5 \cdot 10^{15} \text{ cm}^{-2}$ ausgesetzt worden. Die Dimensionierung der Schichtdicken ist mit 10 nm Si-Cap, 50 nm $\text{Si}_{0.5}\text{Ge}_{0.5}$ Spacer und 55 nm Si-Kanal bewußt höher angesetzt, um eine exakte SIMS-Analyse im Rahmen der Meßungenaugigkeit gewährleisten zu können. Das extrahierte Dotierungsprofil (Abbildung 6.6) zeigt eine auffallende Singularität der Dotierstoffverteilung an den Heteroübergängen des Schichtsystems. Hervorgerufen durch elastische Kernstöße und unelastische Wechselwirkungen mit Elektronen an den Grenzflächen, ändert sich die Reichweitenverteilung des Dotierstoffes und damit die Konzentration auffallend. Deutlich ist auch eine Abhängigkeit von der Donatordichte in den Dotierschichten erkennbar. Die hochdotierte Versorgungsschicht auf der Vorderseite des Kanals bewirkt eine ausgeprägtere Singularität als die niedrigere Rückseitendotierung. Die Unstetigkeit beim Übergang vom Si-Cap zum SiGe-Spacer ist wenig auffällig, was auf das mangelnde Auflösungsvermögen der SIMS-Analyse in diesem Bereich zurückgeführt werden kann.

Aufgrund des linearen Zusammenhanges zwischen den Reichweitenparametern und der Implantationsenergie sind die oben dargestellten Werte für R_p und ΔR_p bei 200 keV auf jede beliebige Energie herunterskalierbar. Dadurch läßt sich das ermittelte SIMS-Profil mit dem simulierten Konzentrationsverlauf, unter Berücksichtigung der modifizierten Reichweitenparameter, für den jeweiligen Schichtabschnitt vergleichen. Abbildung 6.6 zeigt eine gute Übereinstimmung im prinzipiellen Verlauf des Profils, jedoch kann die inhomogene Dotierstoffverteilung

an den Grenzflächen nur ansatzweise wiedergegeben werden. Der Grund für diese Abweichungen liegt wieder an der hohen Dichte von Antimonatomen in den Dotierschichten, die durch die hier verwandte Theorie keine Berücksichtigung findet. Trotzdem erlaubt diese Methode die Abschätzung der Dotierstoffkonzentration für heterogene Systeme und bietet damit eine Möglichkeit, die Implantationsparameter an die jeweilige Schichtstruktur anzupassen.

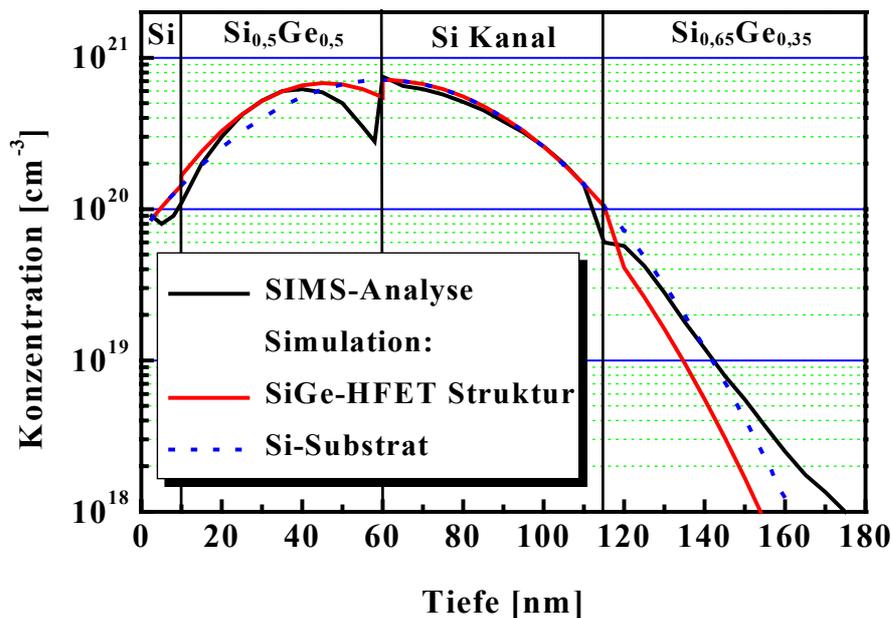


Abbildung 6.6: Vergleich zwischen dem experimentell ermittelten SIMS-Dotierungsprofil eines SiGe-HFETs und dem mit Hilfe der modifizierten Reichweitenparameter R_p und ΔR_p simulierten Verlauf. Als Referenz dient eine Simulation mit Standardparametern für Si-Substrat

Die im Rahmen dieser Arbeit zur Herstellung von SiGe-n-Kanal-HFETs verwendeten Schichtsysteme weisen alle einen Abstand von 14 - 18 nm zwischen Halbleiteroberfläche und Kanal auf. Unter Berücksichtigung eines Materialabtrages am Si-Cap von 1 - 2 nm durch Reinigungs- und Oxidationsschritte ergibt sich der Ladungsschwerpunkt des 2DEGs in einer Tiefe von maximal 25 nm. Simulationen mit den oben extrahierten, modifizierten Reichweitenparametern für SiGe-Schichten ergeben eine optimale Kontaktierung des Kanals für eine Standardimplantation mit Phosphor bei einer Energie von 15 - 20 keV.

Die Bestimmung der optimalen Implantationsdosis ist nicht nur unter der Zielvorgabe eines minimalen Schicht- und Kontaktwiderstandes, sondern auch in Hinblick auf das Temperaturbudget bei der Rekristallisierung ein wichtiger Aspekt. Die durch den Beschuß mit Ionen entstehenden Kristalldefekte bilden mit zunehmender Dosis isolierte amorphe Zonen im Halbleiter. Oberhalb einer bestimmten, materialabhängigen Dosis wachsen die amorphen Gebiete zusammen, so daß schließlich der gesamte oberflächennahe Bereich des Festkörpers eine amorphe Schicht bildet. Beim Überschreiten der amorphisierenden Dosis von etwa $8 \cdot 10^{14} \text{ cm}^{-2}$ kann eine vollständige Aktivierung des Dotierstoffes und die Rekristallisierung des Gitters bereits bei 500 - 600 °C erzielt werden. Dabei dient der ungestörte Kristall als Keim für den Rekristallisierungsprozeß, der sich unter gleichzeitigem Einbau des Dotierstoffes auf regulären Gitterplätzen bis an die Oberfläche fortpflanzt. Bei Hochdosisimplantationen über 10^{16} cm^{-2} kommt es zum Überschreiten der maximalen Löslichkeit des Dotierstoffes in Si und Si-Ge. Da für die implantierten Ionen nicht genügend Gitterplätze zur Verfügung stehen, werden sie erst nach einem Diffusionsprozeß regulär eingebaut. Deshalb sind Kurzzeittemperprozesse von 30 - 45 s oft nicht ausreichend, um eine 100 %ige Aktivierung zu erzielen.

Mit der Zielsetzung eines minimalen Schicht- und Kontaktwiderstandes bei möglichst niedrigem Temperaturbudget sind identische, modulationsdotierte Si-Ge-Heterostrukturen bei unterschiedlichen Implantationsbedingungen untersucht worden. In Abbildung 6.7 und Abbildung 6.8 sind die Ergebnisse für den Schicht- und Kontaktwiderstand in Abhängigkeit von der Aktivierungstemperatur im RTA-System dargestellt. Die Dauer der Temperaturbelastung beträgt jeweils 30 s und läuft unter Stickstoffatmosphäre ab. Die Ermittlung der Widerstände erfolgte mit Hilfe von TLM-Messungen an mindestens zehn unterschiedlichen Stellen auf der Probe. Für den Schichtwiderstand ρ_s sind zusätzlich Vergleichsmessungen mit der Hall-Methode gemacht worden (offene Symbole).

Sowohl die Ergebnisse der TLM- als auch die der Hall-Messung zeigen den niedrigsten Schichtwiderstand bei einer Implantationsenergie von 20 keV. Hinzu kommt bei beiden Implantationsdosen eine schwache Abhängigkeit von der Ausheiltemperatur. In Verbindung mit den Resultaten für den auf die Gateweite normierten, spezifischen Kontaktwiderstand ergeben sich für die Probe mit einer Dosis von $2 \cdot 10^{15} \text{ cm}^{-2}$ bei 20 keV im Mittel die niedrigsten Widerstände. Unter Berücksichtigung der in Abschnitt 6.1 gesteckten thermischen Randbedingungen muß die Temperaturbelastung für eine Heterostruktur mit einem Ge-Gehalt über 40 % unterhalb von 600 °C gehalten werden.

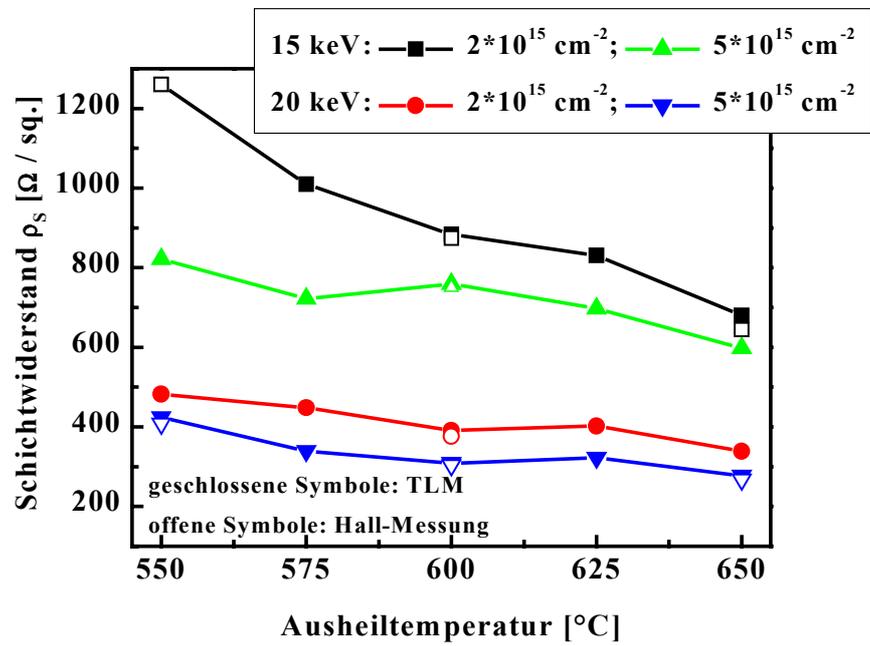


Abbildung 6.7: Schichtwiderstand in Abhängigkeit von der Ausheiltemperatur für verschiedene Implantationsparameter

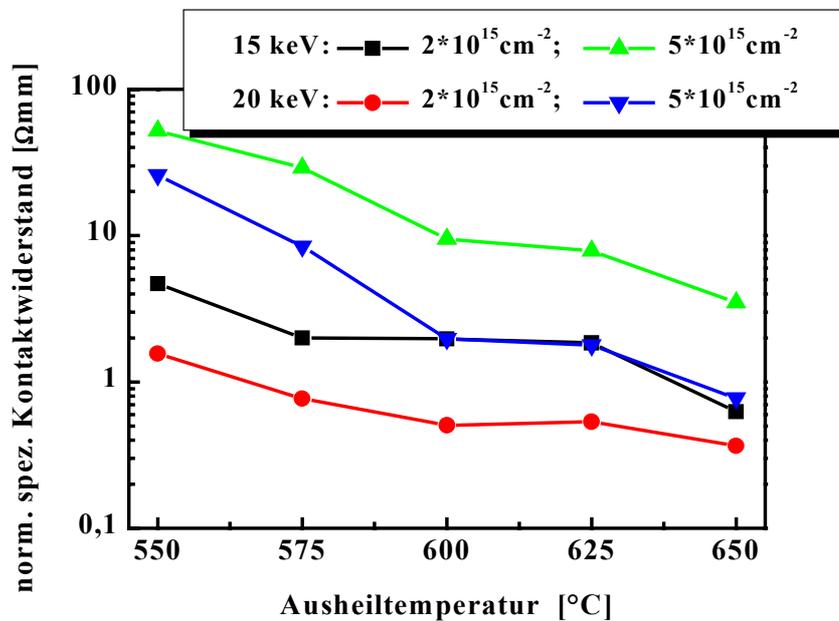


Abbildung 6.8: Normierter spezifischer Kontaktwiderstand in Abhängigkeit von der Ausheiltemperatur für verschiedene Implantationsparameter

In diesem Temperaturbereich erfolgt das Ausheilen des Gitters jedoch nicht mehr so spontan wie bei höheren Temperaturen. Anhand von 4-Zoll Siliziumwafern, die alle mit einer Phosphordosis von $2 \cdot 10^{15} \text{ cm}^{-2}$ bei 20 keV implantiert wurden, zeigt sich mit Hilfe eines Schichtwiderstandsmapping das Fortschreiten der Rekristallisationsfront (Abbildung 6.9). Während bei einer Ausheiltemperatur von $550 \text{ }^\circ\text{C}$ nach einer Dauer von 60 s am Rand des Wafers noch nicht ausgeheilte Bereiche existieren, reichen bei $575 \text{ }^\circ\text{C}$ schon 30 s aus, um den gesamten Wafer zu rekristallisieren. Die bei einer Aktivierungstemperatur von $575 \text{ }^\circ\text{C}$ erzielten Widerstandswerte für die reale SiGe-Heterostruktur liegen mit $\rho_s = 340 \text{ } \Omega/\square$ und $\rho_c = 1,3 \cdot 10^{-5} \text{ } \Omega \cdot \text{cm}^2$, im Vergleich zur CMOS- bzw. GaAs-Technologie, mindestens um eine Größenordnung höher. Als eine weitere Optimierungsoption bietet sich deshalb der Einsatz eines niederohmigen Silizides an.

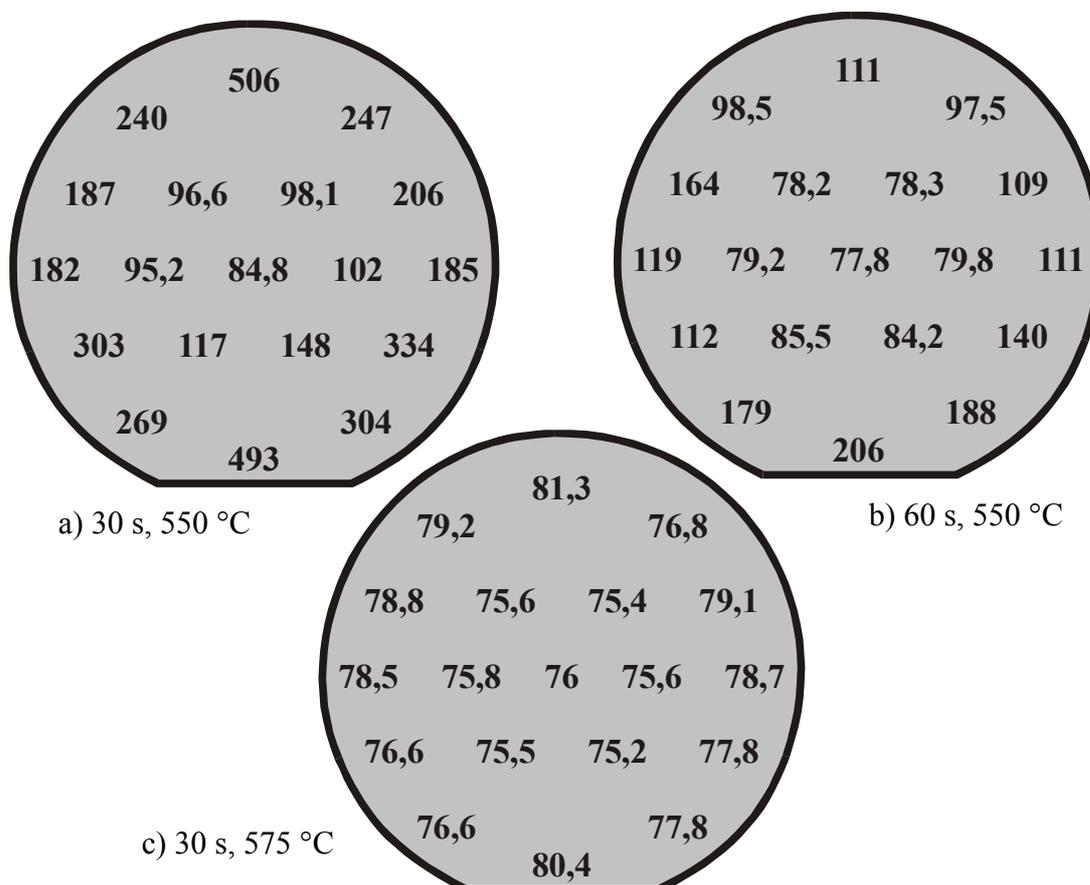


Abbildung 6.9: Schichtwiderstandsmessungen an phosphorimplantierten 4-Zoll p⁻-Siliziumwafern in Abhängigkeit der Ausheilzeit und -temperatur a) 30 s, 550 °C; b) 60 s, 550 °C; c) 30 s, 575 °C

6.2.2 Optimierung der Kontakte durch Silizidierung

Der Einsatz von Siliziden und Polyziden ist ein Standardverfahren in der CMOS-Technologie, um niederohmige Kontakt- und Schichtwiderstände herzustellen [6.13], [6.14]. Nahezu alle Metalle des Periodensystems der Elemente bilden bei der Reaktion mit Silizium eine silizidartige Verbindung mit metallischen Leitungseigenschaften. Für die Anwendung in mikroelektronischen Transistorstrukturen spielt darüber hinaus die Selektivität des Reaktionsprozesses zwischen der Halbleiteroberfläche und dem isolierenden SiO_2 eine entscheidende Rolle. Eine hohe Reaktionsselektivität ermöglicht eine selbstjustierende Silizidierung der Kontaktgebiete mit Hilfe einer vordefinierten SiO_2 -Maske. Unter Einsatz einer selektiven Ätzmischung kann anschließend das nicht silizidierte Metall auf der Maske entfernt werden, um elektrische Kurzschlüsse zu vermeiden.

Im Rahmen der fortschreitenden Miniaturisierung in der Mikroelektronik spielt dieses selbstjustierende Verfahren eine entscheidende Rolle, um den Einfluß der parasitären Widerstände auf das Bauelement- und Schaltungsverhalten zu minimieren. In der CMOS-Technologie hat sich TiSi_2 im Laufe der Zeit eine dominierende Stellung aufgebaut [6.15, 6.13, 6.16, 6.17]. Aufgrund von bestehenden Materialerfahrungen mit Ti und TiN bei der Mehrlagenmetallisierung war der Einsatz als Silizid naheliegend. Des weiteren überzeugt TiSi_2 durch einen niedrigen spezifischen Widerstand, gute Temperaturstabilität und eine hohe Reproduzierbarkeit in der Prozeßführung. Jedoch wirft die progressive Skalierung in der CMOS-Technologie zunehmend Probleme bei der Verwendung von TiSi_2 auf, da der Silizidierungsprozeß nicht nur von der Dotierstoffkonzentration der Implantationsgebiete, sondern auch von der Siliziddicke abhängig ist [6.18]. Bei dünnen Silizidschichten tritt ein Verzögerungseffekt bei der Umwandlung von C49- TiSi_2 in das niederohmige C54- TiSi_2 auf, was zum Ansteigen der Reaktionstemperatur führt. Dadurch ergibt sich für den Einsatz von TiSi_2 bei fortschreitender Skalierung der ohmschen Kontaktgebiete eine zunehmende Einschnürung des Prozeßfensters [6.19].

Als alternative Silizide haben CoSi_2 und NiSi in den letzten Jahren an Aufmerksamkeit gewonnen [6.20], [6.21]. Im Gegensatz zu TiSi_2 ist die Reaktionstemperatur, die zur Bildung der niederohmigen Verbindung führt, bei beiden Materialkombinationen unabhängig von der Siliziddicke. Außerdem hat die Art und Höhe der Siliziumdotierung keinen Einfluß auf den Schichtwiderstand des Silizides [6.19]. Dadurch weisen CoSi_2 und NiSi eine gute Skalierbarkeit auf, bei einem vergleichbar niedrigen spezifischen Widerstand wie er bei TiSi_2 auftritt. Im Pro-

zeßablauf zeigen sie eine hohe Reproduzierbarkeit, und die Resistenz gegenüber HF-haltigen Säuren und trockenchemischen Ätzprozessen erleichtert die Integration in bestehende Standardtechnologien.

| Eigenschaften | TiSi ₂ | CoSi ₂ | NiSi |
|---|---|--|---|
| spez. Widerstand [$\mu\Omega\cdot\text{cm}$] | 13-26 | 13-25 | 10-16 |
| Schottky-Barriere $q\cdot\Phi_B$ [eV] | 0,60 | 0,64 | 0,62 |
| Gitterstruktur | orthorhombisch | kubisch | orthorhombisch |
| Gitterfehlpassung zu Si | 1,8 % | 1,2 % | 0,4 % |
| selektive Ätzlösung zum Ursprungsmaterial | H ₂ SO ₄ :H ₂ O ₂ | HCl:H ₂ O ₂ , H ₂ SO ₄ :H ₂ O ₂ | H ₂ SO ₄ :H ₂ O ₂ |
| Si-Verbrauch bei 1 nm Ursprungsmaterial [nm] | 2,22 | 3,61 | 1,83 |
| Resistenz gegen Säuren | wird von HF-haltigen Säuren angegriffen | verdünnte HF und BHF haben keinen Einfluß auf den Schichtwiderstand | |
| Resistenz gegen Plasmaätzprozesse | Materialabtrag durch Ätzprozesse | keine Einwirkungen auf das Silizid nachweisbar | |

Tabelle 6.1: Gegenüberstellung der wichtigsten charakteristischen Eigenschaften von TiSi₂, CoSi₂ und NiSi [6.21, 22, 23, 24, 25, 26]

In Tabelle 6.1 sind weitere wichtige charakteristische Eigenschaften von TiSi₂, CoSi₂ und NiSi gegenübergestellt. Ein entscheidender Nachteil des CoSi₂ besteht in dem hohen Siliziumverbrauch bei der Silizidbildung. Um 1 nm des Ursprungsmaterials Kobalt in Silizid umzuwandeln, werden 3,61 nm Silizium benötigt. Damit liegt der Siliziumkonsum 63 % höher als beim TiSi₂ und 97 % über dem des NiSi. Da der resultierende Schichtwiderstand mit steigender Dicke des Ursprungsmaterials abnimmt, limitiert der hohe Siliziumverbrauch die Realisierung eines niederohmigen Silizides insbesondere bei flachen Implantationszonen. Ein wichtiges Kriterium für die Anwendbarkeit eines Silizides ist die Lage und Breite

des thermischen Prozeßfensters. Abbildung 6.10 zeigt für Ti, Co und Ni die Ausbildung des niederohmigen Silizides in Abhängigkeit von der Prozeßtemperatur.

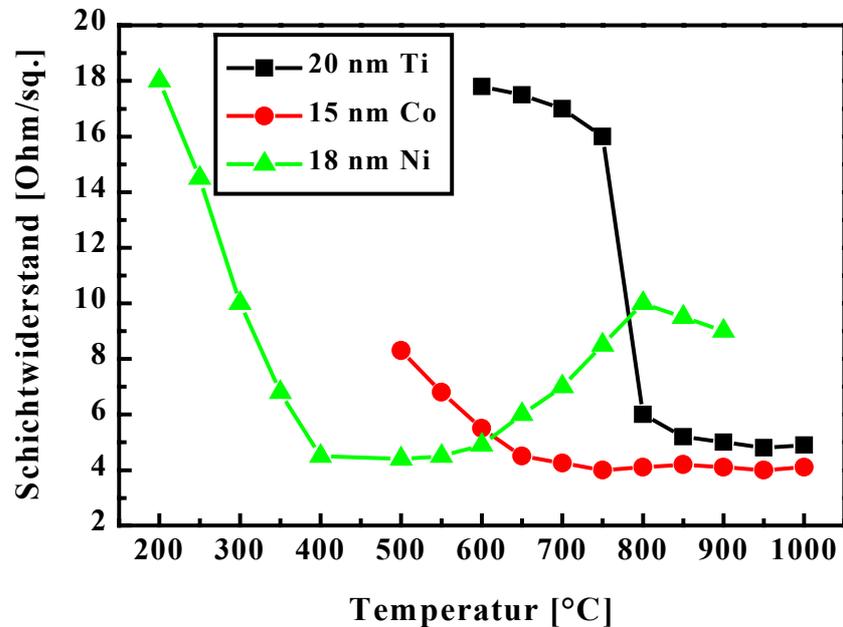


Abbildung 6.10: Schichtwiderstand von TiSi_2 , CoSi_2 und NiSi in Abhängigkeit von der Temperatur

Der zweistufige Prozeß zur Herstellung von TiSi_2 beginnt mit einer Temperung bei 650 °C für 30 s. Dabei reagiert das Ti mit dem Silizium und bildet eine hochohmige C49- TiSi_2 Phase. Gleichzeitig entsteht in der Stickstoffatmosphäre des RTA-Systems TiN. Durch einen naßchemischen Ätzschritt in $\text{H}_2\text{SO}_4:\text{H}_2\text{O}_2$ (1:1) wird sowohl das TiN als auch überschüssiges Ti von der Oberfläche entfernt. Der Übergang des C49- TiSi_2 in die niederohmige C54- TiSi_2 Phase erfolgt in einem zweiten Tempersschritt bei Temperaturen über 800 °C [6.19]. Mit einer Ausgangsdicke von 20 nm Ti lassen sich dadurch minimale Schichtwiderstände von $4,8\ \Omega/\square$ erzielen. Das Auftreten von TiSi_2 -Agglomerationen und der damit verbundene Anstieg des Widerstandes konnte bis zu einer Temperatur von 1000 °C nicht beobachtet werden.

Der ebenfalls zweistufige CoSi_2 -Prozeß zeigt nach dem ersten Temperingschritt bei 450 °C und 30 s eine Durchmischung der beiden Ausgangsmaterialien und die Bildung von CoSi und Co_2Si , die beide einen schlechteren spezifischen Widerstand aufweisen als das endgültige CoSi_2 [6.27]. Im anschließenden Reini-

gungsschritt in $\text{H}_2\text{SO}_4:\text{H}_2\text{O}_2$ (1:1) wird das unverbrauchte Co entfernt, ohne dabei Einfluß auf den Schichtwiderstand zu nehmen. Der zweite RTA-Schritt führt bei Temperaturen über $600\text{ }^\circ\text{C}$ zu einem von Nukleation und Diffusion bestimmten Übergang in die CoSi_2 -Phase. Daraus resultiert eine Minderung des Schichtwiderstandes auf Werte um $4\ \Omega/\square$.

Nickel reagiert mit kristallinem Silizium bereits bei Temperaturen über $200\text{ }^\circ\text{C}$ und bildet dabei Ni_2Si . Unter Einsatz eines gesteigerten Siliziumverbrauches erfolgt bei $400\text{ }^\circ\text{C}$ der Übergang in die niederohmige NiSi -Phase [6.28]. Im Rahmen dieses Prozesses lassen sich mit einer ursprünglichen Metalldicke von 18 nm Schichtwiderstände von minimal $4,4\ \Omega/\square$, unabhängig von der Dotierung des Substrates, generieren. Oberhalb von $650\text{ }^\circ\text{C}$ kann ein deutlicher Anstieg des Widerstandes beobachtet werden, der auf einen weiteren Phasenwechsel schließen läßt. In diesem Bereich wandelt sich das NiSi in ein hochohmiges NiSi_2 um. Gleichzeitig konnten Materialagglomerationen nachgewiesen werden, die ebenfalls einen widerstandserhöhenden Einfluß haben.

Die Anwendbarkeit der oben gegenübergestellten Silizide zur Herstellung von modulationsdotierten SiGe -HFETs muß neben der Skalierbarkeit auch dem Kriterium eines reduzierten Temperaturbudgets standhalten. Während das CoSi_2 aufgrund seiner guten Eigenschaften bei reduzierten Strukturgrößen als zukunftssträchtige Alternative zum TiSi_2 in der CMOS-Technologie eingestuft werden kann, erfüllt NiSi alle wichtigen Voraussetzungen zur Realisierung von niederohmigen Kontakten unter Einsatz einer Niedertemperaturtechnologie. Insbesondere das ermittelte optimale Prozeßfenster zwischen $400\text{ }^\circ\text{C}$ und $650\text{ }^\circ\text{C}$ stimmt sehr gut mit der maximalen Temperaturbelastung von $575\text{ }^\circ\text{C}$ für das Herstellungsverfahren überein.

Um die Verbesserung der Kontakteigenschaften durch den Einsatz von NiSi an realen SiGe -Schichtsystemen quantitativ zu belegen, sind die modulationsdotierten Heterostrukturen aus Abbildung 6.7 und Abbildung 6.8 zusätzlich einem Silizidprozeß unterzogen worden. Dem Experiment lagen wieder die vier Parametervariationen bei der Implantation zugrunde, jedoch erfolgte die thermische Aktivierung diesmal nur bei der optimalen Temperatur von $575\text{ }^\circ\text{C}$. Die Untersuchung berücksichtigt zwei unterschiedliche Ni-Dicken von 10 nm und 20 nm , deren Silizidierung bei $450\text{ }^\circ\text{C}$ für 30 s in einem RTA-System abläuft. Das überschüssige Ni wird anschließend durch eine naßchemische Behandlung in $\text{H}_2\text{SO}_4:\text{H}_2\text{O}_2$ (1:1) von der silizidierten Oberfläche und der Oxidmaske entfernt.

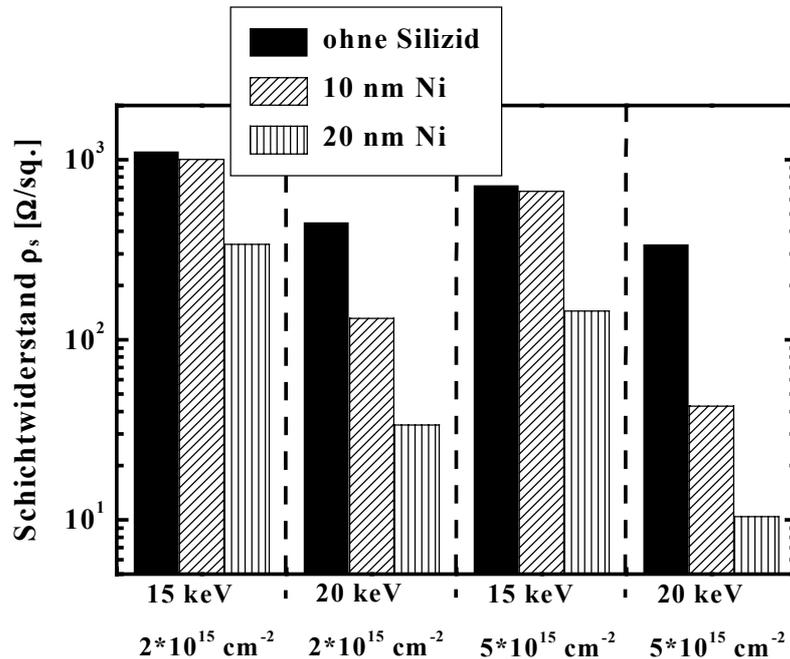


Abbildung 6.11: Schichtwiderstand ohne und mit Silizidierung (10 nm und 20 nm Ni) für verschiedene Implantationsparameter

Die Diagramme in Abbildung 6.11 und Abbildung 6.12 zeigen den Schichtwiderstand und den normierten spezifischen Kontaktwiderstand der mit TLM-Messungen charakterisierten Proben. Die Analyse des Schichtwiderstandes weist wie erwartet eine Abnahme von ρ_s mit zunehmender Ni-Dicke auf. Im Gegensatz zu Untersuchungen auf Si-Substraten, die keine Abhängigkeit von der Dotierung ergaben, ist der Einfluß unterschiedlicher Implantationsparameter bei den SiGe-Heterostrukturen offensichtlich. Insbesondere bei den flach implantierten Proben (15 keV) tritt eine signifikante Reduzierung erst bei einer Nickeldicke von 20 nm auf. Obwohl aus der wissenschaftlichen Literatur keine kristallografischen Analysen über die Ni-Silizidbildung im Si/SiGe-Heterosystem bekannt sind, lassen sich auf der Basis belegter Fakten folgende Erklärungen dafür finden.

Es ist nachgewiesen, daß Ni in Verbindung mit SiGe bei Temperaturen über 300 °C Ni_2SiGe bildet. Dieser Prozeß wird von der Ni-Diffusion bestimmt und zeichnet sich durch einen hohen Nickelverbrauch aus. Oberhalb von 400 °C kommt es zu einer Aufspaltung in eine ternäre NiSiGe-Verbindung und in das niederohmige NiSi [6.29]. Dazu muß mittels Diffusion weiter reines Ni nachgeführt werden. Da der Widerstand der NiSiGe-Komponente deutlich höher ist als

der des reinen NiSi, wächst damit auch der gesamte Schichtwiderstand [6.30]. Der durch den Ge-Gehalt forcierte Einbau von Nickel kann deshalb bei zu dünnen Metallschichten nicht vollständig gesättigt werden, so daß sich der Schichtwiderstand nur in kleinerem Ausmaß verringert. Eine Nickeldicke von 20 nm kommt dem Materialkonsum besser nach, so daß die Reduzierung des Widerstandes signifikanter ausfällt.

Zusätzlich ist zu erwarten, daß sowohl die inhomogene Dotierstoffverteilung an den Übergängen zwischen Si und SiGe als auch die Grenzflächen selber eine hemmende Wirkung auf die Silizidbildung ausüben. Durch eine flachere Implantation rückt das Dotierstoffmaximum näher an die Grenzschicht zwischen SiGe-Spacer und Kanal heran und verstärkt diesen Effekt. Trotzdem führt die Silizidierung von 20 nm Ni für alle Implantationskonfigurationen zu einer drastischen Reduzierung des Schichtwiderstandes von mindestens 70 %. Im optimalen Fall (20 keV, $5 \cdot 10^{15} \text{ cm}^{-2}$) wird sogar eine Abnahme um 90 %, von $339 \Omega/\square$ auf $10 \Omega/\square$, erzielt.

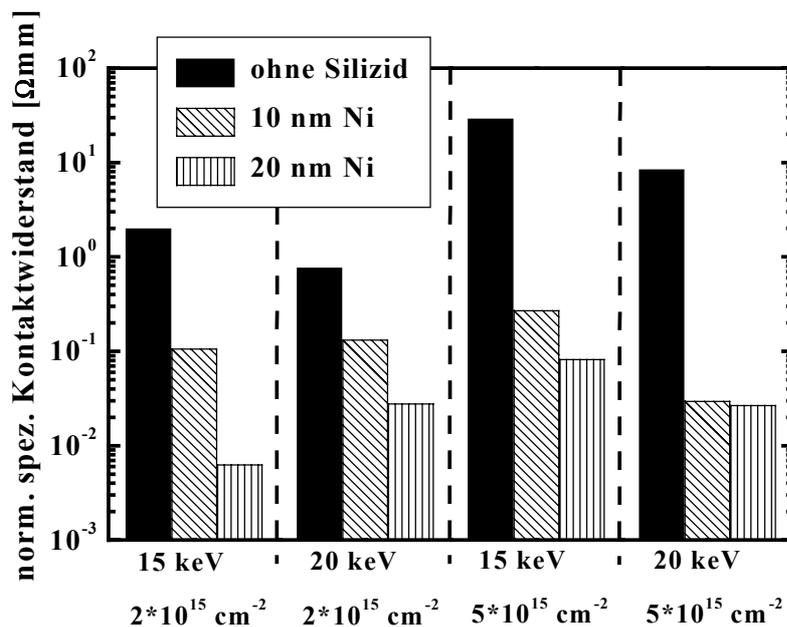


Abbildung 6.12: Normierter spezifischer Kontaktwiderstand ohne und mit Ni-Silizidierung für verschiedene Implantationsparameter

Die erzielten Kontaktwiderstände weisen nur eine schwache Abhängigkeit von den Implantationsparametern auf. Jedoch führt auch hier ein größeres Angebot an

aufgebrachtem Metall zu einem kleineren Widerstand. Durchschnittlich sinken die Kontaktwiderstände um etwa zwei Größenordnungen durch eine Ni-Silizidierung und liegen damit in einem Bereich, der in der CMOS- und III/V-Technologie als Standard gilt. Zwar ergibt sich der kleinste spezifische Kontaktwiderstand bei einer Implantationsenergie von 15 keV und einer Dosis von $2 \cdot 10^{15} \text{ cm}^{-2}$, allerdings ist für die Reduzierung des Gesamtwiderstandes sowohl ein niedriger Schichtwiderstand als auch ein geringer Kontaktwiderstand erforderlich. Eine kombinierte Betrachtung der beiden Kriterien favorisiert eine größere Implantationsenergie von 20 keV bei gleichzeitig erhöhter Dotierstoffdosis.

Die im Rahmen dieser Arbeit entwickelten Integrationskonzepte bedienen sich unterschiedlicher Abfolgen bei der Durchführung der Silizidierung. Im Fall des konventionellen Herstellungsablaufes besteht zum einen die Möglichkeit, das Nickel direkt nach der Implantation aufzubringen und nach dem Lift-off der Implantationsmaske die Aktivierung des Dotierstoffes und den Silizidprozeß in einem Temperaturschritt durchzuführen. Die zweite Variante nutzt eine Hilfsschicht aus SiO_2 , um diese Schritte zu entkoppeln. Die nach der Feldoxidstrukturierung aufgebraachte SiO_2 -Schicht wird mit Hilfe der Implantationsmaske naßchemisch strukturiert, um Fenster für den Dotierungsprozeß auf der Mesa zu öffnen. Nach der Implantation und dem Entfernen der Fotolackmaske findet zuerst die Aktivierung des Dotierstoffes statt, bevor die Abscheidung des Nickels für die Silizidierung erfolgt. Die Umwandlung des Metalls in ein niederohmiges Silizid läuft anschließend in einem separaten RTA-Schritt ab. Beim selbstjustierenden Integrationskonzept ist die Reihenfolge genau umgekehrt. Hier wird nach der Strukturierung des Dummy-Gates zuerst das Nickel aufgedampft und anschließend silizidiert und danach kommt es zur Implantation der ohmschen Kontakte und zur Rekristallisierung des Gitters.

Um die Frage zu klären, ob diese unterschiedlichen Versionen der Silizidierung qualitativ und quantitativ gleichwertig sind, wurden die verschiedenen Prozeßabläufe an realen modulationsdotierten SiGe-Heterostrukturen untersucht. Auf der Basis des konventionellen Integrationskonzeptes erfolgte die Herstellung realer Transistorstrukturen jedoch ohne Gate. Für die Drain/Source-Spannung V_{DS} des HFETs gilt bei in Flußrichtung betriebener Schottky-Diode [6.31]:

$$V_{\text{DS}} = (R_{\text{Kanal}} + R_{\text{s}} + R_{\text{d}}) \cdot I_{\text{D}} + (\alpha \cdot R_{\text{Kanal}} + R_{\text{s}}) \cdot I_{\text{G}} \quad (6.6)$$

Für Transistoren ohne Gateelektrode ($I_G = 0$) fällt der letzte Term weg, so daß der Anlaufwiderstand im linearen Bereich der Ausgangskennlinie direkt die Summe aus Kanalwiderstand R_{Kanal} und den Widerständen R_s und R_d angibt. Die nach Abzug von R_{Kanal} extrahierten Werte für R_s und R_d sind in Tabelle 6.2 für die verschiedenen Silizidierungsvarianten dargestellt.

| Widerstand R_s/R_d [$\Omega \cdot \text{mm}$] | einstufiger Prozeß: Impl./Silizid/ Temperung | zweistufiger Prozeß: Impl./Temperung/ Silizid/Temperung | zweistufiger Prozeß: Silizid/Temperung/ Impl./Temperung |
|---|--|---|---|
| $W_G = 20 \mu\text{m}$ | 0,30 4,13 (Schottky) | 0,28 | 0,32 |
| $W_G = 50 \mu\text{m}$ | 0,34 4,85 (Schottky) | 0,37 | 0,35 |

Tabelle 6.2: Aus dem Anlaufwiderstand der Ausgangskennlinie extrahierte Kontaktwiderstände der unterschiedlichen Prozeßvarianten

Ein Vergleich der ermittelten Gesamtwiderstände für R_s bzw. R_d zeigt deutlich, daß für die zweistufigen Technologieversionen keine Abhängigkeit von der gewählten Prozeßabfolge besteht. Die für verschiedene Gateweiten bestimmten Werte liegen alle innerhalb eines schmalen Intervalls, was zusätzlich das gute Skalieren der Widerstände mit den geometrischen Abmessungen demonstriert. Beim einstufigen Prozeß hingegen treten bei etwa 75 % der vermessenen Transistorstrukturen überproportional hohe Widerstände auf, die mehr als zehnmal größer sind als die typischen Werte. Der Stromverlauf im Anlaufgebiet der Ausgangskennlinie ist dabei nicht linear, sondern zeigt einen parabolischen Anstieg. Außerdem liegt der Sättigungsstrom ca. 40 % unter dem Maximum der zweistufigen Prozesse. Dieses Verhalten deutet auf eine Kontaktierung des Kanals durch einen schottkyartigen Übergang hin. Aufgrund des einstufigen Temperaturschrittes müssen die Aktivierung des Dotierstoffes und die Silizidbildung parallel ablaufen. Dabei startet die Rekristallisierung vom ungestörten Halbleiter ausgehend und läuft in Richtung Siliziumoberfläche. Der Silizidierungsprozeß hingegen beginnt an der Grenzfläche zwischen Halbleiteroberfläche und Nickel und dringt dann weiter in das amorphisierte Silizium ein. Im Gegensatz zur Silizidierung eines bereits rekristallisierten Halbleiters kann das Nickel hier tiefer in das ungeordnete Material hineinlaufen. Dadurch besteht die Möglichkeit, daß die

Nickelsilizidgrenze den Implantationsbereich lateral und vertikal überschreitet und einen Schottky-Kontakt bildet. Aufgrund der hohen Ausfallrate der einstufigen Silizidierungsvariante ist sie für eine stabile und reproduzierbare Prozeßführung ungeeignet. Im Fall des konventionellen Integrationskonzeptes muß deshalb auf die zweistufige Version unter Einsatz einer SiO₂-Hilfsschicht zurückgegriffen werden.

Zusammenfassend läßt sich festhalten, daß es durch die laterale Optimierung der SiGe-HFET-Struktur unweigerlich auch zu einer Skalierung der Implantationszonen kommt. Unter der Randbedingung, das Temperaturbudget im Herstellungsablauf unterhalb von 600 °C zu halten, ergeben sich dadurch hohe Anforderungen bei der Dimensionierung und Realisierung der ohmschen Kontaktgebiete. Durch eine gezielte Anpassung der Reichweitenparameter an die spezifischen Eigenschaften der SiGe-Heterostruktur konnten die Implantationsparameter optimal auf die Bauelementkonfiguration abgestimmt werden. Dabei ergab sich bei Implantationsenergien von 20 keV die effektivste Kontaktierung des Si-Kanals. Um die Widerstandscharakteristik der Source/Drain-Gebiete weiter zu verbessern, kam ein Silizidprozeß zum Einsatz. Dabei demonstrierte Nickel aus elektrischen, thermischen und prozeßbedingten Gründen ein ideales Zusammenspiel mit dem SiGe-System. Durch die Silizidierung von 20 nm Ni gelang es, den Bahnwiderstand des Dotiergebietes auf Werte um 10 Ω/□ zu senken. Gleichzeitig verringerte sich der normierte spezifische Kontaktwiderstand signifikant auf unter 0,01 Ω·mm. Für die Integrationsfähigkeit des optimierten Prozeßmoduls wurde abschließend nachgewiesen, daß nur eine Entkopplung von Implantationsaktivierung und Silizidierung die gewünschten niedrigen Source- und Drainwiderstände erzielen kann.

6.3 Strukturierung des Dummy-Gates

Eine minimale Gatelänge und ein möglichst niedriger Gatewiderstand sind die essentiellen Voraussetzungen für einen leistungsfähigen Hochfrequenztransistor. Da die Skalierung von L_G aber auch Einfluß auf den Gatequerschnitt hat, erhöht sich dadurch wiederum der Widerstand. Dieser gegenläufige Prozeß macht den Einsatz eines Metallgates bei kleinen Gatelängen unumgänglich. Die Gatekonstruktion des SiGe-HFETs basiert auf einer Au- oder Al-Elektrode mit einem Pt/Si-Übergang als Schottky-Diode. Das im Rahmen des Herstellungsprozesses durchlaufene Temperaturbudget würde unweigerlich zu einer unkontrollierten

Silizidierung der Gategrenzfläche führen und dadurch die Transistoreigenschaften drastisch verschlechtern oder sogar eine Zerstörung des Bauelementes zur Folge haben. Deshalb kann die Realisierung der metallischen Gateelektrode erst nach der Aktivierung des Dotierstoffes stattfinden. Bei der lateralen Optimierung der Transistorstruktur ist jedoch die selbstjustierende Anordnung der Source/Drain-Gebiete relativ zum Gate die wesentliche Maßnahme, den Einfluß parasitärer Elemente zu minimieren. Deshalb ist die Dummy- oder Replacement-Gatekonstruktion die zentrale Komponente im selbstjustierenden Integrationskonzept des SiGe-HFETs. Die Realisierung des Dummy-Gates durchläuft dabei zwei wichtige Phasen.

1. optische Definition der Anfangsgatelänge und trockenchemische Strukturierung der Dummy-Struktur
2. naß- und trockenchemische Skalierung des Gatesteiges

6.3.1 Definition und trockenchemische Ätzung des Dummy-Gates

Die Definition und Realisierung der Dummy-Gate-Struktur bildet die essentielle Grundlage für die selbstjustierende Anordnung der Source/Drain-Gebiete. Die vom Maskenlayout vorgegebene Breite des Gatesteiges legt den primären Abstand der ohmschen Kontakte fest. Sowohl der nachfolgende Silizidierungsprozeß als auch die Implantation nutzen die maskierende Wirkung des Dummy-Gates, um die Distanz zwischen der späteren Gateelektrode und den Source/Drain-Bereichen zu minimieren. Nach der isotropen Schrumpfung der Dummy-Struktur auf die endgültige Gatelänge und die anschließende Passivierung muß der in SiO₂ eingebettete Gatestieg wieder entfernt werden. Dabei überträgt sich sein Profil in Gestalt eines Grabens in die Isolationsschicht. Die Kontur des Grabens bestimmt bei der nachfolgenden Auffüllung mit Metall die Länge und Form des Gatefusses. Die daraus erwachsenen Anforderungen an den Strukturierungsprozeß sind eine maßhaltige Übertragung der vorgegebenen Abmessungen und ein ideal vertikales Profil des Gatesteiges. Weiterhin sollte sich der Herstellungsschritt durch eine gute Homogenität über die gesamte Waferfläche, eine hohe Reproduzierbarkeit und eine ausreichende Selektivität zum Substrat auszeichnen. Nur durch den Einsatz von trockenchemischen Ätzverfahren sind alle diese Anforderungen zu erfüllen. Obwohl naßchemische Strukturierungsmethoden sich durch eine größere Selektivität auszeichnen, sind der hohe Grad an Anisotropie und die flexible Einstellung des gewünschten Ätzprofils über die Prozeßparameter die entscheidenden Vorteile der Trockenätzverfahren.

Durch die gezielte Auswahl des Materials für die Dummy-Gate-Struktur kann schon im vorhinein entscheidend auf die dargestellten Anforderungen Einfluß genommen werden. Das Material sollte einfach und mit homogener Dicke aufzubringen sein und sich in einem stabilen trockenchemischen Ätzprozeß einfach strukturieren lassen. Aufgrund der Einbettung in eine SiO₂-Schicht muß die Temperaturstabilität ausreichend für den eingesetzten PECVD-Schritt sein.

| Eigenschaften | Poly-Si | Si ₃ N ₄ | WSi | PI | Fotolack |
|--|---|---|---|---|---|
| Haftung / Temperaturstabilität | ↑ / ↑ | ↑ / ↑ | ↑ / ↑ | ↑ / ↔ | ↑ / ↓ |
| trockenchemische Ätzung | SF ₆ , CF ₄ | CF ₄ /CHF ₃ | CF ₄ | O ₂ | O ₂ |
| Selektivität zu Si SiO ₂ | 2:1 14:1, 2:1 | 2:1 1:1 | 1:1 2:1 | > 100:1 > 100:1 | > 100:1 > 100:1 |
| naßchemische Ätzung | HNO ₃ /HF | H ₃ PO ₄ | HNO ₃ / NH ₄ F | H ₂ SO ₄ / H ₂ O ₂ | H ₂ SO ₄ / H ₂ O ₂ |
| Selektivität zu Si SiO ₂ | 1:1 0,1:1 | > 100:1 15:1 | 2:1 0,1:1 | 50:1 > 100:1 | 50:1 > 100:1 |
| Bewertung | schlechte Selekt. zu SiO ₂ => modifizierter Prozeß | geringe Selektivität zu Si und SiO ₂ | schlechte Selektivität zu Si und SiO ₂ | gut geeignet | geringe Temperaturstab. bei SiO ₂ Abscheidung |

Tabelle 6.3: Eigenschaften verschiedener Werkstoffe für den Einsatz als Dummy-Gate-Material (↑ = hoch, ↓ = niedrig)

Bei der Entfernung des Dummy-Gates sollte der verwendete Ätzprozeß eine genügend hohe Selektivität zum Oxid aufweisen. Außerdem ist dabei eine minimale Schädigung des Substrates von essentieller Wichtigkeit. Bei der Realisierung eines Hetero-MOSFETs besteht die Möglichkeit, das Gateoxid bereits vor

dem Dummy-Gate aufzubringen. In diesem Fall muß die Ätzung mit erhöhter Selektivität und ohne negativen Einfluß auf das Gateoxid erfolgen. Unter diesen Bedingungen ist der Einsatz einer komplett naßchemischen Ätzung sinnvoll, da sie für eine schädigungsarme Entfernung des Gates sorgt. Aber auch eine Kombination der beiden Verfahren ist denkbar. Als Material für das Dummy-Gate wurden verschiedene Standardwerkstoffe getestet (Tabelle 6.3). Dabei empfiehlt sich, unter Berücksichtigung der oben diskutierten Kriterien und unter den gegebenen technologischen Möglichkeiten, die Anwendung von Polyimid (PI).

Alternativ ist der Einsatz von Polysilizium als Replacement-Gate möglich. Da seine naßchemische Ätzung jedoch keine Selektivität zu SiO_2 aufweist, bietet sich hier ein modifizierter Prozeßablauf an. Dabei wird entweder das Gateoxid erst nach der Entfernung der Dummy-Gate-Struktur erzeugt oder durch das Aufbringen eines high- k Materials ersetzt.

Für das anisotrope Ätzen organischer Polymerschichten wird in der Halbleitertechnologie standardmäßig das reaktive Ionenätzen (RIE) mit Sauerstoff als Reaktionsgas eingesetzt [6.32], [6.33]. Aufgrund der hohen Anforderungen an die Flankensteilheit des Polyimidsteges muß die Ätzmaske bestimmte Voraussetzungen erfüllen. Da sich bei gegebener Selektivität $S = r_{\text{Maske}}/r_{\text{PI}}$ der Anstellwinkel der Maske α_{Maske} über die Beziehung $\tan \alpha_{\text{Maske}} = S \cdot \tan \alpha_{\text{PI}}$ auf das zu ätzende Polyimid überträgt, kann nur durch eine hohe Selektivität zwischen Maske und Polyimid oder durch einen Anstellwinkel $\alpha_{\text{Maske}} = 90^\circ$ das geforderte Profil erzeugt werden [6.34]. Aus diesen Gründen scheidet der Einsatz einer konventionellen Lacktechnologie zur anisotropen Strukturierung des Polyimides aus. Bereits während des Lithografieprozesses kommt es durch Temperaturschritte zum Verlaufen des Lackprofils und damit zur Verringerung der Flankensteilheit. Zusätzlich ist die Lackoberfläche während der Trockenätzung ständig dem Angriff der hochenergetischen Ionen ausgesetzt, was neben einem Materialabtrag auch zur Abrundung des Lackprofils führt. Schließlich sorgt die ungenügende Selektivität zwischen Fotolack und Polyimid ($S = r_{\text{Fotolack}}/r_{\text{PI}} \approx 1$) ebenfalls für eine Abflachung der Flanken während des Ätzprozesses und zum kontinuierlichen Abbau der Maske.

Um diese Probleme zu umgehen, erfolgt die Strukturübertragung bei der Polyimidätzung mit Hilfe der Dreilagentechnik [6.35], [6.36]. Das dabei verwendete Dreischichtsystem besteht aus einem Bottom-Layer mittlerer Dicke, der hier durch das Polyimid repräsentiert wird, einer anorganischen Zwischenschicht, die

sich durch eine hohe Selektivität zum Bottom-Layer auszeichnet und einem dünnen fotoempfindlichen Lack an der Oberfläche. Das Polyimid wird durch eine „spin-on“ Technik bei guter Homogenität mit einer Dicke von 1,2 μm auf die Waferoberfläche aufgebracht. Temperaturschritte bei 100 °C (1 min), 125 °C (1 min) und 350 °C (20 min) sorgen für eine gute Haftung zum Substrat und führen zum Aushärten des Bottom-Layers. Die Auswahl des Materials für die Zwischenschicht richtet sich nach den Eigenschaften des Bottom-Layers. In Verbindung mit Polyimid bieten sich z. B. SiO_2 , Al und Ti an. Der Nachteil von Al besteht darin, daß für eine trockenchemische Strukturierung Cl-haltige Ätzgase notwendig sind, die im Rahmen dieser Arbeit nicht zur Verfügung standen. Eine naßchemische Ätzung scheidet ebenfalls aus, da selbst bei dünnen Al-Filmen von 50 nm Dicke die Maßübertragung von der Maske unreproduzierbar ist. SiO_2 und Ti zeigen unter den gegebenen Randbedingungen gute Hafteigenschaften, haben eine hohe Selektivität zu Polyimid und versprechen dadurch eine präzise Maßhaltigkeit. Der entscheidende Vorteil von Ti liegt in der Ausbildung eines TiO_x im O_2 -Plasma, was eine erhöhte Resistenz gegen Sputtereinwirkungen zeigt und zusätzlich eine gesteigerte Selektivität zum Polyimid aufweist [6.37].

Im einem ersten Schritt wird die obere Lackschicht mit einem konventionellen Lithografieprozeß strukturiert. Aufgrund der geringen Lackdicke kommt es zu einer hochauflösenden Strukturübertragung von der Maske mit minimalen Liniendicken von 0,4 μm . Eine nachfolgende CF_4 -basierte RIE-Ätzung überträgt die Lackmaske in die 50 nm dicke Zwischenschicht aus Ti. Der abschließende Ätzprozeß in einem O_2/Ar -Plasma führt, aufgrund der hohen Selektivität zwischen der Ti-Maske und dem Polyimid ($S > 500$), zu einer anisotropen Strukturierung der Gatestege.

Für den Trockenätzprozeß stand eine RIE-Anlage mit induktiv gekoppeltem Plasma (ICP) zur Verfügung. Der spezifische Aufbau des Ätzreaktors erlaubt die kontrollierte Variation des Plasmas mit Hilfe einer hochfrequenzgespeisten Spulenordnung oberhalb des Ätzobjektes. Das Plasma zeichnet sich durch eine große Dichte an Ionen und reaktiven Spezies aus, so daß die erzielten Ätzraten deutlich höher liegen als beim konventionellen reaktiven Ionenätzen. Gleichzeitig erlauben die hohen erzielbaren Plasmadichten eine Reduzierung des Prozeßdruckes, was den Grad der Anisotropie und der Maßhaltigkeit steigert. Der zu ätzende Wafer liegt auf der unteren Elektrode des Reaktors, die kapazitiv an eine zweite Hochfrequenzquelle gekoppelt ist. Dieses Anlagenkonzept ermöglicht es, Plasmadichte und Ionenenergie in physikalisch sinnvollen Grenzen unabhängig

voneinander zu steuern und dadurch die Ätzcharakteristik an die geforderten Kriterien anzupassen.

Im folgenden sollen die Einflüsse der Ätzparameter auf das Profil des Dummy-Gates näher diskutiert werden. Eine entscheidende Bedeutung für das Ätzverhalten hat die Wahl der Gaszusammensetzung. Standardmäßig kommt für die anisotrope Strukturierung organischer Schichten Sauerstoff als Ätzgas zum Einsatz. Durch die Beimischung eines Edelgases, wie z. B. Argon, stabilisiert sich die Plasmareaktion, und es kommt zu einer Verstärkung der physikalischen Ätzwirkung. Des weiteren erlaubt die Variation des Partialdruckes für das inerte Gas eine Optimierung der Uniformität. Aufgrund der fehlenden Polymerbildung steigt die Anisotropie A des O_2 -Ätzprozesses mit sinkendem Partialdruck des Sauerstoffes (Abbildung 6.13). Bei niedrigerem Druck erhöht sich die Substratvorspannung V_T , und die freie Weglänge der Gasteilchen im Plasma vergrößert sich. Dadurch können die Elektronen und Ionen eine höhere kinetische Energie aufnehmen und in den Ätzprozeß einbringen, was zwar die Selektivität verringert, die Flankensteilheit der Strukturen jedoch ansteigen läßt.

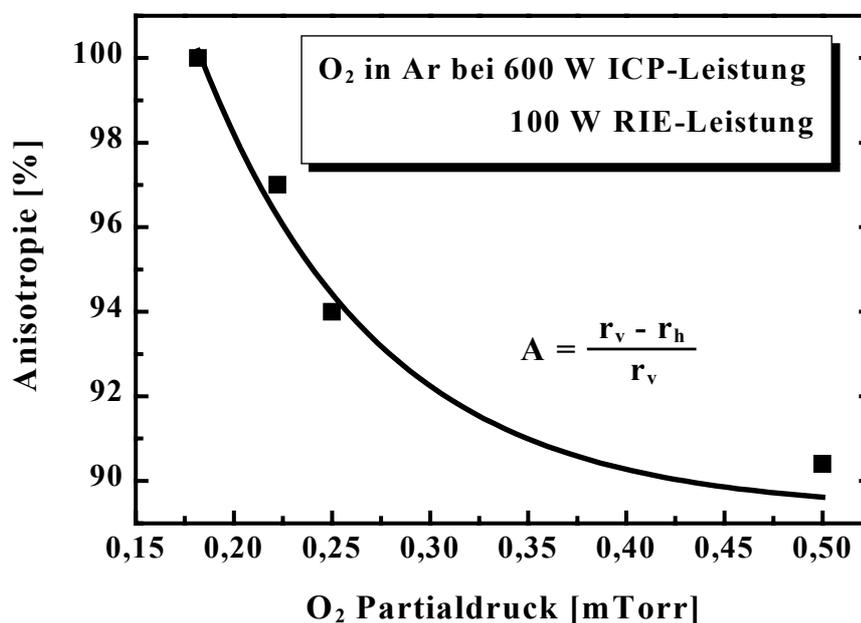


Abbildung 6.13: Abhängigkeit der Anisotropie des O_2/Ar -Ätzprozesses vom Partialdruck des Sauerstoffes bei 600 W ICP-Leistung und 100 W RIE-Leistung

Eine weitere effektive Möglichkeit, auf das Profil der Gatestege Einfluß zu nehmen, ohne dabei den Prozeßdruck zu ändern, besteht durch eine Variation der RIE-Leistung. Die an der unteren Elektrode eingespeiste Hochfrequenzspannung sorgt dafür, daß sich dort ein stark negatives Potential gegenüber dem Plasma bildet. Die positiven Ionen, die im Plasma erzeugt werden, erreichen deshalb den zu ätzenden Wafer mit einer erhöhten kinetischen Energie, was den Materialabtrag ansteigen läßt und die Anisotropie erhöht. Eine thermische Grenze als Limit für die RIE-Leistung spielt bei dem hier betrachteten Dreilagelackprozeß eine untergeordnete Rolle, da beim Ätzen des Polyimides die temperaturunempfindliche Ti-Schicht die maskierende Wirkung übernimmt. Zusätzlich bietet sich bei dem hier verwendeten Ätzverfahren die Möglichkeit, über die ICP-Leistung die Ionengeneration im Plasma zu steuern. Über die variierte Anzahl von Reaktionspezies kann damit die Ätzrate eingestellt werden. Das Verhältnis aus ICP- und RIE-Leistung stellt damit ein zusätzliches Instrument dar, den Charakter der Ätzung zwischen isotrop und anisotrop zu variieren. Die Abhängigkeit der Ätzrate r_{PI} und der Substratspannung V_T von der RIE-Leistung ist in Abbildung 6.14 für verschiedene ICP-Leistungen dargestellt.

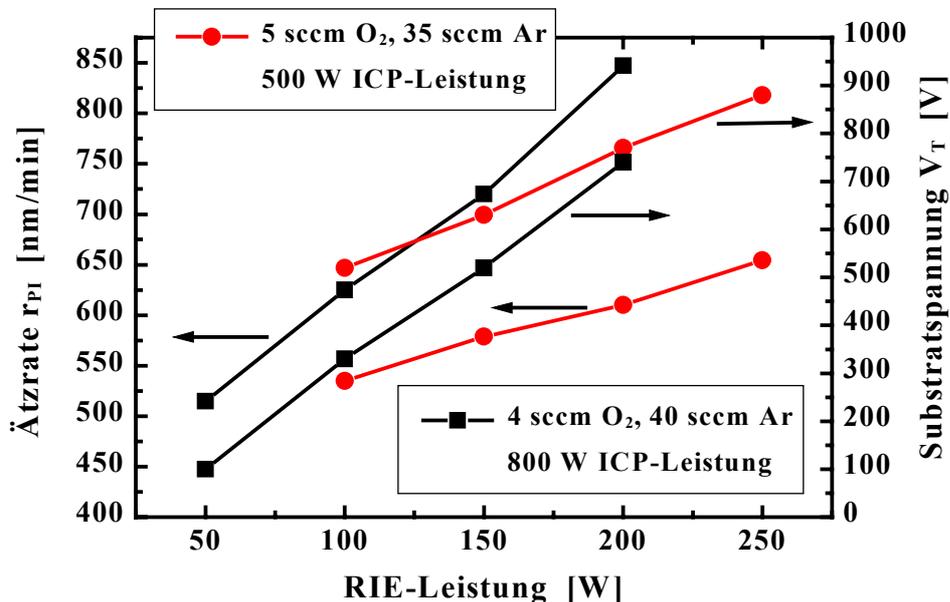


Abbildung 6.14: Abhängigkeit der Ätzrate r_{PI} und der Substratspannung V_T von der RIE-Leistung mit der ICP-Leistung als Parameter

Aus der Vielzahl von durchgeführten Parametervariationen für den Ätzprozeß des Dummy-Gates sind in Abbildung 6.15 drei charakteristische Ergebnisse in Form von Gatequerschnitten dargestellt.

Ätzprofile von Polyimid-Gatestegen

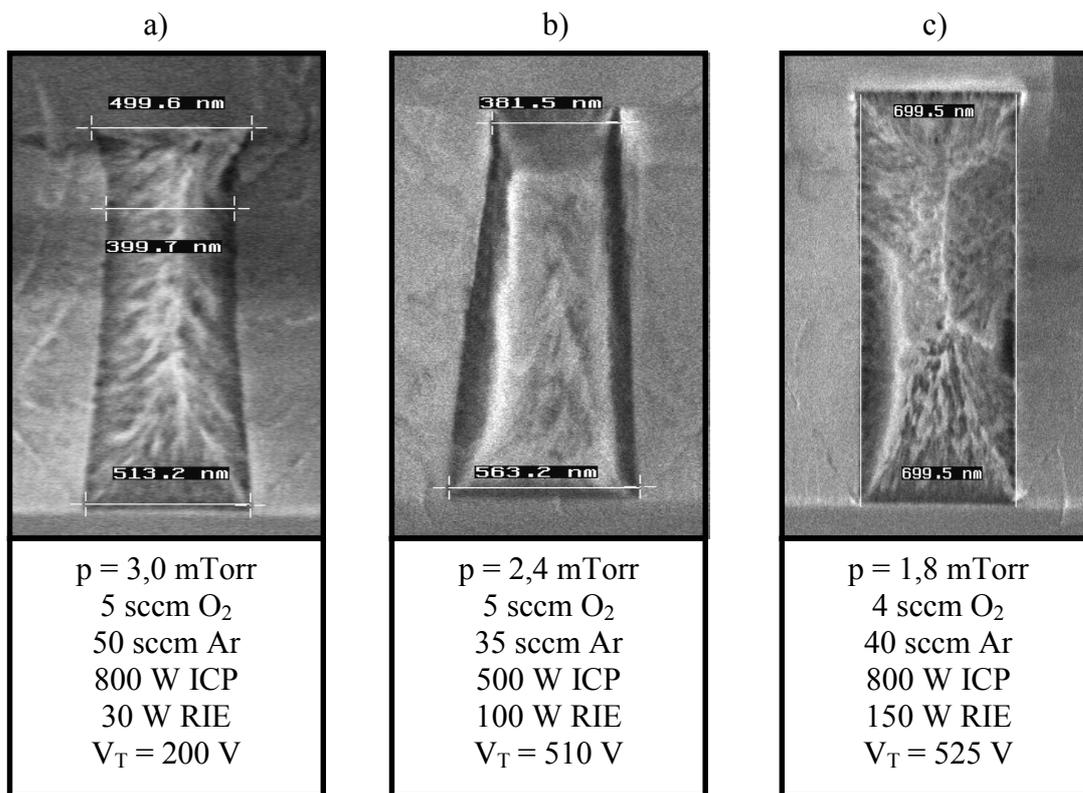


Abbildung 6.15: Charakteristische Polyimidprofile für verschiedene Ätzparameter

Abbildung a) zeigt deutlich eine Unterätzung des Gatekopfes und den daraus resultierenden schrägen Flankenverlauf des Polyimidsteges. Aufgrund der geringen RIE-Leistung und einer Substratspannung von nur 200 V haben die Ionen eine kleine kinetische Energie, so daß der physikalische Anteil der Ätzwirkung reduziert ist. Die hohe ICP-Leistung generiert zwar ein dichtes Plasma, jedoch haben die Gasteilchen wegen des erhöhten Gesamtdruckes von 3 mTorr nur eine verkürzte freie Weglänge. Der isotrope Charakter der Ätzung führt deshalb zu einem gesteigerten Materialabtrag unterhalb der Ätzmaske. Der Gatesteg in Abbildung b) weist zwar keine Unterätzung auf, bietet allerdings nur eine ungenügende Flankensteilheit. Dies ist auf einen relativ hohen O₂-Partialdruck zurückzuführen.

ren, der sich mit einer zu geringen RIE-Leistung paart. Beides ergibt schließlich einen mangelnden Grad an Anisotropie. Abbildung c) stellt schließlich mit einer optimalen Maßhaltigkeit des Ätzprozesses und einer Anisotropie von 100 % den Idealzustand dar.

Der ökonomische und reproduzierbare Einsatz der Mehrlagenlacktechnik und die schädigungsarme Realisierung der Dummy-Gate Struktur verlangen nach einer in-situ Prozeßkontrolle zur Endpunktdetektion. Nach der Definition des Top-Layers mit Hilfe einer konventionellen Kontaktlithografie kann die nachfolgende trockenchemische Strukturierung der Ti-Schicht und des Polyimides in einem Ätzschritt erfolgen, falls eine geeignete Methode zur Endpunkterkennung zur Verfügung steht. Außerdem reduziert sich durch den Einsatz eines entsprechenden Verfahrens die Überätzzeit, die für das Entfernen der Eckenverrundung am Fuß des Polyimidsteges erforderlich ist, auf ein Minimum. Dadurch kann die Schädigung des Heteroschichtsystems durch Strahleneinwirkung und Sputterprozesse verringert werden, und der Linienbreitenverlust hält sich in Grenzen.

Von den beiden am häufigsten verwendeten Verfahren nutzt die Emissionsspektroskopie Informationen über die chemische Zusammensetzung des Plasmas [6.38]. So ist zum Beispiel beim Übergang vom Polyimid zum Si/SiGe-Schichtsystem die Änderung der Konzentration von Reaktanden (hier O_2) oder Ätzprodukten (z. B. CO , CO_2 , C_2H_2) zu beobachten. Die zeitaufgelöste Darstellung charakteristischer Linien oder Bänder des Spektrums gibt dabei Aufschluß über die integrale Komposition des Plasmas und detektiert den Endpunkt. Die Steilheit der gemessenen Intensitätsflanken ist dabei ein Maß für die Gleichmäßigkeit des Ätzprozesses. Da die Änderung des Meßsignals proportional zur ätzenden Fläche ist, eignet sich die Emissionsspektroskopie sehr gut für Ätzprozesse mit einem geringen Bedeckungsgrad der Maskierungsebene.

Das im Rahmen dieser Arbeit eingesetzte Verfahren ist die Laserinterferometrie [6.39]. Bei dieser Methode wird das zu ätzende Schichtsystem mit einem Laser abgetastet. Der reflektierte Laserstrahl setzt sich bei transparenten Schichten aus zwei Anteilen zusammen, einer Komponente, die von der Oberfläche zurückgeworfen wird und einer zweiten, die von der Grenzschicht zum Substrat stammt. Die Phasendifferenz der beiden Anteile hängt von der Dicke der zu ätzenden Schicht ab. Je nach Phasenlage verstärken sich die beiden Strahlen oder schwächen sich ab. Die Intensität des Signals ändert sich deshalb periodisch mit der Schichtdicke.

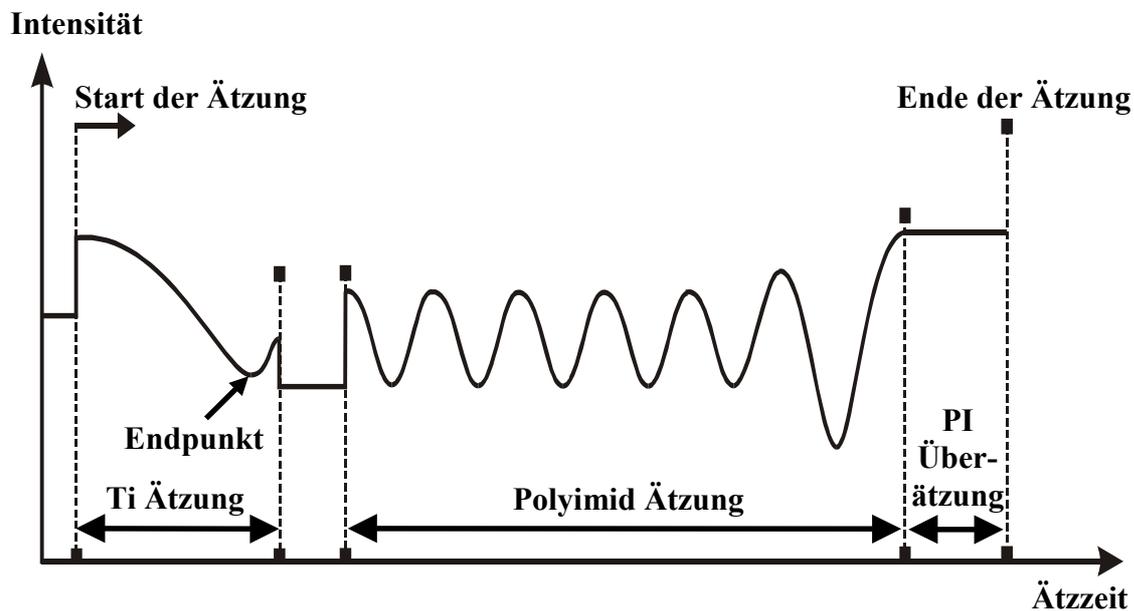


Abbildung 6.16: Laserinterferometersignal bei der Ätzung des Dummy-Gates

Abbildung 6.16 zeigt den schematischen Signalverlauf des Interferometers bei der Ätzung des Dummy-Gates mit der Dreilagentechnik. Der gesamte Verlauf läßt sich in drei Abschnitte einteilen. Den ersten Abschnitt kennzeichnet eine große Wellenlänge, was darauf hindeutet, daß die Ätzrate der 50 nm dicken Titanschicht im CF_4/Ar -Plasma mit ca. 26 nm/min einen niedrigen Wert einnimmt. Am Endpunkt der Titanätzung ändert sich die Reflektivität der Oberfläche, was durch einen Knick im Signalverlauf offensichtlich wird. Eine 10 %ige Überätzung gewährleistet die sichere Entfernung der Metallmaske über den ganzen Wafer. Da die Ätzrate des Polyimides im O_2/Ar -Plasma mit 720 nm/min deutlich höher liegt als beim Titan, verkürzt sich die Periodendauer des Signals im zweiten Abschnitt signifikant. Der Endpunkt der PI-Ätzung zeigt sich unverkennbar an der Grenzlinie zu Bereich 3. Die mangelnde Transparenz des Si/SiGe-Substrates reduziert die Intensität der Interferenz zu Null und läßt nur noch eine waagerechte Linie erscheinen. Um die Verrundung des Polyimidsteges an der Grenzfläche zum Substrat zu eliminieren, findet eine 20 %ige Überätzung statt, die zu einem vertikalen Auftreffen der Gateflanke auf das Si führt.

6.3.2 Naß- und trockenchemische Skalierung des Dummy-Gates

Die trockenchemische Strukturierung des Replacement-Gates erlaubt es, mit Hilfe der Dreilagentechnik reproduzierbar anisotrope Polyimidprofile mit einer Linienbreite von $0,4\ \mu\text{m}$ zu erzeugen. Diese Abmessung spiegelt den späteren Abstand L_{SD} der Source/Drain-Gebiete wider. Nach dem Ätzprozeß wird, gemäß der Ergebnisse aus Abschnitt 6.2.2, eine 10 - 20 nm dicke Nickelschicht ganzflächig mit Hilfe eines Elektronenstrahlverdampfers auf den Wafer aufgebracht. Bei dem nachfolgenden Salizidierungsprozeß wandelt sich das Ni in Kontakt mit dem Si/SiGe-Schichtsystem in ein niederohmiges Silizid um. Da das aktive Transistorgebiet von einem Feldoxid umschlossen ist, läuft der Silizidierungsprozeß an der unbedeckten Mesaoberfläche selbstjustierend zur Dummy-Gate Struktur und dem SiO_2 ab. Auf den isolierenden Flächen bleibt das Ni in seiner metallischen Ursprungsform erhalten und kann anschließend naßchemisch, selektiv zum Silizid entfernt werden. Als allgemein anerkannte Ätzlösung für den Salizidierungsprozeß mit Nickel gilt eine Mischung aus Schwefelsäure und Wasserstoffperoxid [6.40], [6.19]. Während die damit erzielte Selektivität zu SiO_2 beliebig hoch ist, läßt sich für das PI ein merklicher Abtrag feststellen. Da bei einem Mischungsverhältnis von 1:1 die Ätzrate des Nickels etwa 4- bis 5-mal höher liegt als die des Polyimides, hält sich die Schrumpfung des Gatesteges beim Entfernen des Metalles in Grenzen. Es konnte nachgewiesen werden, daß sich das Profil des Gates durch den isotropen Ätzangriff nicht verändert und damit der hohe Grad an Flankensteilheit erhalten bleibt. Jedoch kann sich der Effekt der lateralen Skalierung des Gatesteges positiv auf die Ausbeute und Reproduzierbarkeit des Prozesses auswirken. Um in jedem Fall eine ohmsche Kontaktierung des Si-Kanals und damit des 2DEGs zu erzielen, muß gewährleistet werden, daß die laterale Ausdehnung des Nickels während der Salizidierung die Grenzen des Implantationsgebietes nicht überschreitet. Durch eine gezielte naßchemische Schrumpfung des Polyimidsteges reduzieren sich gleichzeitig die Abmessungen der Implantationsmaske, so daß sich der Dotierbereich der ohmschen Kontakte weiter in Richtung Gatefuß ausdehnen kann als das Silizid.

Abbildung 6.17 stellt die mit Hilfe des Rasterelektronenmikroskops ermittelte laterale Skalierung des Dummy-Gates in Abhängigkeit der Ätzzeit dar. Daraus ergibt sich ein linearer Zusammenhang zwischen dem Materialabtrag und der Dauer des Ätzangriffes mit einer mittleren Ätzrate von etwa 10 nm/min. Unter der Annahme, daß der Siliziumkonsum während des Silizidierungsprozesses in lateraler und vertikaler Richtung gleich ist, ergibt sich bei einer Nickeldicke von 20 nm eine Ausdehnung des Silizides von maximal 40 nm über den ursprüngli-

chen Grenzbereich hinaus. Diese Distanz muß durch eine entsprechend lange Schrumpfung des Gatesteiges in $\text{H}_2\text{SO}_4:\text{H}_2\text{O}_2$ ausgeglichen werden, um die ohmsche Kontaktierung des Kanals zu sichern. Die hier gemachten Annahmen verfolgen ein „worst-case“-Szenario, das sowohl die laterale Diffusion des Dotierstoffes bei der Aktivierung als auch den Einfallswinkel der Ionen bei der Implantation unberücksichtigt läßt.

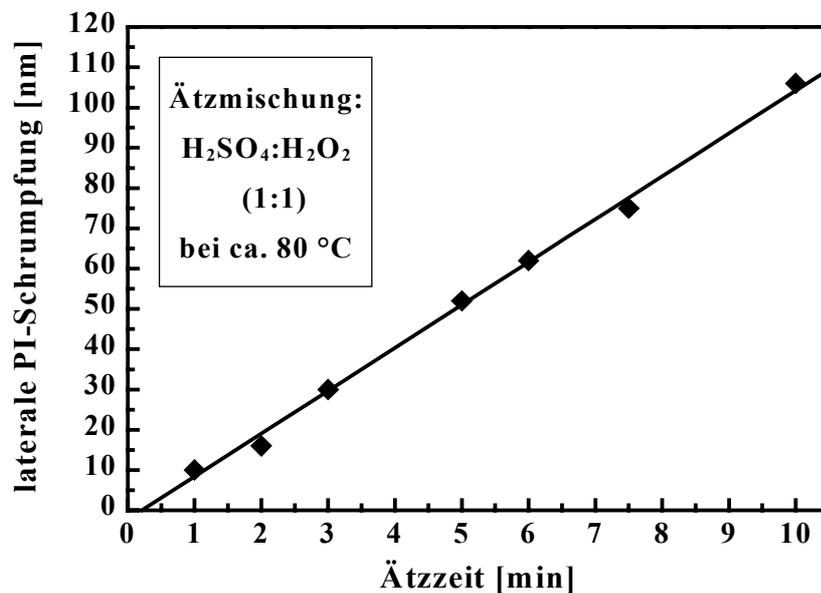


Abbildung 6.17: Lateraler Ätzabtrag am PI-Gate bei der naßchemischen Skalierung in Abhängigkeit der Zeit

Nach der Implantation der Source/Drain-Gebiete erfolgt die Skalierung des Replacement-Gates auf die endgültige Gatelänge. Der dabei verwendete Ätzprozeß muß sich durch eine exakte Konservierung der vertikalen Polyimidflanken, eine gute Homogenität über die Waferfläche und ein hohes Maß an Reproduzierbarkeit auszeichnen. Da Versuche mit einem naßchemischen Verfahren bei entsprechend langen Ätzzeiten nur eine unzureichende Gleichmäßigkeit der Ätzrate über die gesamte Waferfläche ergaben, fiel die Entscheidung auf einen trockenchemischen Prozeß. Wie bei der Strukturierung des Gatesteiges kommt auch bei der Skalierung eine sauerstoffbasierte Ätzung zum Einsatz. Jedoch erfordert der überwiegend isotrope Charakter des Ätzschrittes eine komplette neue Anpassung der Prozeßparameter. Eine drastische Erhöhung des Druckes und eine signifi-

kante Reduzierung der RIE-Leistung sorgen dafür, daß die Elektronen und Ionen nur wenig an kinetischer Energie gewinnen und die chemische Ätzwirkung mehr in den Vordergrund tritt. Durch die gleichzeitige Steigerung des O₂-Partialdruckes und der ICP-Leistung tritt eine verstärkte Generierung reaktiver Spezies ein, die das Niveau der Ätzrate hoch halten und für einen gleichmäßigen Ätzangriff an horizontalen und vertikalen Flanken sorgen. Tabelle 6.4 zeigt diverse Parametervariationen bei der Skalierung des Dummy-Gates und deren charakteristische Merkmale.

| Parameter | Variation 1 | Variation 2 | Variation 3 | Variation 4 | Variation 5 |
|--|-------------------------------|-----------------------------|-----------------------------|----------------------------------|------------------------------|
| Gasgemisch O ₂ /Ar [sccm] | 25/25 | 5/5 | 5/5 | 5/5 | 5/5 |
| Leistung ICP/RIE [W] | 400/15 | 200/0 | 400/0 | 400/15 | 400/30 |
| Druck/V _T [mTorr] / [V] | 60/400 | 78/30 | 78/30 | 78/170 | 78/210 |
| Ätzrate [nm/min] r _{lat.} /r _{vert.} | 21/43 | 48/0 | 61/0 | 21/42 | 52/81 |
| Merkmale | inhomogene Ätzrate über Wafer | schlechte Profilkonformität | schlechte Profilkonformität | gutes Profil mittlere Ätzrate | gutes Profil hohe Ätzrate |

Tabelle 6.4: Parametervariation bei der trockenchemischen Dummy-Gate-Skalierung

Die Gegenüberstellung der verschiedenen Ätzvarianten für die Skalierung und die daraus resultierenden Profile der Dummy-Gates (Abbildung 6.18) zeigen deutlich, daß für eine konforme Schrumpfung des Polyimidsteges eine präzise Abstimmung der Prozeßparameter nötig ist. Insbesondere das Verhältnis aus ICP- und RIE-Leistung hat großen Einfluß auf die Flankensteilheit. Bei einem zu hohen RIE-Anteil (> 50 W) wächst die vertikale Ätzrate stark an, und die laterale Komponente verliert kontinuierlich an Einfluß. Sinkt die RIE-Leistung unter 10 W, überwiegt der isotrope Ätzcharakter, und es ergeben sich signifikante Unterätzungen (Variation 2 und 3). Aufgrund der wenig ausgeprägten Topographie und des geringen Bedeckungsgrades des Wafers ist die Gleichmäßigkeit der Ät-

zung sehr stark von Gasfluß und Druck abhängig. Erst nach einer 30 %igen Erhöhung des Prozeßdruckes und einer drastischen Verringerung der Gasflüsse um 80 %, stellte sich eine ausreichende Homogenität der lateralen Ätzrate von +/- 10 % über die gesamte Waferfläche ein.

Parametervariation bei der Dummy-Gate Skalierung

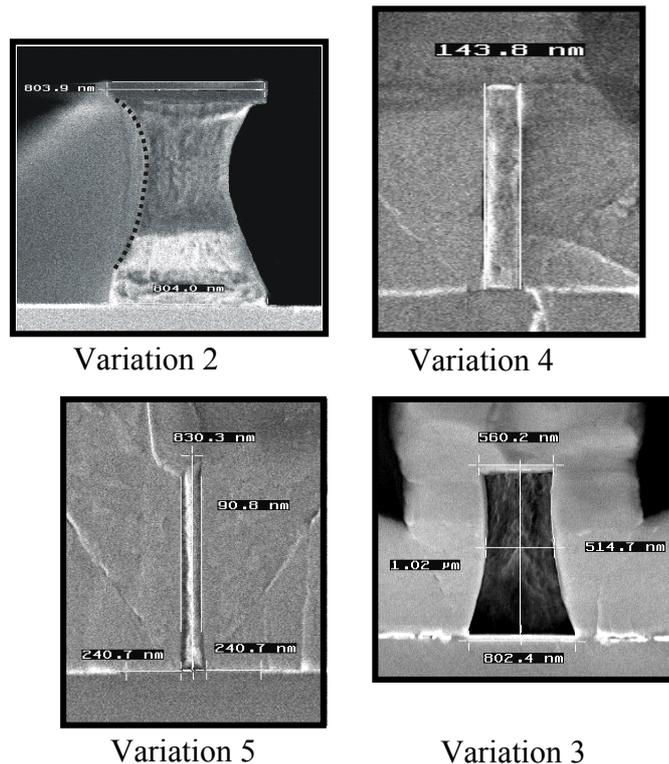


Abbildung 6.18: Querschnitte von in Oxid eingebetteten Polyimidstegen nach der Schrumpfungszätzung bei unterschiedlichen Parametervariationen

Für Parametervariation 4 ist in Abbildung 6.19 der Verlauf der lateralen und vertikalen Skalierung in Abhängigkeit von der Ätzzeit dargestellt. Beide Komponenten zeigen eine lineare Zunahme mit steigender Prozeßdauer bei einer mittleren Ätzrate von 21 nm/min in lateraler und 42 nm/min in vertikaler Richtung. Diese Resultate konnten in mehreren Versuchsreihen bei gleichen Randbedingungen bestätigt werden, so daß auch die Reproduzierbarkeit in ausreichendem Maße gesichert ist. Damit stellt der entwickelte Ätzschritt ein effektives Instrument dar, um Dummy-Gate-Strukturen mit einer ursprünglichen Breite von 0,5 - 0,4 µm reproduzierbar und mit guter Homogenität ohne einen Verlust an Flankensteilheit auf Gatelängen von bis zu 75 nm zu skalieren.

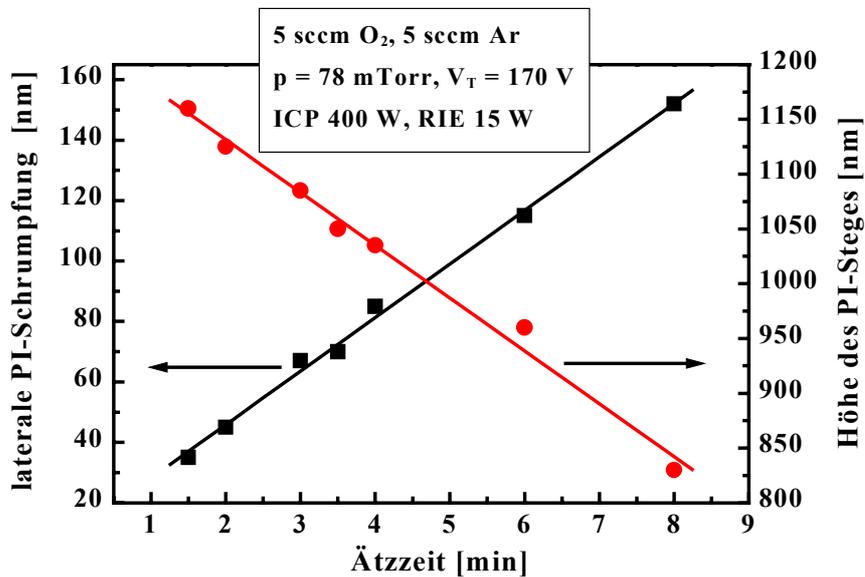


Abbildung 6.19: Laterale und vertikale Skalierung des Dummy-Gates in Abhängigkeit von der Ätzeit für Parametervariation 4 (5 sccm O₂, 5 sccm Ar, $p = 78$ mTorr, $V_T = 170$ V, ICP 400 W, RIE 15 W)

6.4 Planarisierungskonzepte des SiGe-HFETs

Mit der permanent fortschreitenden Miniaturisierung in der Halbleitertechnologie und den daraus resultierenden hohen Packungsdichten wächst auch die Bedeutung effektiver Planarisierungskonzepte. Viele typische Prozeßmodule zur Herstellung mikroelektronischer Bauelemente und Schaltungen, wie z. B. Isolationschritte, die Definition von Gate- und Emitterstrukturen, die Kontaktlochätzung oder die Realisierung von Metallisierungsebenen, hinterlassen eine charakteristische Topographie auf der Halbleiteroberfläche. Um bei zunehmender Skalierung der minimalen Abmessungen trotzdem eine reproduzierbare Integration zu gewährleisten, ist die Einebnung der Oberfläche eine primäre Forderung der ULSI-Technologie. Die Randbedingungen und zentralen Anforderungen an einen Planarisierungsprozeß werden dabei von den Bereichen Lithografie, Bauelementisolation und Mehrlagenmetallisierung definiert.

In der optischen Lithografie liefert die derzeit noch weit verbreitete DUV-Technik mit einer Wellenlänge von 248 nm eine maximale Tiefenschärfe von 0,7 μm . Beim Übergang in den FUV-Bereich (193 nm) reduziert sich dieser Wert

auf $0,55\ \mu\text{m}$ [6.41]. Um bei zunehmender Strukturverkleinerung weiterhin eine bestmögliche Auflösung zu erzielen, muß der verwendete Fotolack sehr dünn und homogen in der Schichtdicke sein. Eine planarisierte Oberfläche ist deshalb die Grundvoraussetzung, um Linienbreitenschwankungen an topografischen Kanten zu vermeiden und eine maßhaltige Strukturübertragung durch die Lithografie zu gewährleisten.

Die Isolation des aktiven Bauelementes greift meistens auf Standardverfahren, wie die Trench-Isolation oder die LOCOS-Technik zurück. Diese Technologien sind nicht in der Lage, in-situ eine planare Oberfläche zu erzeugen. Da die Isolation bereits in einem frühen Stadium des Prozeßablaufes stattfindet, ist ein ebenes Fundament für den weiteren Herstellungsprozeß und insbesondere die nachfolgenden Metallisierungsebene erstrebenswert.

Die zunehmende Packungsdichte der aktiven Bauelemente erfordert zusätzliche Metallisierungsebenen, um die gewünschte Schaltungskonfiguration realisieren zu können und durch funktionale Gliederung der Verdrahtungsebenen kürzere und schnellere Signalfade zu generieren. Die dabei erzeugten Topographien werden von den aufgetragenen Isolationsschichten oft nur ungenügend planarisiert, so daß die Metalleitungen an den Kantenübergängen schlechte Bedeckungsgrade zeigen. Bei abnehmenden Leiterquerschnitten und gleichzeitig hohen Stromdichten steigt die Gefahr des Materialtransportes durch Elektromigration, was schließlich zu Ribbildungen oder Leitungsunterbrechungen führen kann. Um die daraus resultierenden Einbußen in der Zuverlässigkeit und der Ausbeute zu minimieren, sind Planarisierungsverfahren für Prozesse mit Mehrlagenmetallisierungen essentiell wichtig.

Die derzeit verfügbaren Planarisierungskonzepte unterteilen sich in drei unterschiedliche Verfahrensprinzipien:

- PVD- und CVD-SiO₂ Verfahren
- Spin-Coating-Techniken, z. T. gepaart mit Rückätzungsschritten
- CMP (chemical mechanical polishing)

Die Verwendung von PVD- und CVD-SiO₂ Verfahren zur Herstellung planarierter Isolationsschichten für Mehrlagenmetallisierungen ist stark von dem Temperaturbudget und den spezifischen Eigenschaften des Prozesses abhängig.

Da der Anteil an metallischen Leitbahnen aus Aluminium in der Halbleitertechnologie weiterhin hoch ist, beschränkt sich die maximale Prozeßtemperatur auf ca. 500 °C. Deshalb konzentriert sich die Anwendung zumeist auf plasmaunterstützte Silanverfahren, dotierte CVD-Oxide, eine modifizierte Form der TEOS-Methode und gesputterte SiO₂-Schichten. PECVD-Oxide zeichnen sich neben einer niedrigen Reaktionstemperatur durch eine geringe Rißbildung an topographischen Kanten aus, was auf die Zuverlässigkeit der nachfolgenden Metallisierung großen Einfluß hat [6.42]. Eine Beimischung von Phosphin (PH₃) oder/und Diboran (B₂H₆) während der Abscheidung erhöht das Fließvermögen der Schichten bei entsprechender Temperaturbehandlung [6.43]. Die daraus resultierende planarisierende Wirkung des Flow-Glases setzt bei Temperaturen von 900 - 1000 °C ein, was seine Anwendung auf den Zeitraum vor der Metallabscheidung beschränkt. Eine modifizierte Form des TEOS-Prozesses erlaubt es, durch Zugabe von O₃ die Abscheidetemperatur bei gleichbleibender Konformität auf unter 450 °C zu senken und stellt damit eine gute Alternative zum PECVD-Oxid dar [6.41]. Der nur mäßige Planarisierungsgrad kann durch eine Kombination mit Spin-on-Materialien gesteigert werden. Das Hochfrequenz-Bias-Sputtern ist eine rein physikalische Depositionsmethode, die deshalb bei niedrigen Temperaturen stattfinden kann. Sie nutzt dabei eine unterschiedliche Abscheidungsrate an lateralen und vertikalen Flächen zur Einebnung der Topographie aus [6.44], [6.45]. Im Gegensatz zu den CVD-Verfahren zeigen die Schichten eine erhöhte mechanische und chemische Stabilität.

Die Spin-Coating-Technik profitiert von der planarisierenden Wirkung viskoser, organischer Stoffe, wie z. B. Fotolack, Spin-On-Glas (SOG) oder Polyimid. Nach dem konventionellen Aufschleudern der Flüssigkeiten werden sie einer Temperaturbehandlung ausgesetzt. Dabei verflüchtigt sich der Lösungsmittelanteil, und es findet eine Umwandlung bzw. Vernetzung des Grundstoffes zu einer festen Isolierschicht statt. In Abhängigkeit von der Viskosität, der Schichtdicke, den Aufschleuderparametern und der Topographie ergibt sich eine breite Spanne von Planarisierungsgraden zwischen 40 % und 70 % [6.46], [6.47]. Um eine weitere Einebnung der Oberfläche zu erzielen, bietet sich die Kombination mit einem Rückätzschritt an. Dabei wird zuerst eine konventionelle SiO₂-Schicht im PVD- oder CVD-Verfahren abgeschieden und anschließend durch Aufbringen des Spin-on-Materials eine Planarisierung durchgeführt [6.48]. Durch eine isotrope Rückätzung mit einer Selektivität von eins überträgt sich das Profil in die SiO₂-Schicht und läßt so eine ebene Oberfläche entstehen. Eine Variante dieses Verfahrens nutzt eine Hilfsschicht, die eine inverse Topographie zum bestehenden Profil erzeugt und so vor dem eigentlichen Planarisierungsschritt eine Anglei-

chung in den Höhenunterschieden der Oberfläche durchführt [6.49]. Die hier beschriebenen Modifikationen der konventionellen Spin-Coating-Technik sind in der Lage, den Planarisierungsgrad auf über 80 % zu steigern.

CMP ist das derzeit dominierende Planarisierungsverfahren für Dielektrika und Metalle. Dabei werden die Unebenheiten der Waferoberfläche durch chemisch-mechanisches Polieren begradigt. Die dabei verwendete Poliersubstanz enthält neben Quarz- und Oxidzusätzen, die den mechanischen Abtrag verstärken sollen, eine chemische Komponente, die in ihrer Ätzwirkung auf das zu bearbeitende Material abgestimmt ist [6.41]. Da die Selektivität zu benachbarten Schichten begrenzt ist, werden spezielle Barrierschichten oder Techniken zur Endpunkterkennung eingesetzt, um die Prozeßkontrolle zu erhöhen. Das Verfahren zeichnet sich durch eine einfache und gut kontrollierbare Prozeßführung aus und weist eine hohe Skalierbarkeit auf. Die damit erzielbaren Planarisierungsgrade sind strukturabhängig und liegen für typische Anwendungen, wie z. B. der Definition von W-Plugs, der Planarisierung von ILDs und Feldoxiden oder der Damascene-Technologie deutlich über 80 %. Die Nachteile des Prozesses gehen primär von der mechanischen Komponente aus und zeigen sich in Form von Erosionen im Dielektrikum und durch lokale konkave Unebenheiten im bearbeiteten Metall [6.50]. Außerdem kann eine verstärkte Kontamination der Oberfläche mit Partikeln festgestellt werden, was zusätzliche Reinigungsschritte nötig macht [6.51].

Bei der Herstellung des SiGe-HFETs scheidet aufgrund der begrenzten technologischen Laborausstattung eine Planarisierung mit dem CMP- und SOG-Verfahren aus. Außerdem besteht keine Möglichkeit, ein PVD- oder TEOS-Oxid zu realisieren. Die im folgenden dargestellten Planarisierungskonzepte für die Mesaisolation und die Passivierung des Bauelementes greifen dabei auf ein PECVD-Oxid zurück und nutzen Spin-Coating-Techniken in Verbindung mit Rückätzungsschritten, um die gewünschte Einebnung der Oberfläche zu gewährleisten.

Die primäre Motivation zur Realisierung einer selbstjustierenden, planaren Anordnung des Feldoxides liegt neben der Erzeugung einer ebenen Oberfläche in der Vermeidung einer Feldoxidüberlappung auf die Mesaoberfläche (Abbildung 6.20). Bei nicht-planarisierten Feldoxidstrukturen wird auf der Mesa ein Fenster geöffnet, um das Oxid dort partiell zu entfernen. Dabei bleibt am Mesarand ein Bereich erhalten, wo die Oxidschicht weiterhin die Mesaoberfläche überlappt. In diesem Randbereich hat das Gate durch die mehrere 100 nm dicke Oxidschicht

aber keine Steuerwirkung mehr auf die Ladungsträger im Kanal, so daß der Stromfluß bei „normally-on“-Transistoren nicht vollständig unterdrückt werden kann. Dadurch ergeben sich signifikante Leckströme, die das R_{on}/R_{off} -Verhältnis der Transistors maßgeblich verschlechtern.

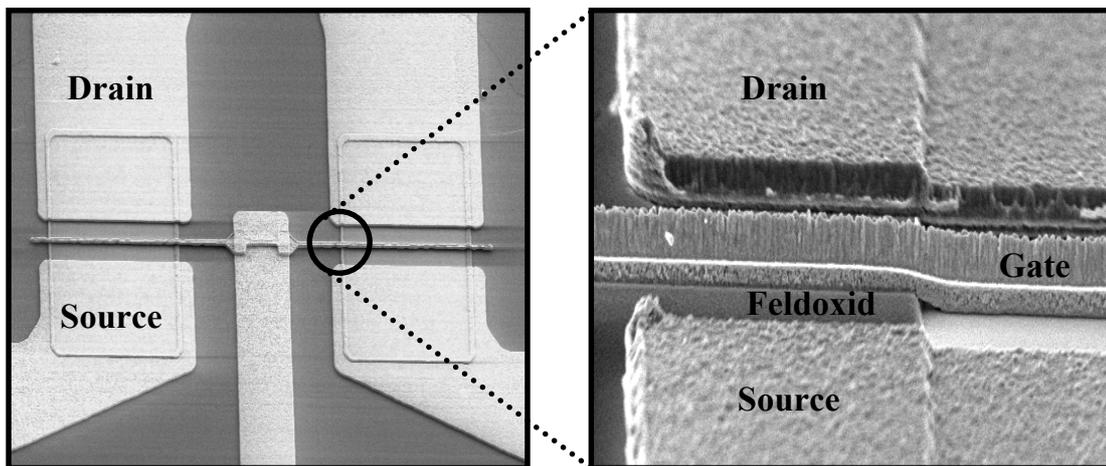


Abbildung 6.20: Konventioneller Doppelmesa SiGe-HFET mit überlappender Feldoxidanordnung

Abbildung 6.21 zeigt zwei entwickelte Prozeßsequenzen zur Herstellung eines selbstjustierenden Feldoxides. Beide Varianten sind sowohl für das konventionelle als auch das selbstjustierende Integrationskonzept einsetzbar. Als Ausgangssituation ergibt sich eine Mesastruktur, die je nach Gateweite des Transistors eine Fläche zwischen $22 \cdot 14 \mu\text{m}^2$ und $22 \cdot 54 \mu\text{m}^2$ umfaßt.

Der Prozeß FOX 1 beginnt mit der ganzflächigen Abscheidung einer PECVD-Schicht, deren Dicke d_{FOX} mit der Mesahöhe d_{Mesa} übereinstimmt. Anschließend wird eine $1,2 \mu\text{m}$ dicke Fotolackschicht aufgeschleudert und bei $130 \text{ }^\circ\text{C}$ für 5 min planarisiert. Mit Hilfe eines RIE-Prozesses erfolgt die Rückätzung des Fotolackes. Dabei sorgt eine in-situ Endpunktkontrolle mit dem Laserinterferometer für eine präzise Freilegung der Oxidoberfläche oberhalb der Mesa. Seitlich der Mesa bleibt der Lack auf dem Oxid bestehen und dient als selbstjustierende Maske für die nachfolgende SiO_2 -Ätzung mit gepufferter Flußsäure. Da die planarisierte Lackdicke eine Funktion der Strukturdimensionen und des topographischen Profils ist, zeigen sich insbesondere im Randbereich des Wafers nach der Rückätzung auf großflächigen Strukturen Lackreste auf dem Oxid. Zur vollständigen Entfernung des Oxides müssen diese Mesaoberflächen mit einer Zusatzmaske nachbehandelt werden.

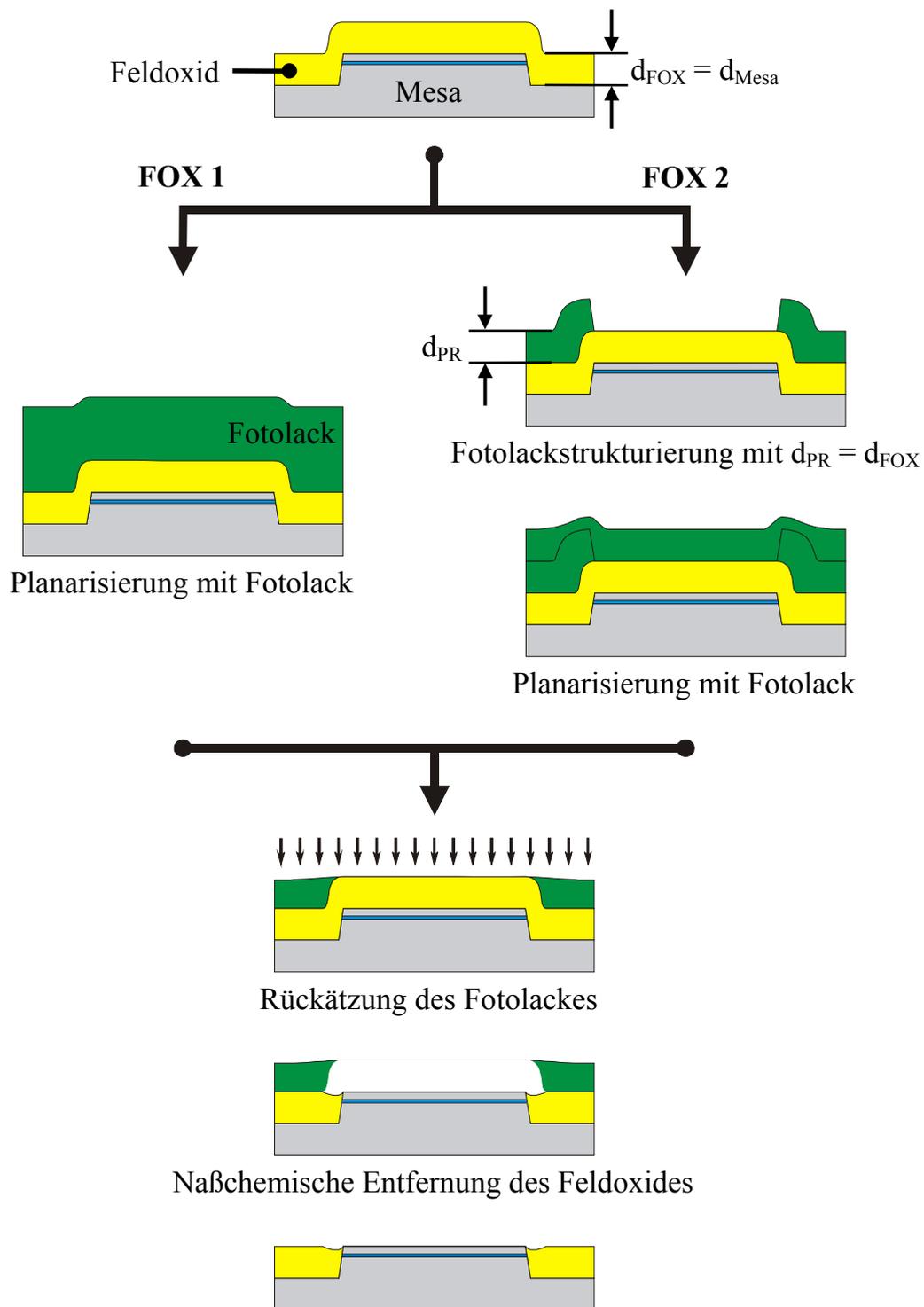


Abbildung 6.21: Prozeßsequenzen FOX 1 und FOX 2 zur Herstellung einer selbstjustierenden, planaren Feldoxidanordnung

Dieser zusätzliche Prozessschritt läßt sich durch die Einführung einer Hilfsschicht im Prozeßablauf FOX 2 vermeiden. Die Basis stellt hier wieder die ganzflächig passivierte Mesastruktur dar. Mit Hilfe einer Fotolackschicht, deren Dicke mit der Mesahöhe übereinstimmt, wird eine Schichtstruktur mit einem komplementären Profil zur ursprünglichen Topographie erzeugt. Durch eine Temperaturbehandlung bei 150 °C verliert der Lack seine Lichtempfindlichkeit und baut eine Resistenz gegenüber Lösungsmitteln ein. Dadurch kann nun eine zweite Lackschicht zur eigentlichen Planarisierung der Gesamtstruktur aufgeschleudert werden. Wie bei der ersten Methode erfolgt nun eine Rückätzung, um die passivierende Oxidschicht auf der Mesa wieder freizulegen. Die abschließende naß-chemische Entfernung des Feldoxides nutzt die angrenzende Lackschicht als Ätzmaske, so daß nur die Mesaoberfläche vom Oxid befreit wird. Aufgrund der vorstrukturierten Hilfsschicht und der daraus resultierenden reduzierten Topographie sinken bei diesem Verfahren die Anforderungen an den eigentlichen Planarisierungsschritt. Zudem entkoppelt sich der Planarisierungsgrad von der Strukturgröße und erlaubt deshalb auch die komplette Freilegung größerer Mesaflächen.

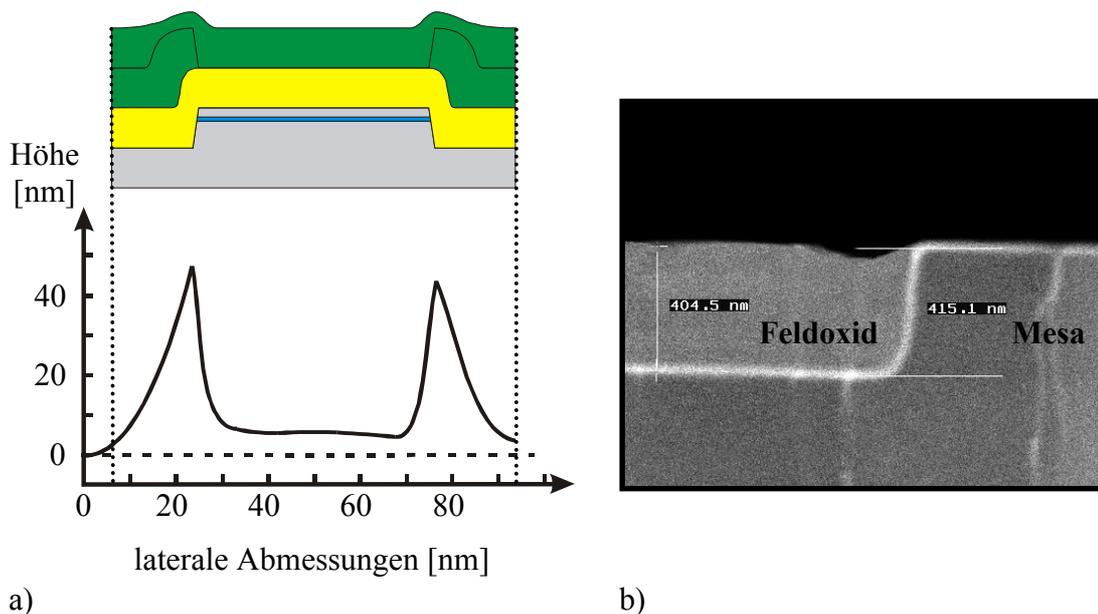


Abbildung 6.22: a) Höhenprofilmessung mit dem DEKTAK nach der Lackplanarisierung für Prozeß FOX 2; b) Querschnitt einer selbstjustierenden, planaren Feldoxidstruktur im REM

Abbildung 6.22 a) zeigt das mit dem DEKTAK gemessene Höhenprofil für Prozeß FOX 2 nach der Lackplanarisierung. Die dabei am Rand der Mesa hervor-

tretenden Signalspitzen ergeben sich aus den Kanten der vorstrukturierten Hilfschicht und übertragen sich reduziert in die zweite Lackschicht. Der dabei erzielte Planarisierungsgrad ermittelt sich aus dem Höhenunterschied der Struktur vor und nach der Einebnung und liegt zwischen 80 % und 90 %. Eine Querschnittsansicht des Grenzbereiches zwischen Mesa und Feldoxid nach der Planarisierung zeigt Abbildung 6.22 b). Es ist deutlich zu erkennen, daß die Feldoxidschicht einen nahezu planaren Übergang zur Mesaoberfläche schafft und mit dieser an der oberen Kante bündig abschließt. Dadurch ergibt sich eine vollständige Steuerung des Kanals durch das Gate über den gesamten aktiven Bereich der Mesa und eine ebene Oberfläche als Basis für die weiteren Integrationsschritte.

Die zweite Anwendung eines Planarisierungsverfahrens beim SiGe-HFET betrifft die Passivierung des aktiven Bauelementbereiches (Abbildung 6.24). Dazu wird nach der Skalierung des Replacement-Gates auf die endgültige Gatelänge die gesamte Transistorstruktur mit einer 1000 nm dicken PECVD-Oxidschicht überzogen. Dabei gilt es, den auf etwa 800 -1000 nm Höhe geschrumpften Gatestege komplett zu isolieren. Zur Planarisierung kommt eine 1,2 µm dicke Polyimidschicht zum Einsatz. Sie wird mit der Spin-Coating-Technik bei 4000 min^{-1} aufgeschleudert und nachfolgend bei $100 \text{ }^\circ\text{C}$ (1 min), $125 \text{ }^\circ\text{C}$ (1 min) und $350 \text{ }^\circ\text{C}$ (20 min) eingebnet und ausgehärtet.

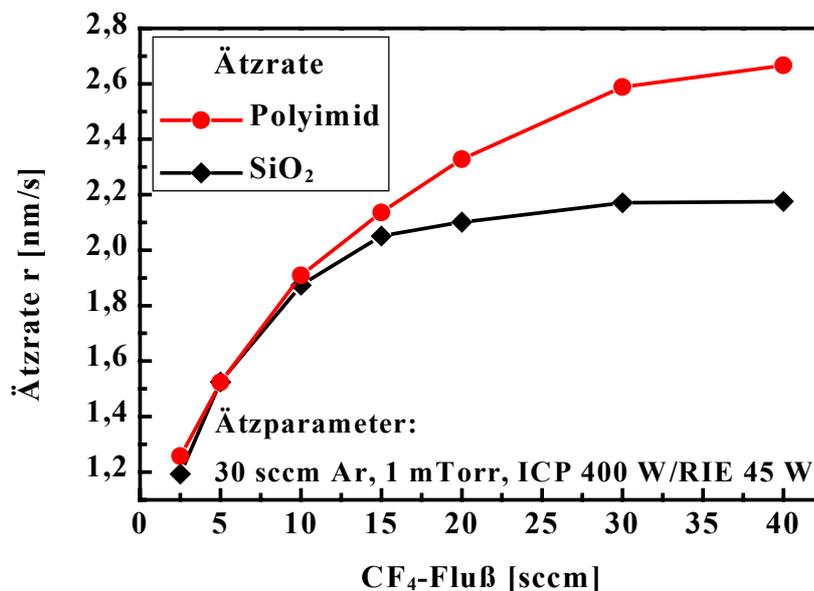


Abbildung 6.23: Abhängigkeit der Ätzrate des Polyimides und des SiO₂ vom CF₄-Fluß bei der Planarisierungsätzung

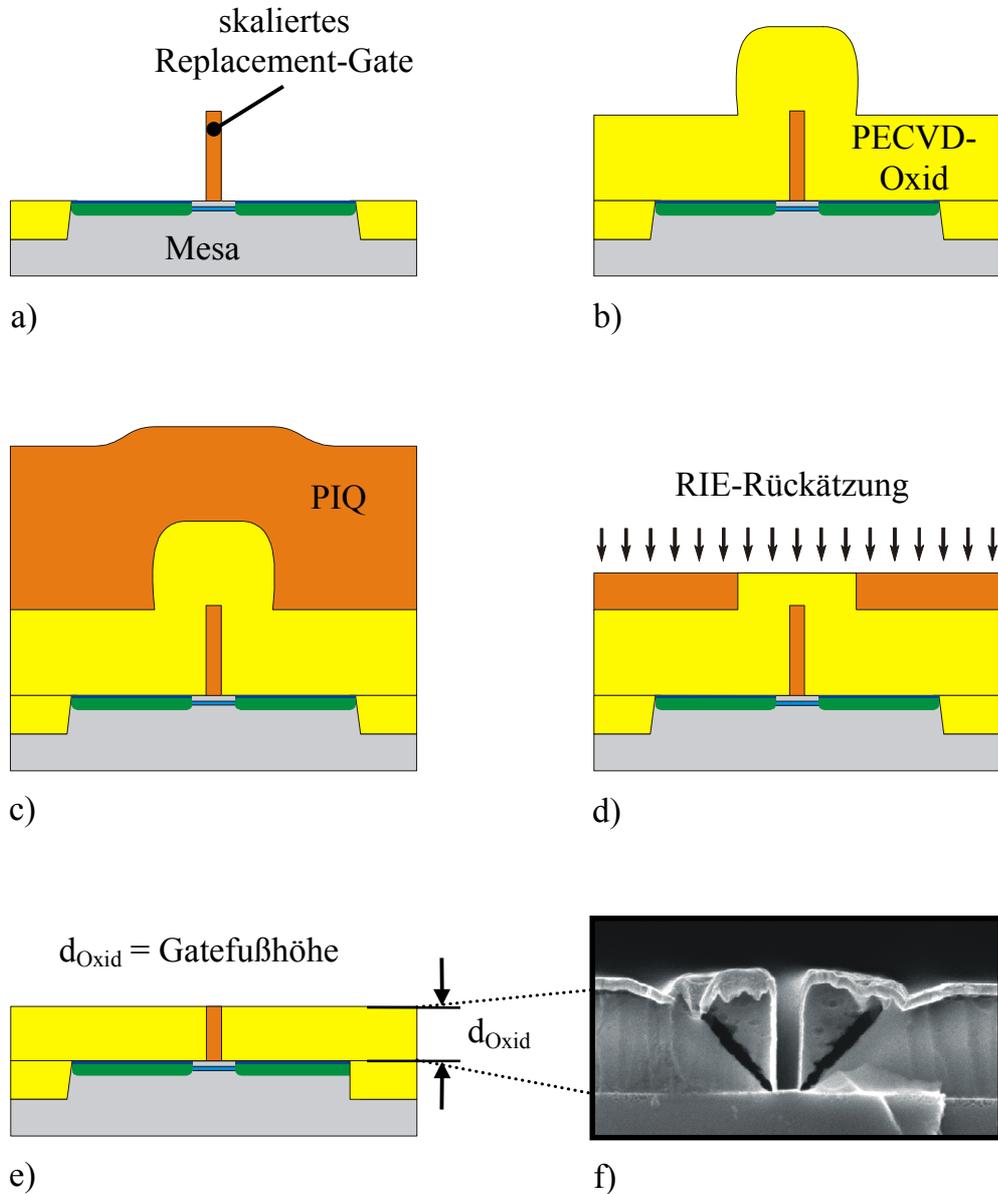


Abbildung 6.24: Prozeßablauf zur Planarisierung der Bauelementpassivierung

Durch eine Rückätzung mit einem RIE-Prozeß überträgt sich die planarisierte Oberfläche in das Oxid. Dabei ist es wichtig, daß die Ätzselektivität zwischen Oxid und Polyimid durch die Anpassung des CF_4 -Flusses bei eins gehalten wird, um eine ebene Oberfläche nach dem Prozeß zu gewährleisten (Abbildung 6.23). Die Kontrolle der Ätzung und die Bestimmung des Endpunktes übernimmt dabei das Laserinterferometer. Aufgrund der entkoppelten Realisierung von Gatefuß

und Gatekopf läßt sich über die Dauer des RIE-Prozesses die Dicke der verbleibenden Oxidschicht und damit die Höhe des späteren Gatefusses individuell einstellen. Zusätzlich sichert die begradigte Topographie der Oxidschicht eine hohe Tiefenschärfe für den Lithografieschritt zur Definition des Gatekopfes und ermöglicht dadurch eine maßhaltige Strukturübertragung. Der bei diesem Prozeß erzielbare Planarisierungsgrad liegt, trotz einer ausgeprägten Topographie mit abrupten Stufen von bis zu 1000 nm, bei etwa 70 %.

Die im Rahmen dieser Arbeit entwickelten Planarisierungsprozesse gestatten es, mit Hilfe einfacher Standardverfahren der Halbleitertechnologie die Integrationsfähigkeit, insbesondere des selbstjustierenden SiGe-HFET Konzeptes, für Schaltungsanwendungen mit Mehrlagenmetallisierungen zu verbessern. Die im Laufe des Prozeßablaufes auftretenden Topographien, z. B. durch die Mesaisolation und die T-Gate-Konstruktion, werden mit einfachen Methoden effektiv planarisiert und realisieren so einen hohen Integrationsgrad.

6.5 Gaterealisierung

Die elektrischen Eigenschaften und das Leistungspotential des SiGe-HFETs werden maßgeblich von dem Schottky-Gate als Steuerelektrode bestimmt. Dabei spielen nicht nur die Bauart und der Herstellungsprozeß für die Realisierung hochfrequenztauglicher Gatekonstruktionen eine entscheidende Rolle, sondern auch Materialeigenschaften und Grenzschichtreaktionen, um Stabilitäts- und Reproduzierbarkeitskriterien zu erfüllen. Die zentralen Forderungen an das Schottky-Gate lassen sich wie folgt zusammenfassen:

- ultrakurze Gatelänge
- niedriger Gatewiderstand
- hohe Schottky-Barriere zu Silizium
- gute elektrische und thermische Stabilität
- optimale Haftung und mechanische Festigkeit
- einfacher, kostengünstiger und reproduzierbarer Herstellungsprozeß

Die physikalischen und technologischen Maßnahmen zur Erfüllung dieser Ansprüche weisen teilweise gegensätzliche Wirkungsrichtungen auf, so daß die Optimierung der Eigenschaften wie so oft einen Kompromiß erfordert.

Die folgenden Abschnitte beschäftigen sich mit der optimalen Umsetzung der vorgestellten Gatekonzepte. Dazu wird zunächst der Schottky-Übergang mit der Zielsetzung einer möglichst effektiven Barriere, bei gleichzeitig hoher elektrischer und thermischer Stabilität untersucht. Danach erfolgt die Beschreibung des Herstellungsprozesses für die konventionelle Gatekonstruktion mit e-Beam Lithografie. Abschließend werden unterschiedliche technologische Verfahren zur Realisierung des selbstjustierenden Gatekonzeptes diskutiert und auf der Basis der in Abschnitt 6.3 und 6.4 geschaffenen Randbedingungen eine erfolgreiche Methodik erarbeitet und umgesetzt.

6.5.1 Entwicklung temperaturstabiler Schottky-Übergänge auf Pt/Si-Basis

Die Strom-Spannungs-Charakteristik idealer Metall-Halbleiter-Übergänge zeigt bei Raumtemperatur eine starke Abhängigkeit von den involvierten Materialparametern. Dabei spielt nicht nur das verwendete Metall mit seiner spezifischen Austrittsarbeit Φ_M eine große Rolle, sondern auch die Art und Höhe der Halbleiterdotierung N_D . Sie hat insbesondere signifikanten Einfluß auf den dominierenden Strommechanismus in der Diodenstruktur. Unter Vernachlässigung von Rekombinations- und Generationsvorgängen wird der Ladungsträgertransport durch die Metall-Halbleiter-Grenzschicht von drei physikalischen Prozessen bestimmt: a) thermionische Emission, b) quantenmechanische Tunnelströme und c) thermionische Feldemission [6.52].

Die thermionische Emission ist der bestimmende Stromanteil bei geringen Halbleiterdotierungen. Dabei weisen die Ladungsträger eine ausreichend große thermische Energie auf, um die an der Metall-Halbleiter-Grenzschicht existierende Energiebarriere $q \cdot V_B$ zu überwinden und in das Metall überzugehen. Bei einer sehr hohen und oberflächennahen Dotierung des Halbleiters rückt das Fermi-Niveau näher an das Valenz- bzw. Leitungsband heran und verringert so die Schottky-Barriere Φ_B . Zusätzlich reduziert sich die Weite der Raumladungszone, so daß ein quantenmechanischer Tunnelprozeß durch die schmale Barriere den Stromfluß beherrscht. Moderate Dotierstoffkonzentrationen im Halbleiter führen zu einer Kombination aus thermionischer Emission und Tunnelströmen. Dabei ist die Energie der Ladungsträger nicht ausreichend, um die Barriere zu überwinden, sie ermöglicht jedoch ein Durchtunneln oberhalb des Fußpunktes bei geringerer Barrierenhöhe und -breite. Dieser Mischfall des Stromflusses wird als thermionische Feldemission bezeichnet.

Im Fall des n-Kanal-SiGe-HFETs trifft die Metallelektrode des Gates auf eine undotierte, epitaktisch abgeschiedene Si-Cap-Schicht, die zusätzlich durch undotiertes SiGe von der vorderen Dotierschicht getrennt ist. Bei ausreichendem Abstand zwischen Schottky-Übergang und Vorderseitendotierung ist der vorherrschende Strommechanismus die thermionische Emission (Gleichung 6.7) [6.53], [6.54].

$$I_{\text{Schottky}} = A^* \cdot T^2 \cdot \exp\left(-\frac{q \cdot \Phi_B}{k \cdot T}\right) \cdot \left[\exp\left(\frac{q \cdot V_D}{k \cdot T}\right) - 1\right] \quad (6.7)$$

Wichtige Gemeinsamkeit aller Ladungsträgertransportprozesse ist die exponentielle Abhängigkeit des Sperrstromes von der Schottky-Barriere Φ_B . Deshalb muß für den Schottky-Übergang der Gateelektrode eine möglichst hohe Barriere angestrebt werden, um zum einen den Arbeitsbereich und damit auch das Leistungspotential des Transistors so groß wie möglich zu halten und zum anderen eine leistungsarme Steuerung zu erzielen.

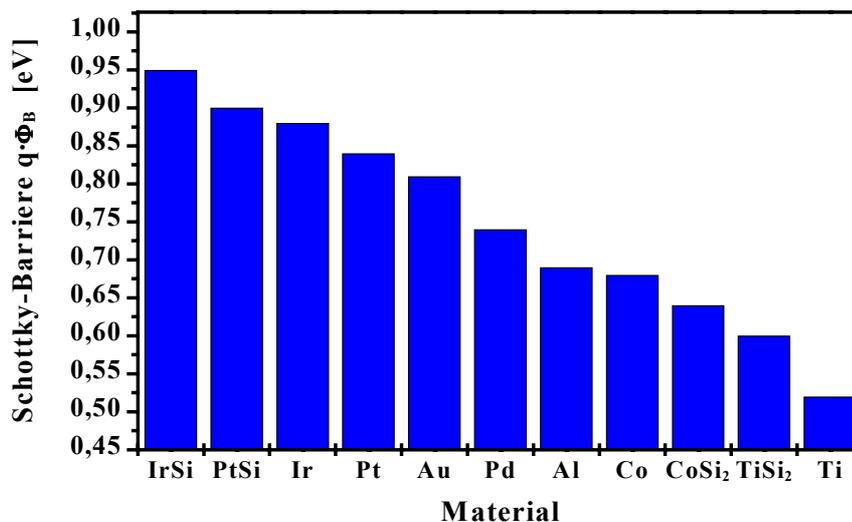


Abbildung 6.25: Schottky-Barrieren verschiedener Metall und Silizide auf n-Si

Die in Abbildung 6.25 dargestellten Schottky-Barrieren verschiedener Metalle und Silizide zeigen deutlich, daß Pt und Ir oder ihre Silizide die beste Wahl für einen gleichrichtenden Metall-Halbleiter-Kontakt auf n-Silizium sind [6.55]. In Hinblick auf eine nachfolgende Integration des Einzelbauelementes in eine Schaltung oder den Betrieb des Transistors bei erhöhten Temperaturen ist zusätzlich die elektrische und thermische Stabilität des Schottky-Kontaktes von

erhöhtem Interesse. Zur Untersuchung dieses Sachverhaltes sind mesaisolierte Diodenstrukturen mit Pt- und Ir-Schottky-Übergängen hergestellt und verglichen worden. Der Aufbau der Diode besteht aus einem niederohmigen n⁺-Si-Substrat mit einer epitaktisch aufgebracht, intrinsischen Si-Deckschicht definierter Dicke. Die Bauelementstruktur ist komplett mit SiO₂ passiviert und weist eine effektive Diodenfläche von 30 · 30 μm² auf, die über eine Metallisierung aus WTi/Al kontaktiert wird. Eine Temperung dieser Schottky-Kontakte für 5 min bei Temperaturen zwischen 100 °C und 300 °C weist eine deutliche Zunahme der Sperrströme mit der Temperaturbelastung auf (Abbildung 6.26). Während sich für den Platinkontakt ein Anstieg des Stromes in Sperrichtung erst bei 200 °C einstellt, zeigt die Iridiumdiode bereits bei 150 °C deutliche Qualitätseinbußen. Die vollkommene Degradation der Diodeneigenschaften setzt bei beiden Materialien bei Temperaturbelastungen über 250 °C ein.

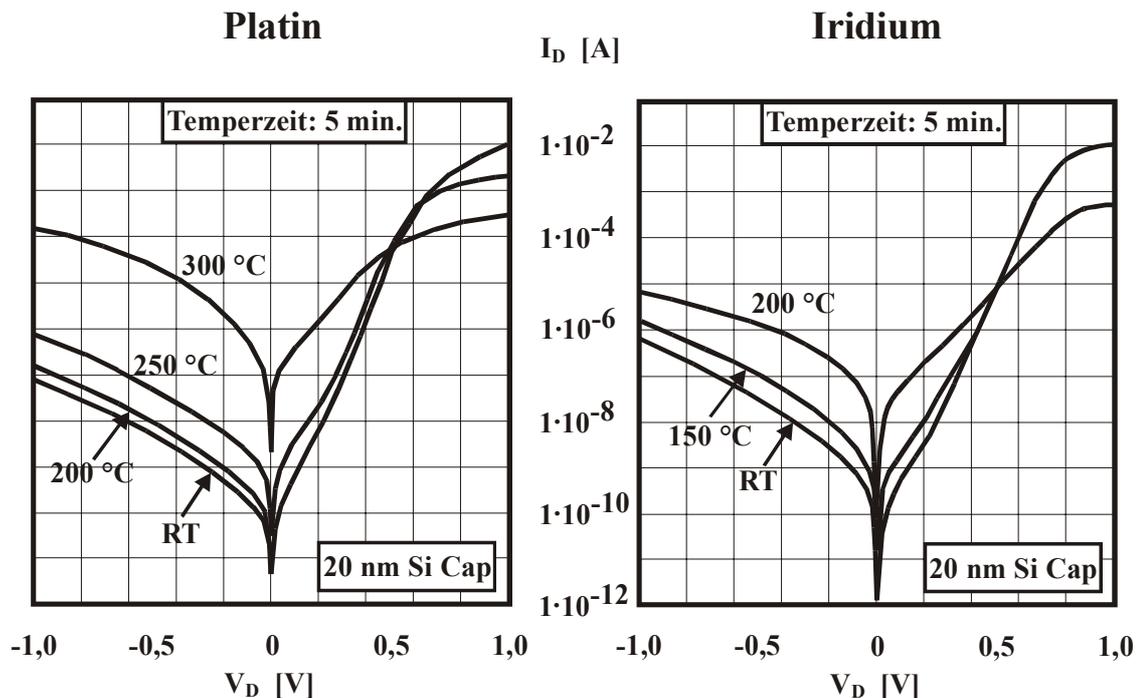


Abbildung 6.26: Einflüsse verschiedener Temperaturbelastungen zwischen RT und 300 °C auf die Strom-Spannungskennlinie von Pt- und Ir-Schottky-Dioden

Die physikalische Erklärung dieser mangelnden Temperaturstabilität liegt in der niedrigen Reaktionstemperatur zum Si. Sowohl Platin als auch Iridium bilden bereits bei Temperaturen oberhalb von 200 °C Silizide an der Grenzschicht zum

Halbleiter. Dadurch reduziert sich die dünne Si-Cap-Schicht, und es treten verstärkt Tunnelströme zwischen der Metallelektrode und dem hochdotierten Substrat auf. Bei zunehmender Temperaturbelastung wird die intrinsische Si-Schicht komplett silizidiert und es entsteht ein Übergang mit ohmschen Charakter.

Für Platin ist bereits bei Raumtemperatur die spontane Entstehung einer silizidartigen Verbindung nachgewiesen worden, die sich durch Interdiffusion von Pt und Si im Grenzbereich der beiden Materialien ausbreitet [6.56], [6.57]. Bei Temperaturen über 200 °C steigt die Diffusion von Pt in das Silizium stark an und führt zur Bildung der Pt₂Si-Phase. Diese geht nachfolgend unter weiterer Energiezufuhr in die stabile PtSi-Komponente über und bewirkt dabei einen Anstieg der Schottky-Barriere auf 0,90 eV [6.58], [6.59]. Der Siliziumverbrauch des Reaktionsprozesses liegt für 1 nm Pt bei 1,33 nm Si und resultiert in etwa 2 nm PtSi. Die Silizidierung des Iridiums beginnt bei ca. 200 °C und durchläuft in Abhängigkeit der Temperatur drei binäre Phasen: IrSi, Ir₃Si₅ und IrSi₃ [6.60]. Dabei haben die erste und letzte Verbindung metallischen Charakter, während die zweite Phase einen Halbleiter darstellt. Im Gegensatz zum Pt können beim Ir alle auftretenden Phasen parallel existieren, jedoch bildet sich keine stabile Einzelkomponente, was dem Silizid eine instabile Eigenschaft verleiht.

Für die Integration des SiGe-HFETs in eine Schaltungstechnologie ist die demonstrierte thermische Instabilität des Schottky-Überganges nicht zu tolerieren. Deshalb soll im folgenden der kontrollierte Einsatz von PtSi als Barrierenmaterial untersucht werden. Die Motivation für dieses Vorgehen liegt in der Temperaturstabilität des Silizides bis zum eutektischen Punkt von 830 °C. Zusätzlich liefert der Silizidierungsprozeß einen Anstieg der Schottky-Barriere und zeichnet sich durch einen niedrigen Siliziumverbrauch aus. Um die Silizidierung auf die limitierte Dicke des Si-Caps von 5 - 6 nm zu beschränken und übermäßige Leckströme des Gates durch Tunnelprozesse zu vermeiden, werden zwei unterschiedliche Prozeßstrategien verfolgt.

Die erste Methode verfolgt die Silizidierung sehr dünner Platinschichten, um den Siliziumkonsum so weit wie möglich zu beschränken. Dabei kommen Platindicken von 0,5 - 2,5 nm zum Einsatz, deren resultierende Schottky-Barrieren wieder mit Hilfe von Diodenstrukturen ermittelt werden. Beim Aufbringen der Platinschicht im Elektronenstrahlverdampfer hat es sich dabei als vorteilhaft herausgestellt, das Platin direkt mit einer etwa 50 nm dicken Deckschicht aus Titan zu überziehen. Zum einen ergibt sich dadurch beim Rückspütern ein niedriger

Kontaktwiderstand zur nachfolgenden Metallisierung, zum anderen steigt die Reproduzierbarkeit der ermittelten Barrieren. Die Silizidierung des Platinfilms erfolgt in Stickstoffatmosphäre bei Temperaturen über 400 °C. Die dabei ermittelten Werte für die Schottky-Barrieren zeigen für Platindicken unter 2 nm einen proportionalen Anstieg mit der Schichtdicke und eine starke Streuung über die Waferfläche (Abbildung 6.27). Dieses Ergebnis läßt darauf schließen, daß sich beim Aufdampfen keine geschlossene Platindecke ergibt, sondern eine verstärkte Kornbildung des Materials in den Vordergrund tritt. Die daraus resultierenden Barrieren stellen eine flächenanteilige Kombination aus Platin (0,84 eV) und Titan (0,52 eV) dar und weisen eine große Streuung entsprechend des statistischen Prozesses der Agglomeration auf. Bei Platindicken ≥ 2 nm stabilisiert sich die Schottky-Barriere bei etwa 0,81 eV und zeigt darüber hinaus eine verringerte Streuung über den Wafer. Für die reproduzierbare Herstellung stabiler PtSi/Si-Schottky-Übergänge ist deshalb eine Mindestdicke von 2,5 nm Platin anzustreben, um eine stabile Barriere über 0,8 eV zu gewährleisten.

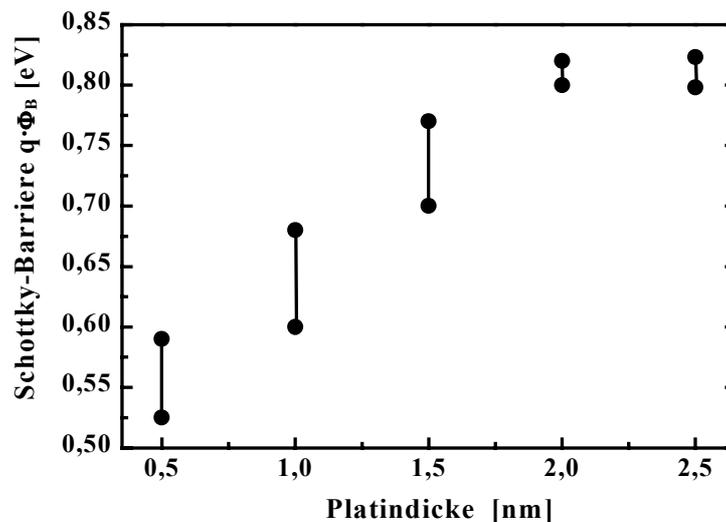


Abbildung 6.27: Abhängigkeit der Schottky-Barriere von der aufgetragenen Platindicke. Die dargestellten Balken geben die Streuung der Werte über die Fläche eines halben 4-Zoll Wafers wieder

Die zweite Methode versucht, auf die Silizidierung des Si-Caps vollkommen zu verzichten und nutzt statt dessen eine zusätzlich aufgetragene α -Silizium-Schicht, um den Materialkonsum bei der Silizidbildung zu decken. Dabei wird vorausgesetzt, daß der primär von Diffusion dominierte Silizidierungsprozeß bevorzugt

im amorphen Silizium stattfindet, da dort der Diffusionskoeffizient höher ist, als in der kristallinen Cap-Schicht. Das entscheidende Kriterium für den Einsatz dieses Verfahrens ist die endgültige Lage der PtSi/Si-Grenzfläche. Dabei muß durch präzise Kontrolle der Schichtdickenverhältnisse während des Aufdampfens eine maximale Annäherung der Silizidfront an das Si-Cap erzielt werden, ohne daß ein übermäßiger Materialkonsum an der Oberfläche der Heterostruktur stattfindet. Durch die Analyse von Diodenstrukturen mit Hilfe von Rasterelektronenmikroskopuntersuchungen und Schichtdickenmessungen läßt sich der anteilige Siliziumverbrauch innerhalb der Mehrschichtanordnung ermitteln und die Silizidgrenzschicht lokalisieren. Dabei ergibt sich ein Verhältnis von 1:2 zwischen kristallinem und amorphem Siliziumkonsum während der Silizidierung. Für den Fall a) in Abbildung 6.28 bedeutet dies, daß zur PtSi-Bildung 40 nm des α -Siliziums und 20 nm der kristallinen Si-Deckschicht verbraucht werden. Um den Verbrauch des Si-Caps weiter zu reduzieren, bietet sich die Möglichkeit an, die Platinschicht in zwei α -Silizium-Schichten einzubetten und die untere in der Dicke so abzustimmen, daß der Silizidierungsprozeß direkt auf dem kristallinen Si stoppt (Abbildung 6.28 b).

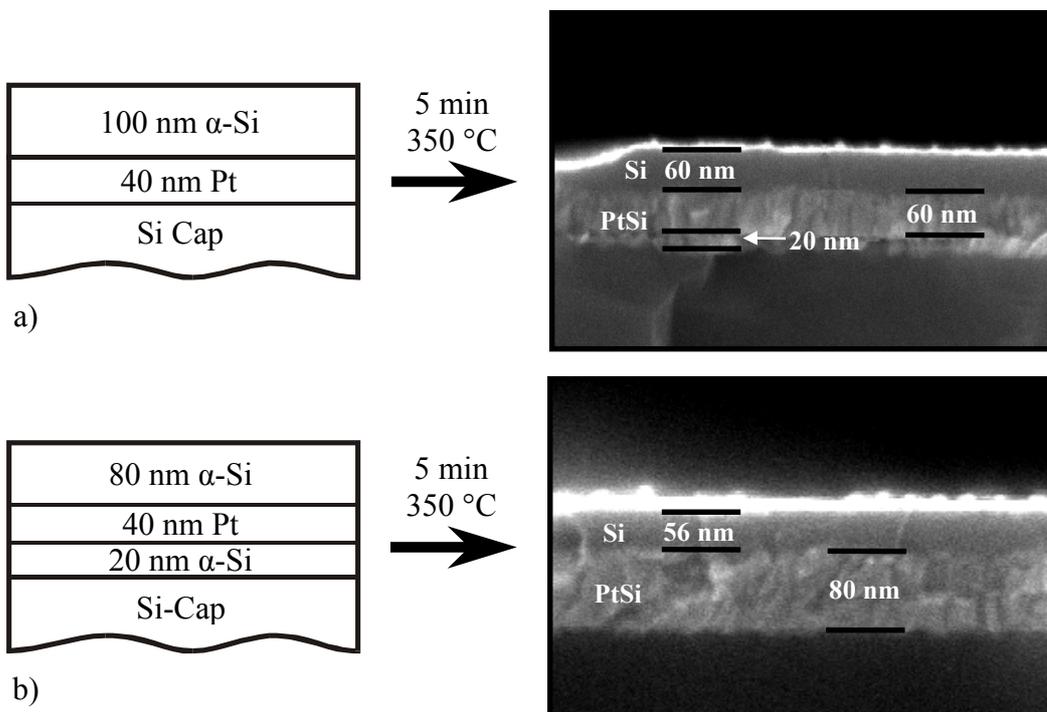


Abbildung 6.28: Ergebnisse der Schichtdickenanalyse nach der Silizidierung von a) Pt/ α -Si und b) α -Si/Pt/ α -Si-Mehrschichtsystemen zur Erzeugung eines temperaturstabilen Schottky-Überganges auf PtSi/Si-Basis

Dieses Verfahren konnte an realen Strukturen erfolgreich praktiziert werden und erzielte reproduzierbar Schottky-Barrieren von 0,82 eV. Aufgrund eines Überschusses an α -Silizium nach der Silizidierung erfordert das Aufbringen der nachfolgenden Gatematerialien zwangsläufig einen Rücksputterprozeß, um den Kontaktwiderstand zum PtSi zu minimieren. Dieses Vorgehen ist nicht zu vereinbaren mit der Lift-off-Technik der e-Beam-Lithographie, da das Lackprofil durch die Einwirkungen des Ionenbeschusses stark deformiert würde. Deshalb eignet sich die zweite Methode zur Herstellung des Schottky-Kontaktes nur für das selbstjustierte Integrationskonzept, das sich durch die vordefinierte Gatestruktur eigenständig organisiert.

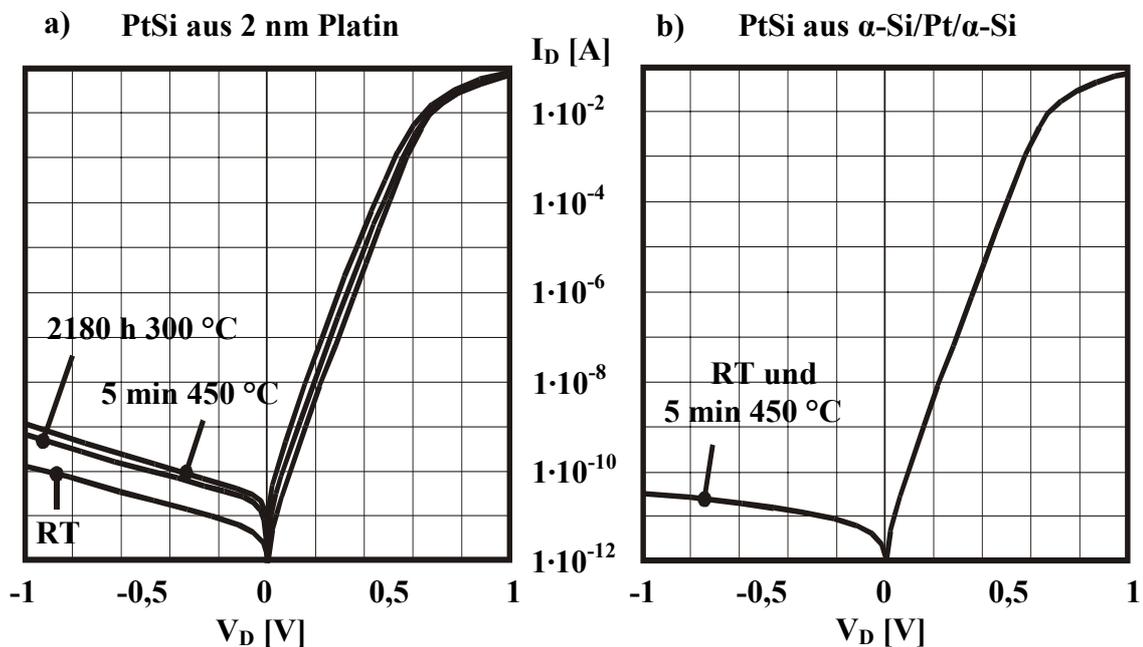


Abbildung 6.29: Temperaturtests an PtSi-Diodenstrukturen, realisiert durch die gezielte Silizidierung einer dünnen 2 nm dicken Platinschicht a) und die Anwendung einer α -Si/Pt/ α -Si-Schichtstruktur b)

Die mit den beiden entwickelten Methoden prozessierten Schottky-Übergänge auf PtSi/Si-Basis sind einer Reihe von Tempertests zur Ermittlung der Temperaturstabilität unterzogen worden. Die maximale Temperaturbelastung limitiert sich dabei wegen der Verwendung einer Aluminiummetallisierung auf 450 °C. Die Ergebnisse dieser Untersuchungen in Abbildung 6.29 zeigen eine deutlich verbesserte elektrische und thermische Stabilität der Dioden selbst bei 450 °C. Zwar ergibt sich bei der Silizidierung sehr dünner Platinschichten ein leichter

Anstieg des Sperrstromes bei Temperaturbelastungen von 450 °C oder bei Langzeittests, jedoch liegen diese Schwankungen im marginalen Bereich und sind im Rahmen des geplanten Einsatzgebietes als Gateelektrode tolerierbar.

Zusammenfassend ist festzuhalten, daß Schottky-Übergänge auf Basis der Materialkombination Silizium/Platin sich durch eine hohe Barriere auszeichnen, jedoch aufgrund von Silizidierungsmechanismen, die schon bei niedrigen Prozeßtemperaturen einsetzen, eine geringe elektrische und thermische Stabilität mit sich bringen. Durch die Entwicklung zweier neuer Herstellungsmethoden, die beide die verbesserten Materialeigenschaften von PtSi nutzen, konnte sowohl die Schottky-Barriere als auch die Temperaturbelastbarkeit des Metall-Halbleiter-Überganges entscheidend verbessert werden. Ein tolerierbares Temperaturbudget von bis zu 450 °C eröffnet nicht nur ein breiteres Anwendungsspektrum des Transistors, sondern erhöht gleichzeitig die Integrationsfähigkeit des Einzelbauelementes in ein Schaltungskonzept.

6.5.2 Gateherstellung mit Hilfe der e-Beam-Lithografie

Voraussetzung für die Herstellung von Halbleiterstrukturen im Nanometerbereich ist die Verfügbarkeit eines hochauflösenden Strukturierungsverfahrens. Während für die präzise Dimensionierung der Schichtsysteme in vertikaler Richtung CVD-, PVD- und Epitaxieverfahren zur Verfügung stehen, ist die Lithografie die Schlüsseltechnologie für die exakte Strukturübertragung in lateraler Richtung. Durch die permanent fortschreitende Skalierung der minimalen Abmessungen in der Mikroelektronik wachsen auch die Ansprüche an die verwendeten Lithografieverfahren. Für die Herstellung mesoskopischer Strukturen, wie z. B. der T-förmigen Gatekonstruktion des HFETs, liegt die geforderte Auflösung im Bereich einiger 10 nm. Dadurch ist die Auswahl potentieller Verfahren stark eingeschränkt. Mit einer Auflösungsgrenze der lichteoptischen Lithografie, die etwa im Bereich der verwendeten Photonenwellenlänge liegt, wird die Forderung nach attraktiven Alternativen deutlich. Neben der unausgereiften Röntgen- und Ionenstrahlithografie stellt der fokussierte Elektronenstrahl mit einem Durchmesser von ca. 10 nm derzeit das Schreibwerkzeug mit der höchsten Auflösung dar. Die wesentlichen Vorteile der e-Beam-Lithografie sind die große Flexibilität und die maskenlose Übertragung beliebiger Muster. Aufgrund der Sequentialität des Verfahrens müssen diese Vorzüge jedoch mit einer hohen Schreibzeit erkauft werden.

Im Rahmen des Prozeßablaufes des konventionellen Integrationskonzeptes stellt die Elektronenstrahlolithografie ein effektives Werkzeug dar, um auf die zentralen Bauelementeigenschaften des SiGe-HFETs Einfluß zu nehmen. Aufgrund der hohen Justiergenauigkeit des Elektronenstrahls von ± 20 nm kann der Gatefuß [6.61], nur begrenzt durch den Überhang des Gatekopfes, sehr nah an die source-seitige Implantationszone herangeführt werden. Dadurch verkürzt sich der ungesteuerte Kanalbereich zwischen Gateelektrode und Sourcegebiet und reduziert gleichzeitig R_s , was eine positive Wirkung auf die Steilheit des Transistors und seine Hochfrequenzeigenschaften hat. Weitere Bedingungen für ein großes Leistungspotential des Bauelementes bei hohen Frequenzen sind eine minimale Gatelänge L_G und ein niedriger Gatewiderstand R_g . Die Umsetzung einer T-förmigen Gatekonstruktion entkoppelt diese gegenläufigen Forderungen an die geometrische Form der Steuerelektrode. Dabei wird durch einen schmalen Gatefuß ein kleines L_G erzielt und gleichzeitig durch einen großflächigen Kopf der Widerstand gering gehalten.

Die technologischen Probleme bei der Realisierung des T-Gatekonzeptes werden mit Hilfe einer 3-Lagen-Lackstruktur im Rahmen der Elektronenstrahlolithografie gelöst. Dabei ermöglicht die Kombination aus Lackebenen unterschiedlicher Strahlungssensitivität und die „Belichtung“ des Kopf- und Fußbereiches des Gates mit verschiedenen Elektronendosen die Definition eines präzisen T-förmigen Lackprofils. Als „Arbeitspferd“ der hochauflösenden Elektronenstrahlolithografie hat sich im Laufe der Zeit Polymethylmethacrylat (PMMA) als Positivresist bewährt. Er zeichnet sich durch eine hohe Auflösung < 5 nm und eine große Molekülmasse $M_W \approx 10^6$ g/mol aus [6.62]. Über die molekularen Eigenschaften können sowohl die Steilheit des Hell/Dunkel-Überganges als auch die Empfindlichkeit des Fotolackes bestimmt werden. Da mit zunehmender Molekülmasse M_W bei gleicher Dichte ρ die Volumendichte der zu fragmentierenden Molekülketten mit $1/M_W$ abnimmt, ist nur eine reduzierte Bestrahlungsenergie E_{Strahl} notwendig, um die chemischen Eigenschaften des Lackes zu verändern. Damit läßt sich über unterschiedliche Molekülmassen die Entwicklungselektivität der einzelnen Lackebenen bei gleicher Bestrahlungsdosis einstellen [6.63].

Durch den Beschuß der Lack- und Halbleiteroberfläche mit Elektronen hoher kinetischer Energie treten atomare Streuprozesse in den Materialien auf. Die möglichen Wechselwirkungen lassen sich dabei in zwei Klassen aufteilen. Zum einen finden elastische Streuungen an Atomkernen mit der ausschließlichen Konsequenz einer Richtungsänderung statt. Zum anderen werden bei unelastischen Streuprozessen Sekundärelektronen mit typischen Energien von einigen 10 eV

ausgelöst. Diese niederenergetischen Elektronen sind es schließlich, welche für die chemische Veränderung der Lackstruktur verantwortlich sind. Als Konsequenz für das resultierende Lackprofil zeigt sich eine Aufweitung des Primärelektronenstrahls im Fotoresist aufgrund der Vorwärtsstreuung und eine Rückstreuung der Ladungsträger in den Lackfilm durch elastische Wechselwirkungen im Substrat. Dieser Mechanismus führt zu einer unerwünschten Bestrahlung des Lackes durch die sogenannte Hintergrunddosis (Proximity-Effekt) [6.64]. Monte-Carlo-Simulationen der Elektronenstreuung im Lack-Halbleitersystem haben gezeigt, daß bei erhöhter Elektronenenergie die Vorwärtsstreuung und damit die Aufweitung des Primärstrahles drastisch verringert werden kann, gleichzeitig läßt sich die maximal erzielbare Stromdichte des Strahles proportional mit der Elektronenenergie erhöhen [6.65]. Die Effizienz der Elektronenrückstreuung ist über die Variation der Primärstrahlenergie nicht zu beeinflussen, da sie durch das Halbleitermaterial festgelegt wird, jedoch erlaubt die Erhöhung der Elektronenenergie die Verteilung des Rückstreubeitrages auf eine größere Resistfläche. Aus diesen Gründen erfolgte die Entwicklung des Lithografieprozesses auf der Basis einer Primärelektronenenergie von 50 keV.

Die theoretische und praktische Entwicklung des verwendeten Elektronenstrahl-lithografieprozesses erfolgte in Zusammenarbeit mit dem Fraunhofer Institut für „Angewandte Festkörperphysik“ in Freiburg [6.66]. Dabei mußte das für III/V-Halbleiter qualifizierte Herstellungs-konzept auf die spezifischen Eigenschaften des SiGe-Materialsystems und die Technologie des HFETs adaptiert werden. Der Prozeßablauf beginnt mit einer gründlichen Reinigung der Halbleiteroberfläche mit Isopropanol. Zur Verbesserung der Hafteigenschaften des Lackes kommt ein Haftvermittler (HMDS) zum Einsatz. Als erste Lackschicht, die später zur Definition der Gatelänge dienen soll, wird ein 100 nm dicker PMMA (950 K) mit $M_w = 950000$ g/mol aufgebracht. Danach folgt ein Copolymerlack (PMAA) von 350 nm Dicke, der sich durch eine etwa 20-mal höhere Empfindlichkeit zum PMMA unterscheidet. Den Abschluß des 3-Lagen-Lacksystems bildet wieder ein Polymethylmethacrylat, diesmal jedoch mit einer Molekülmasse von 50000 g/mol bei einer Schichtdicke von 80 nm.

Für den Belichtungsprozeß der Lackstruktur kommt eine Elektronenstrahlanlage der Firma Leica vom Typ EBP-5 HR zum Einsatz. Die Zentralbelichtung des Gatefusses erfolgt gepulst mit einer Dosis von ca. $1600 \mu\text{C}/\text{cm}^2$ und einem Strahldurchmesser von 30 - 40 nm. Die Schreibfrequenz liegt bei der gewählten Dosis und Auflösung bei etwa 400 kHz. Um eine vollständige Belichtung des Lackes zu gewährleisten, werden zwei parallele Fußlinien geschrieben, deren

Elektronenpulse in vertikaler und lateraler Richtung einen Abstand von 10 nm aufweisen. Durch die hohe Elektronendosis kommt es zu einer Überbelichtung des Fußlackes, was zu einer resultierenden Breite des Lackgrabens nach der Entwicklung von 90 -110 nm führt [6.66]. Aufgrund der vielfach höheren Empfindlichkeit des Copolymers reduziert sich die Elektronendosis bei der Belichtung der beiden Flügellinien auf ca. $500 \mu\text{C}/\text{cm}^2$. Der dabei verwendete Strahldurchmesser von 60 - 80 nm erlaubt eine Erhöhung der Elektronenpulsschrittweite auf 50 nm. Die stark unterschiedliche Elektronensensitivität von PMMA und Copolymerlack verhindert bei der zweiten Bestrahlung eine weitere Aufweitung des bereits belichteten Profils im Fußlack und sorgt für eine selektive Fragmentierung des Kopflackes. Die Übergänge der schmalen Gatestege auf die großen Kontaktflächen der ohmschen Metallisierung werden durch dreieckige Anschlußpads realisiert. Um die Schreibzeit dieser Flächen in Grenzen zu halten, kommt ein Strahldurchmesser von 100 nm zum Einsatz, der in Verbindung mit einer verringerten Belichtungs-dosis eine hohe Schreibfrequenz von 8 MHz zuläßt.

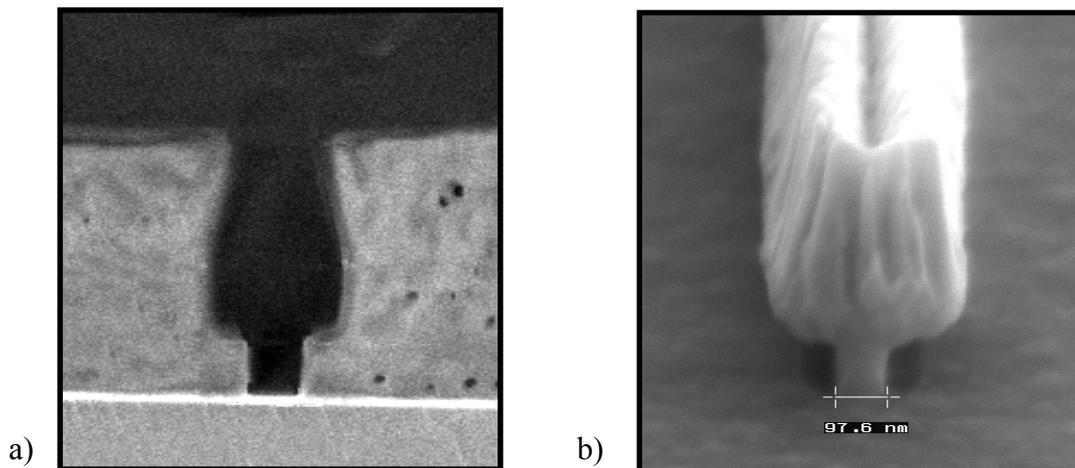


Abbildung 6.30: a) 3-Lagen-Lackprofil nach der Elektronenstrahlolithografie und dem Entwicklungsprozeß; b) T-Gate mit 100 nm Gatelänge

Der nachfolgende Entwicklungsprozeß in einer Lösung aus Methylisobutylketon (MIBK) und Isopropanol im Verhältnis 1:3 entfernt die durch den Elektronenstrahl chemisch veränderten Lackbereiche und hinterläßt das in Abbildung 6.30 a) dargestellte Lackprofil. Um die Qualität des Lift-off-Prozesses zu erhöhen und Gateabrisse und aufsteigende Metallfahnen zu vermeiden, ist die Ausbildung einer überhängenden Lackflanke im oberen Bereich des Gatekopfes von großer Wichtigkeit. Dafür sorgt die dünne und unempfindliche Decklackschicht

aus PMMA 50 K. Sie erfährt durch die beiden Belichtungen eine geringere Aufweitung als der Kopflack und erzeugt so ein negatives Lackprofil. Beim Auffüllen des Lackgrabens mit Metall besteht deshalb nur wenig Kontakt zwischen dem Copolymer und der Gatemetallisierung, was die Ergebnisse des Lift-offs und die T-Form des Gates verbessert (Abbildung 6.30 b). Vor dem Aufdampfen der Gatemetallisierung muß die Kontaktfläche zum Halbleiter von residualen Lackresten und dem natürlichen SiO₂ befreit werden. Dazu erfolgt zuerst ein sanfter Reinigungsschritt im O₂-Plasma und anschließend ein Bad in gepufferter 4,7 %iger Flußsäure (BHF).

Um den Gatewiderstand niedrig zu halten, sollte die nachfolgende Metallisierung so dick wie möglich ausfallen. Für den Lift-off-Prozeß hat sich eine Schichtkombination aus Pt und Au mit einer Gesamtdicke von 400 nm bewährt. Der aus Hochfrequenzmessungen extrahierte Gatewiderstand erzielt dabei einen niedrigen Wert von 70 Ω /mm für eine Gatelänge von 130 nm. Eine, wie in Abschnitt 6.5.1 dargestellte, modifizierte Prozeßführung zur Herstellung eines temperaturstabilen Schottky-Überganges zwischen Gatemetallisierung und Halbleiter ist beim konventionellen Integrationskonzept mit e-Beam-Lithografie nicht möglich. Durch den zwingend notwendigen Rückspatterschritt vor dem Aufbringen der Hauptmetallisierungsschicht würde es zu einer Aufweitung und Degradation des in Abbildung 6.30 a) präsentierten Lackprofils kommen, was die reproduzierbare Prozessierung einer präzisen T-Gatestruktur unmöglich macht.

6.5.3 Selbstjustierendes T-Gatekonzept

Die Realisierung der selbstjustierenden Gatekonstruktion zeigt komplett andere Randbedingungen als das konventionelle Herstellungskonzept. Während bei der traditionellen Strukturierungsmethode die charakteristische Gateform mit Hilfe der Elektronenstrahlithografie und einer Dreilagelacktechnologie mit anschließendem Lift-off-Prozeß umgesetzt wird, steht durch den Einsatz des Replacement-Gates beim selbstjustierenden Fertigungsablauf bereits eine vorgefertigte Struktur zur Implementierung des Gatefusses zur Verfügung. Die geometrischen Dimensionen dieser Anordnung sind im Laufe des Technologieablaufes durch verschiedene Prozeßmodule unabhängig voneinander festgelegt worden. Die Gatelänge resultiert dabei aus dem Grad der Skalierung des Dummy-Gates, und die Höhe des Gatefusses läßt sich durch die Rückätzung des Oxides bei der Freilegung des Replacement-Gates einstellen. Die Nutzung einer festen vordefinierten Struktur bietet darüber hinaus den Vorteil einer erhöhten mechanischen Sta-

bilität. Dadurch ergibt sich die Möglichkeit, Gatefuß und -kopf im Rahmen physikalisch sinnvoller Grenzen getrennt voneinander zu dimensionieren, ohne an statische Restriktionen gebunden zu sein, wie es bei einer freistehenden Gatekonstruktion der Fall ist. Das Problem der Gateherstellung zerfällt damit in zwei unabhängige Teilprozesse, zum einen die Realisierung des Gatefusses und zum anderen die Strukturierung des Gatekopfes. Durch die Entkopplung dieser Prozesse eröffnet sich dem Entwickler eine höhere Anzahl von Freiheitsgraden bei der Umsetzung eines optimierten Bauelementkonzeptes.

Die Basis für ein leistungsfähiges Metall/Halbleiter-Gatekonzept ist ein thermisch und elektrisch stabiler Schottky-Übergang mit möglichst hoher Barriere. Die physikalischen Grundlagen und technologischen Prozeßabläufe dafür wurden bereits in Abschnitt 6.5.1 erarbeitet, so daß nun die Adaption auf die reale Bauelementstruktur im Mittelpunkt steht. Die Randbedingungen für diese Aufgabe ergeben sich aus den geometrischen Abmessungen des Gategrabens, der durch das Replacement-Gate aus Polyimid seine Form erhalten hat und an dessen Fuß nun der Schottky-Übergang zur Si-Cap-Schicht der Heterostruktur entstehen soll. Da die freistehende, elektronenstrahlgeschriebene Gatestruktur des konventionellen Integrationskonzeptes eine natürliche Luftisolation aufweist, erhöhen sich für den selbstjustierenden Transistor maßgeblich die Überlappkapazitäten des Gates mit der höheren Dielektrizitätskonstante der passivierenden Oxidschicht. Um die parasitären Einflüsse dieser Kapazitäten auf das Hochfrequenzverhalten des Bauelementes zu mindern, muß die Isolatorstärke, bei gleicher Gatekopfbreite, etwa den vierfachen Wert der nominellen Gatefußhöhe der freistehenden Konstruktion einnehmen. Dadurch ergeben sich für den zu füllenden Gategraben bei einer Gatelänge von etwa 100 nm Aspektverhältnisse von bis zu 6:1. Die technologischen Anforderungen an die Realisierung des Schottky-Überganges und die Füllung des Gatefusses wachsen damit extrem an.

Der folgende Abschnitt beschreibt zuerst die technologischen Maßnahmen zur Umsetzung des Metall-Halbleiter-Überganges, bevor anschließend auf die homogene Füllung des Gatefusses eingegangen wird. Primäres Ziel bei der Herstellung des Schottky-Überganges ist die gleichmäßige Belegung der Siliziumoberfläche am Fuß des Gategrabens mit der geforderten Metallkombination. Dabei müssen jedoch die strukturellen und morphologischen Voraussetzungen zur nachfolgenden Füllung des Grabens erhalten bzw. verbessert werden. Mit Hilfe des Aufdampfverfahrens lassen sich, aufgrund des geringen Prozeßdruckes von $< 10^{-7}$ Torr und der relativ großen Distanz zwischen Materialquelle und Wafer im Rezipienten, geradlinige Materialströme mit geringen Einfallswinkeln erzeugen.

Die daraus resultierende schlechte Konformität ist für diesen Anwendungsfall ausgesprochen erwünscht, um bevorzugt Metallablagerungen am Boden des Grabens zu erzielen und an den Seitenwänden zu vermeiden und dadurch das Aspektverhältnis zu reduzieren. Die Ausführung der Transistorpassivierung und damit des Gategrabens aus SiO_2 begünstigt dabei dieses Ziel und ermöglicht zusätzlich eine selbstjustierende Plazierung des PtSi/Si-Schottky-Überganges am Fuße des Grabens. Aufgrund des edlen Charakters des Platins besitzt es, ebenso wie Gold, eine schlechte Haftung auf Isolierschichten wie Si_3N_4 und SiO_2 [6.67], [6.68]. Deshalb erfolgt während der Abscheidung keine Materialanlagerung an den Seitenwänden des Grabens, sondern nur an horizontalen Oberflächen.

Nach dem Aufdampfprozeß löst sich das Platin ohne Einwirkung einer Ätzmischung von der SiO_2 -Oberfläche und bildet nur zum Silizium der Cap-Schicht eine feste Verbindung. Voraussetzung für die Nutzung dieser Eigenschaft ist, daß Pt die unterste Schicht bei der Materialkombination für den Schottky-Übergang bildet. Deshalb wurde in Verbindung mit dem selbstjustierenden Integrationskonzept primär die Silizidierung dünner Platinschichten eingesetzt. Die zur Abdeckung des Platins eingesetzte Ti-Schicht weist nicht diese vorteilhaften Eigenschaften auf und zeigt dadurch auch eine verstärkte Anlagerung an den Seitenwänden und an der oberen Öffnung des Oxidgrabens. Aufgrund der geringeren Masse des Titans führen die Kollisionen mit den Restteilchen im Vakuum zu einer größeren Abweichung des Materialstromes von der Substratnormalen als beim Platin. Dadurch fächert sich der Fluß der Ti-Atome weiter auf und erhöht durch seine Haftung an den vertikalen Flächen auch das Aspektverhältnis des Grabens. Deshalb sollte die Ti-Deckschicht auf dem Pt-Film so dünn wie möglich sein und gegebenenfalls durch eine Kombination aus Ti und Pt ersetzt werden. Die 10 - 20 nm dicke Ti-Zwischenschicht dient dabei als Diffusionsbarriere für das Pt. Der untere Pt-Film steht mit seiner definierten Dicke von etwa 2,5 nm zur kontrollierten Silizidierung mit dem Si-Cap zur Verfügung, und die obere Lage von 50 nm wirkt als Opferschicht für den nachfolgenden Rückspalterprozeß. Um das Aspektverhältnis für die nachfolgende Grabenfüllung günstiger zu gestalten, kann diese Pt-Deckschicht auch dicker ausgeführt werden.

Zur Füllung von Kontaktlöchern und Gräben mit hohem Aspektverhältnis haben sich in der Halbleitertechnologie verschiedene Verfahren durchgesetzt. Eine Standardmethode ist die Wolframabscheidung aus der Gasphase [6.69]. Diese Technik ist aufgrund ihrer exzellenten Stufenbedeckung in der Lage, Aspektverhältnisse von bis zu 8:1 homogen zu füllen. Anwendung findet sie in Kombination mit der Aluminium- oder Kupfer-Mehrlagenverdrahtung für hochintegrierte

Logik-ICs [6.70]. In Verbindung mit dem Einsatz von Kupferleitbahnen hat in den letzten Jahren die galvanische Abscheidung zunehmend an Bedeutung gewonnen. Sie ermöglicht zum einen die Umsetzung hoher Depositionsraten für große Metallisierungsdicken und zum anderen die Realisierung einer Dual-Damascene-Technik mit Aspektverhältnissen von bis zu 9:1 [6.71], [6.72]. In Zusammenhang mit Al- und Cu-Metallisierungen kommt zum Teil auch das sogenannte „Reflow-Verfahren“ zum Einsatz [6.73]. Dabei findet ebenfalls eine konforme Metallabscheidung statt, die jedoch nicht zu einer kompletten Füllung des Kontaktloches führt, deshalb wird mit Hilfe von thermischer Energie das Metall in einen viskosen Zustand versetzt und teilweise unter Einwirkung von äußerem Überdruck in die Verbindungskanäle gepreßt. Abschließend soll hier noch die gerichtete Kathodenzerstäubung zur Darstellung kommen. Bei dieser Methode führt entweder ein großer Abstand zwischen Target und Substrat, gepaart mit einem geringen Prozeßdruck, oder ein Kollimatorgitter zu einer senkrechten Ausrichtung des Materialflusses und damit zur Voraussetzung für eine homogene Füllung tiefer Gräben und Löcher [6.74].

Die Anwendung der Wolframabscheidung und der gerichteten Kathodenzerstäubung zur Realisierung des selbstjustierenden T-Gates scheitert an der mangelnden Verfügbarkeit dieser Techniken. Darüber hinaus übersteigt die „Reflow“-Methode, mit Temperaturen um 500 °C, das thermische Budget des PtSi/Si-Schottky-Überganges. Deshalb konzentriert sich die Umsetzung des Gatekonzeptes auf die Verwendung einer elektrolytischen Abscheidung, jedoch kommt anstatt Kupfer hier Gold zum Einsatz [6.75].

Die Galvanisierung erfolgt in einem zyanidischen Elektrolyt ($\text{KAu}(\text{CN})_2$) mit einem Goldgehalt von 10 g/l Flüssigkeit bei einer Prozeßtemperatur von ca. 50 °C. Die von einer Konstantstromquelle eingestellte Stromdichte J bestimmt die Abscheiderate G_R , die sich nach Gleichung (6.8) berechnet. Dabei berücksichtigen \ddot{A}_{Au} und $\rho_{\text{M,Au}}$ jeweils das elektrochemische Äquivalent und das spezifische Gewicht von Gold, und α_K gibt die kathodische Effizienz des Prozesses an.

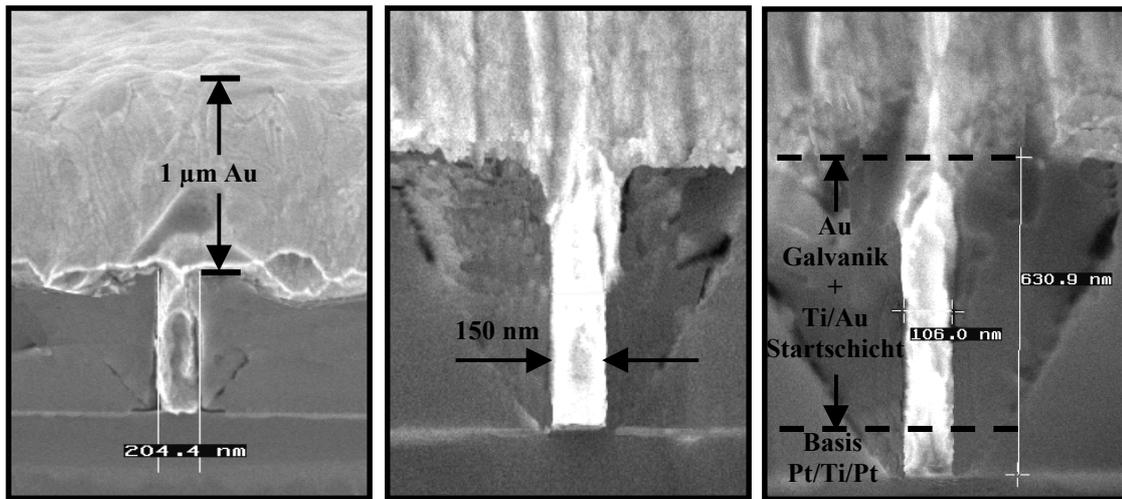
$$G_R = \frac{J \cdot \ddot{A}_{\text{Au}} \cdot \alpha_K}{\rho_{\text{M,Au}}} \quad \text{mit} \quad \ddot{A}_{\text{Au}} = 0,6812 \cdot 10^{-3} \frac{\text{g}}{\text{A} \cdot \text{s}} \quad (6.8)$$

Unter Zuführung einer nominellen Stromdichte von 3 mA pro cm^2 zu galvanisierender Fläche ergibt sich dadurch eine Abscheiderate von 1,85 nm/s. Um eine ganzflächige Deposition zu gewährleisten, muß eine möglichst konforme Start-

schicht auf den Wafer aufgebracht werden, die als Kathode dient. Dabei kommt eine Schichtkombination aus 10 nm Ti und 20 nm Au zum Einsatz. Der dünne Titanfilm dient in diesem Fall als Haftvermittler für die Goldschicht auf dem SiO₂. Aufgrund der guten Kantenbedeckung erfolgte die Abscheidung dieser Basisschicht für die nachfolgende Galvanisierung mit Hilfe der Hochfrequenz-Kathodenzerstäubung. Die elektrolytische Au-Deposition unter den oben dargestellten nominellen Bedingungen führt zu einer homogenen Füllung von Gategräben mit einem Aspektverhältnis von bis zu 4:1 (Abbildung 6.31 a).

Da die galvanisch hergestellten Schichten durch Kornwachstum entstehen, lassen sie eine gewisse Oberflächenrauigkeit erkennen. Durch eine Variation der Prozeßparameter konnte nachgewiesen werden, daß die Korngröße eine Abhängigkeit von der eingestellten Stromdichte demonstriert und sich mit geringeren Strömen die Rauigkeit des Goldes erheblich verringert. Deshalb begünstigen niedrigere Stromdichten ein homogeneres und gleichmäßigeres Wachstum des Goldes in den Oxidvertiefungen. Diese Modifikation erfolgt jedoch auf Kosten der Prozessierungsdauer, die sich antiproportional zur Stromdichte verhält. Zusätzlich muß bei der Galvanisierung von Strukturen mit extremen Aspektverhältnissen gewährleistet sein, daß der Elektrolyt gut in die tiefen und engen Gategräben eindringen kann, um sie komplett ohne Einschluß von Luftblasen zu füllen. Aus diesem Grund erfolgt die Benetzung der Waferoberfläche mit dem Goldcyanid vor der eigentlichen Abscheidung unter Vakuumbedingungen in einem Exsikkator. Durch eine Reduzierung des Druckes auf 0,1 bar wird sichergestellt, daß die Flüssigkeit auch in die kleinsten Öffnungen vordringen kann und eine kontrollierte Abscheidung des Materials, startend am Fuß des Grabens, ohne Ausbildung von Hohlräumen erfolgt. Mit dieser Methode und durch eine Verringerung der Stromdichte auf ein Viertel des nominellen Wertes konnte das maximale Aspektverhältnis der Strukturen, die einen 100 %igen Füllungsgrad aufweisen, auf 5:1 gesteigert werden (Abbildung 6.31 b).

Andere Modifikationen des elektrolytischen Depositionsprozesses führen zu keiner weiteren Erhöhung des füllbaren Aspektverhältnisses auf den gewünschten Wert von 6:1. Deshalb muß durch die Abscheidung einer verstärkten Pt-Deckschicht die Höhen/Tiefen-Konstellation des Gategrabens bereits vor der Galvanisierung auf 5:1 reduziert werden, um anschließend eine homogene Auffüllung mit Gold zu gewährleisten. Das Ergebnis dieser Kombination aus einer aufgestockten Schottky-Metallisierung aus Pt/Ti/Pt und der abschließenden elektrolytischen Au-Abscheidung zeigt Abbildung 6.31 c).



a) Aspektverhältnis: 4:1 b) Aspektverhältnis: 5:1 c) Aspektverhältnis: 6:1

Abbildung 6.31: Füllung von Gategräben verschiedener Aspektverhältnisse mit Hilfe der elektrolytischen Au-Abscheidung

Die Ausführung der galvanisch erzeugten Au-Schicht mit einer Dicke von $1\ \mu\text{m}$ ermöglicht nicht nur die Realisierung des Gatefusses, sondern führt außerdem zu einer ausreichenden Au-Deposition auf der SiO_2 -Passivierung (Abbildung 6.31 a), um direkt daraus den Gatekopf zu definieren. Dabei wird mit Hilfe einer lithografisch hergestellten Lackmaske die Breite des Kopfes festgelegt und anschließend naßchemisch mit einer geeigneten Ätzlösung die Struktur in die Goldschicht übertragen. Die Festlegung der nominellen Gatekopfbreite spielt dabei eine entscheidende Rolle für das Hochfrequenzverhalten des Transistors. Da die Überlappkapazitäten C_{gso} und C_{gdo} , zwischen den Flanken des Gatekopfes und der Halbleiteroberfläche bzw. den Ladungen im Kanal, als Teilkapazitäten in C_{gs} und C_{gd} enthalten sind, beeinflussen sie auch die Grenzfrequenzen f_T und f_{max} . Die positiven Auswirkungen eines niedrigen Gatewiderstandes auf f_{max} , durch eine große Querschnittsfläche des Kopfes, können durch eine Zunahme der Überlappkapazitäten kompensiert werden. Die Ermittlung der optimalen Gatekopfbreite für die gegebene T-Gate-Konstruktion kann mit Hilfe von Simulationen, basierend auf extrahierten Kleinsignalersatzschaltbildparametern erfolgen.

Kapitel 7

Elektrische Charakterisierung und Simulation der SiGe-HFETs

Die Modifikation bestehender Technologieprozesse und die Entwicklung neuer Integrationskonzepte läuft stets nach einem iterativen Verfahren ab. Dabei werden zunächst Einzelprozessschritte separat entworfen, getestet und qualifiziert, bevor die Einbindung bzw. Zusammenfassung in Prozeßmodule erfolgt. Die Integration der verschiedenen Prozeßmodule ergibt schließlich den kompletten Technologieablauf für das Bauelement oder die Schaltung. Auf jeder dieser Ebenen stellt die Prozeßcharakterisierung ein wichtiges und effektives Werkzeug zur Analyse, Beschreibung und Beurteilung des Prozesses und seines Endproduktes dar. Die Korrektur der Prozeßparameter auf der Grundlage der experimentell ermittelten Daten ermöglicht die sukzessive Optimierung des Herstellungsverfahrens, innerhalb eines technisch und ökonomisch sinnvollen Prozeßfensters, bis zum Erreichen der geforderten Produktspezifikation. Während auf der Entwicklungsebene der Einzelprozessschritte und Prozeßmodule chemische und physikalische Materialcharakterisierungen und optische Kontrollen im Mittelpunkt stehen (siehe Kapitel 6), erfolgt die Beurteilung des funktionsfähigen Bauelementes zumeist auf der Grundlage elektrischer Meßverfahren. Die Extraktion spezifischer Transistorkennwerten aus den ermittelten Meßergebnissen erlaubt die Bestimmung des aktuellen Leistungspotentials, ermöglicht den wissenschaftlichen Vergleich verschiedener Bauelementausführungen und stellt die Grundlage für eine nachfolgende Bauelementsimulation dar. Die Nachbildung des Bauelementverhaltens in einem geeigneten Modell bietet darüber hinaus die Möglichkeit der Variation verschiedener struktur- und technologierelevanter Faktoren und kann, abgeleitet aus den Simulationsergebnissen, eine gezielte Korrektur der Prozeßparameter nach sich ziehen. Außerdem lassen sich damit gegebenenfalls zeitaufwendige und kostspielige zusätzliche Technologiedurchläufe einsparen und die theoretischen Leistungsgrenzen des Bauelementes ermitteln.

Das folgende Kapitel analysiert und vergleicht die Ergebnisse der elektrischen Charakterisierung unterschiedlicher Bauelementversionen der beiden Integrationskonzepte. Die meßtechnische Analyse des Transistorverhaltens erfolgt dabei unter Gleichstrombedingungen, durch Streuparametermessungen bei Frequenzen bis 50 GHz, durch die Bestimmung spezifischer Rauschparameter im Hochfrequenzbereich und bei kryogenischen Temperaturen. Zusätzlich werden die bei der Hochfrequenzmessung ermittelten S-Parameter zur Modellierung eines, für den betrachteten Frequenzbereich gültigen, Kleinsignalersatzschaltbildes genutzt und die zugehörigen Elemente durch eine analytische Extraktionsroutine ermittelt.

7.1 Analyse des Gleichstromverhaltens

Die Charakterisierung des Gleichstromverhaltens der Transistoren erfolgte mit Hilfe eines halbautomatischen Spitzenmeßplatzes in Verbindung mit einem Parameteranalysator für Halbleiterbauelemente HP 4155. Der computergesteuerte Meßaufbau erlaubt unter anderem das automatische Anfahren beliebiger Bauelementpositionen auf dem Wafer und das sequentielle Aufnehmen mehrerer unterschiedlicher Meßreihen am jeweiligen Transistor. Des weiteren bietet sich die Möglichkeit der automatischen Extraktion verschiedener Gleichstromkenngrößen des Bauelementes und der statistischen Auswertung der Charakteristiken über ein Wafermapping.

Der typische Operationsbereich des SiGe-n-Kanal-HFETs erstreckt sich am Eingang, d.h. zwischen Gate und Source, von ca. -1 V bis etwa +0,7 V. Die obere Grenze der Gateaussteuerung ergibt sich dabei durch die Schottky-Barriere des Metall-Halbleiter-Überganges. Weitere Faktoren, die den Arbeitspunkt des Transistors beeinflussen, sind die Auslegung der SiGe-Heteroschichtstruktur, der Aufbau und die Abmessungen des Layouts und die jeweilige technologische Umsetzung. Abbildung 7.1 zeigt die Transfercharakteristiken von drei unterschiedlichen Transistoren, die mit dem selbstjustierenden Integrationskonzept hergestellt wurden und ein identisches Layout aufweisen. Der einzige Unterschied der Bauelemente liegt in der Auslegung der undotierten SiGe- und Si-Cap-Schichten, die mit ihrer Dicke den Abstand d_{GK} zwischen der Gateelektrode und dem Kanal festlegen. Für die Transistoren A, B und C ist jeweils epitaktisch ein nominales d_{GK} von 23,5, 19 und 16,5 nm angesetzt worden. Die Simulation der Bandstrukturen dieser Transistoren zeigt bei einer Veränderung des Gatepo-

tentials für kleinere d_{GK} eine stärkere relative Verschiebung des Leitungsbandes zum Fermi-niveau und damit einen wachsenden Einfluß auf die Ladungsträger im 2DEG. Daraus resultiert eine effektivere Modulation des Drainstromes, die durch einen steileren Anstieg im linearen Anlaufbereich der Kennlinie sichtbar wird. Gleichzeitig ergibt sich, aufgrund der Definition der Steilheit, für diese Transistoren auch ein größeres Steilheitsmaximum. Begleitet wird diese Veränderung des Bauelementverhaltens bei abnehmenden Gate-Kanal-Abständen von einer Verschiebung der Schwellspannung V_T in Richtung positiver V_{GS} -Werte. Dabei kommt es sogar zu einem Vorzeichenwechsel der Schwellspannung vom negativen Bereich für Transistor A in den positiven Bereich für Transistor C. Die Auslegung der Heteroschichtstruktur, das sogenannte „layer stack engineering“, ermöglicht somit eine gezielte Einstellung des Operationsbereiches, je nach Anwendung des Bauelementes, zwischen Verarmungs- und Anreicherungstyp.

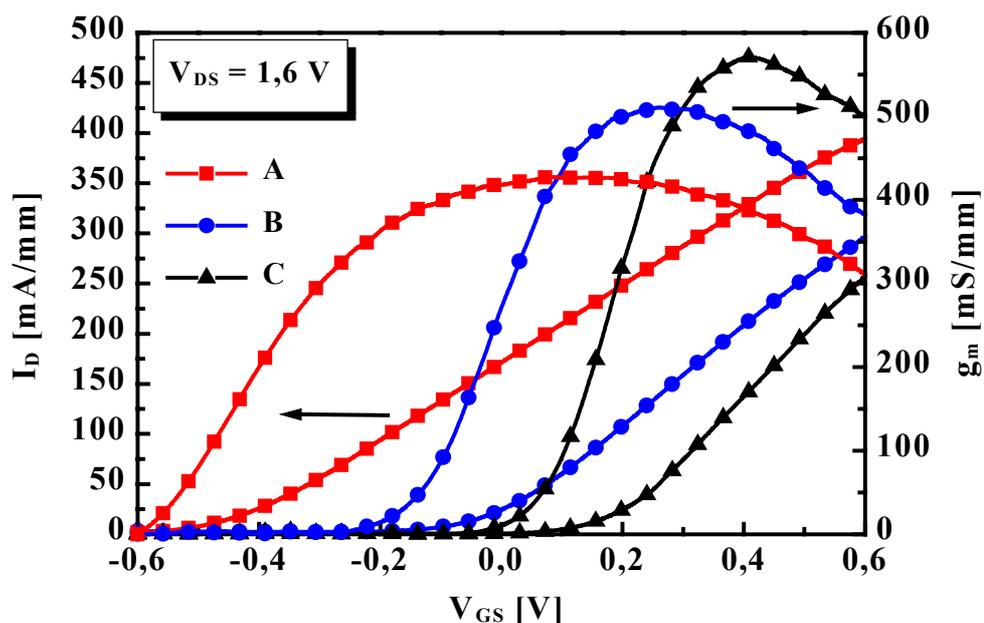


Abbildung 7.1: Transfercharakteristiken von SiGe-HFETs mit $A = 23,5$ nm, $B = 19$ nm und $C = 16,5$ nm Gate-Kanalabstand d_{GK}

Darüber hinaus ermöglicht die Transfer- oder Übertragungskennlinie weitere aufschlußreiche Analysen des Transistorverhaltens. Aus ihr lassen sich neben der Steilheit g_m und der Schwellspannung V_T zusätzliche charakteristische Kenngrößen ablesen, wie die Unterschwellsteilheit S und das I_{on}/I_{off} -Verhältnis des Drainstromes, ebenso wie das Steuer- und Abschnürverhalten des Transistors.

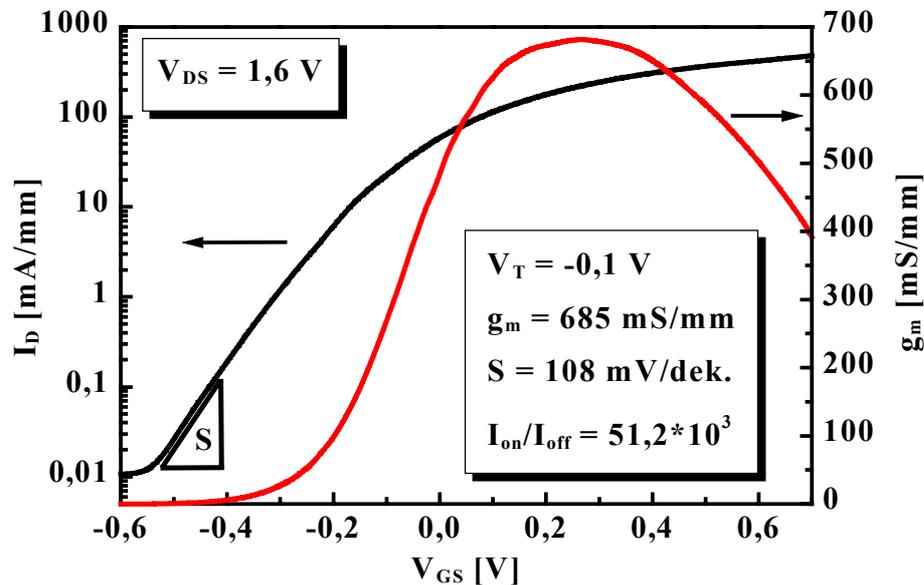


Abbildung 7.2: Logarithmische Transferkennlinie eines SiGe-n-Kanal-HFETs, realisiert mit dem selbstjustierenden Integrationskonzept

Abbildung 7.2 zeigt die logarithmische Transferkennlinie eines SiGe-n-Kanal-HFETs mit 90 nm Gatelänge, der mit dem selbstjustierenden Herstellungskonzept realisiert wurde. Die Aufnahme der Transferscharakteristik und die Bestimmung der Gleichstromkenngrößen erfolgt bei einer Ausgangsspannung von 1,6 V. Der Transistor lässt sich trotz der extrem kurzen Gatelänge ohne Durchbrucherscheinungen bis $V_{DS} = 4$ V aussteuern und weist damit eine gute Spannungsfestigkeit auf. Aufgrund der lateralen Optimierung des Bauelementes und der geringen Gateabmessungen ergibt sich eine maximale Steilheit von 685 mS/mm, was den höchsten extrinsischen Steilheitswert darstellt, der bisher für SiGe-basierte Transistoren veröffentlicht wurde [7.1]. Die Anwendung eines Feldeffekttransistors im Speicher- oder Logikbereich erfordert ein Minimum an Leckströmen im Off-Zustand, um die statische Verlustleistung komplementärer Transistorzusammenschaltungen klein zu halten. Der von der SIA-Roadmap 2001 angesetzte Wert des maximalen I_{off} für eine vergleichbare 100 nm Technologie beträgt bei 25 °C 70 nA/ μ m. Mit einem minimalen I_{off} von 10 nA/ μ m kann der SiGe-HFET diese Anforderung trotz verlustbehaftetem Schottky-Gate noch unterbieten. Gleichzeitig soll sich das Bauelement durch gute Stromtreiberfähigkeiten auszeichnen, um auch große Lasten steuern zu können. Mit einem I_{on}/I_{off} -Verhältnis von $51,2 \cdot 10^3$ unterstreicht der Transistor eine ausgewogene Dimen-

sionierung mit geringen Leckströmen im abgeschnürten Zustand und hoher Stromdichte am Ausgang im Flußbetrieb.

Die Leckströme im Off-Zustand des Transistors werden primär vom Sperrstrom des Schottky-Gates bestimmt. Das belegt die in Abbildung 7.3 dargestellte Schottky-Kennlinie des Transistors aus Abbildung 7.2. Die Diodencharakteristik wurde zwischen dem Source-Anschluß und der Gateelektrode gemessen, dabei lag der Source-Kontakt auf Massepotential. Der Sperrstrom der Schottky-Diode sättigt im gleichen Größenbereich wie der Reststrom des Transistors im Off-Zustand, so daß der Metall-Halbleiter-Übergang als dominierender Faktor beim Abschalten des Transistors angesehen werden muß.

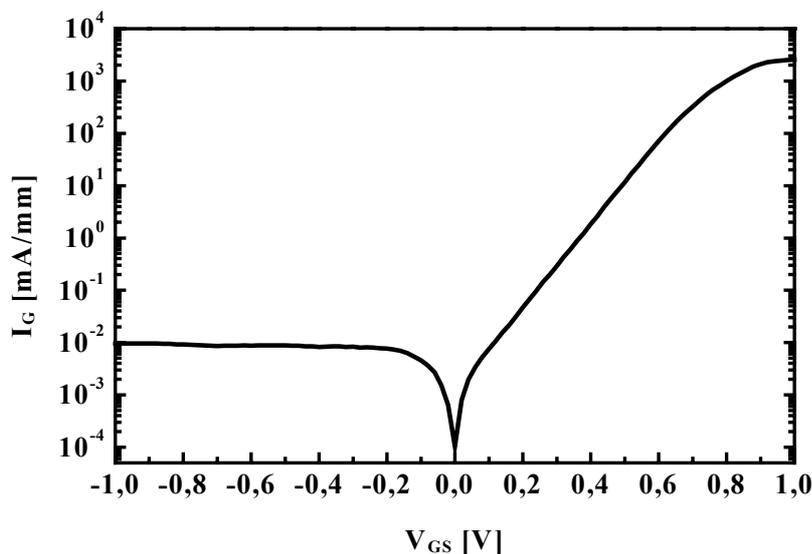


Abbildung 7.3: Typische Schottky-Gate Kennlinie eines SiGe-HFETs

Einen weiteren Beitrag zum Leckstrom liefern Ladungsträgerflüsse, die über den Puffer zwischen benachbarten Bauelementen verlaufen. Zwar stellen Puffer und p⁻-Substrat einen hochohmigen Widerstand dar, jedoch können nachweislich noch Restströme im niedrigen nA-Bereich gemessen werden. Ein Ansatz zur Reduzierung dieser Leckströme ist die Einführung einer p-dotierten Epitaxieschicht zwischen Puffer und der rückseitigen n⁺-Dotierschicht. Sie bildet in Verbindung mit dem aktiven Schichtsystem des HFETs eine sperrende pn-Diode, die Leckströme verringern soll und gleichzeitig zur Absenkung des Fermi-Niveaus auf der Rückseite der Heterostruktur führt. Dadurch wird gewährleistet, daß der

Transistor im Off-Zustand gut abschnürt, und gleichzeitig läßt sich über die Höhe der p-Dotierung auf die Schwellspannung des Bauelementes Einfluß nehmen. Bei starker Bor-Dotierung verschiebt sich die Schwellspannung sogar bis hin zu positiven V_{GS} -Werten und realisiert dadurch einen Anreicherungs-transistor. Bei moderaten Bor-Konzentrationen bleibt der Verarmungscharakter des Bauelementes erhalten.

Wie bereits bei der Vorstellung der Integrationskonzepte in Kapitel 5 beschrieben, nutzt das konventionelle Herstellungsverfahren eine Feldoxidrealisierung mit überlappenden Oxidkanten im Randbereich der Mesa. Die dabei auftretenden nicht steuerbaren Kanalbereiche unterhalb des Feldoxides führen, insbesondere bei Bauelementen mit kleinen Gateweiten, zu signifikanten, nicht vernachlässigbaren Leckströmen im abgeschnürten Zustand. Anschaulich ist dieses Problem in Abbildung 7.4 wiedergegeben. Das dargestellte Ausgangskennlinienfeld eines Transistors mit konventioneller Feldoxidprozessierung zeigt deutlich eine mangelnde Abschnürung bei negativen Gate/Source-Spannungen mit einem Leckstromanteil von bis zu 5 % am maximalen Sättigungsstrom.

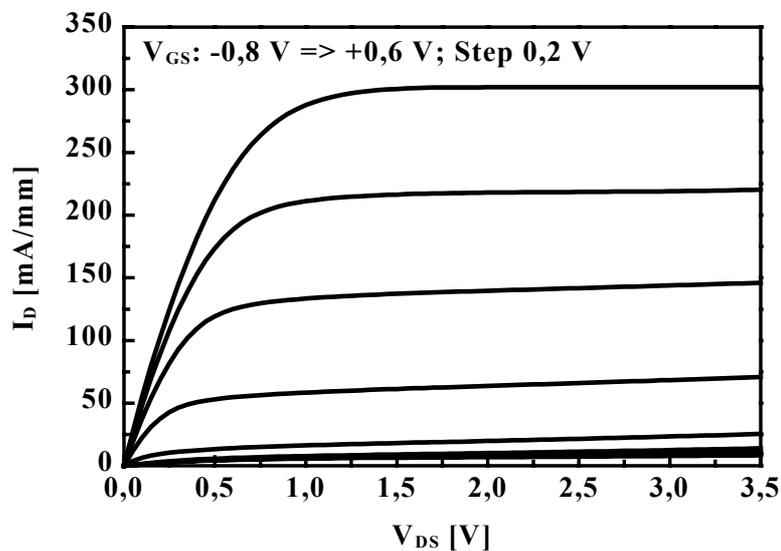


Abbildung 7.4: Ausgangskennlinienfeld eines HFETs der konventionellen Herstellungsvariante mit überlappendem Feldoxid

Bei einer selbstjustierenden Anordnung des Feldoxides, die bündig mit der oberen Mesakante abschließt, fallen diese parasitären Ströme weg. Der Kennlinienverlauf für negative Gate/Source-Spannungen zeigt dadurch keinen Anstieg mit

zunehmenden V_{DS} mehr und mündet im abgeschnürten Zustand im Sperrstrom der Schottky-Diode (Abbildung 7.5).

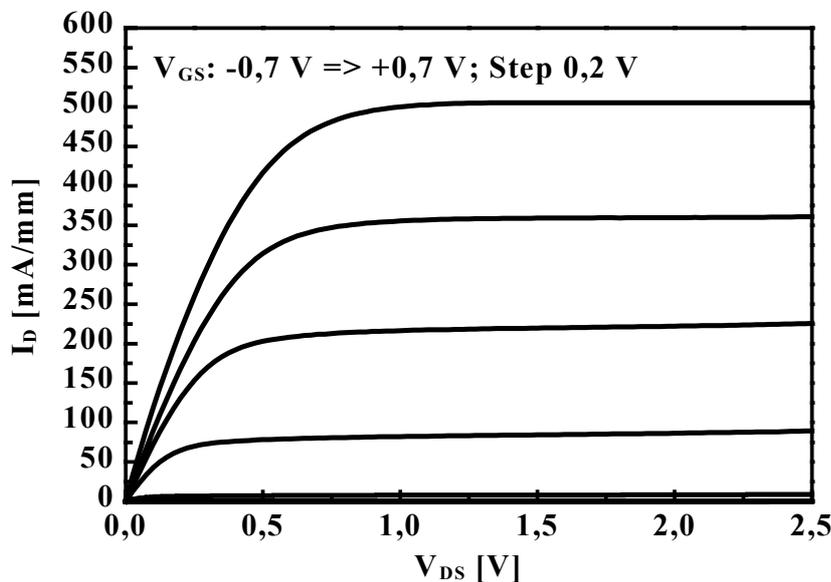


Abbildung 7.5: Ausgangskennlinienfeld eines selbstjustierten SiGe-HFETs mit optimierter Feldoxidanordnung

Beide Ausgangskennlinienfelder zeichnen sich durch eine niedrige Kniespannung von unter 1 V aus. Im linearen Anlaufbereich des Transistors sind neben bauelementspezifischen Abmessungen in erster Linie die Ladungsträgerbeweglichkeit μ_0 und die parasitären Widerstände R_s und R_d für den Anstieg des Drainstromes verantwortlich. Bei vergleichbaren Beweglichkeiten von 1150 - 1200 cm^2/Vs für beide Transistoren unterscheiden sie sich maßgeblich in der lateralen Optimierung. Durch die selbstjustierende Anordnung der Implantationsgebiete zum Gate verkürzt sich der Source/Drain-Abstand von 2,5 μm bei der konventionellen Herstellungsmethode, auf etwa 500 nm beim neuen Integrationskonzept. Die Verkürzung des ungesteuerten Bahngebietes um 2 μm Länge führt bei einem Schichtwiderstand des Kanals von 1000 - 1100 Ω/\square zu einer Widerstandsreduktion von ca. 40 Ω bei einem 50 μm weiten Transistor. Dadurch läßt sich erklären, warum die Kniespannung des selbstjustierten Transistors mit 0,65 V noch einmal um 0,25 V gegenüber dem konventionellen Bauelement verringert werden konnte und der Sättigungsstrom im Arbeitspunkt $V_{DS} = 2,0$ V und $V_{GS} = 0,6$ V eine Steigerung um 43 % auf 433 mA/mm aufweist.

Für analoge Schaltungsanwendungen spielt der Ausgangsleitwert g_{ds} des eingesetzten Transistors eine wichtige Rolle. Zum einen definiert sich über ihn, zusammen mit der Steilheit, die Spannungsverstärkung des Bauelementes (g_m/g_{ds}), und zum anderen bestimmt er maßgeblich die optimale Anpassung zwischen dem Transistorausgang und der Schaltung. Ein minimaler Ausgangsleitwert g_{ds} ist deshalb, auch in Hinblick auf eine hohe Grenzfrequenz der Leistungsverstärkung f_{max} , ein erstrebenswertes Ziel beim Transistordesign. Jedoch führt die permanent vorangetriebene Miniaturisierung der Bauelemente und die dadurch induzierten Kurzkanaleffekte, wie z. B. die Kanallängenmodulation, zu einer Transistorcharakteristik, die immer stärker von dem Idealbild abweicht. Zusätzlich machen sich beim modulationsdotierten HFET parallele Stromflüsse in parasitären Kanälen und Leckströme über den Puffer direkt im Ausgangsleitwert bemerkbar. Eine optimale Auslegung des heterogenen Schichtaufbaus, mit gut aufeinander abgestimmten Dotierstoffkonzentrationen und Schichtdicken, ist deshalb die Basis für eine ausgewogene Bauelementcharakteristik. Beide dargestellten Ausgangskennlinienfelder zeigen eine ausgeprägte Sättigung des Drainstromes, auch bei hohen Drain/Source-Spannungen. Daraus läßt sich schließen, daß Kurzkanaleffekte, trotz der kleinen Gatelängen von unter 100 nm, bisher keinen gravierenden Einfluß auf das Bauelementverhalten haben und ein gutes „Confinement“ des 2DEGs im Potentialtopf vorherrscht. Für den selbstjustierten Transistor, dessen Layout aufgrund der Selbstorganisation der einzelnen Elemente eine höhere Anfälligkeit gegenüber Kurzkanaleffekten vermuten läßt, konnte durch das Einfügen der bereits erwähnten p^+ -Schicht eine drastische Minderung der parasitären Einflüsse erzielt werden.

Das Verhalten des Ausgangsleitwertes g_{ds} bei verschiedenen Gate/Source-Spannungen für den selbstjustierten Transistor aus Abbildung 7.5 zeigt Abbildung 7.6. g_{ds} startet für alle Arbeitspunkte bei hohen positiven Leitwerten und weist dann eine asymptotische Annäherung an den endgültigen Steigungswert des Drainstromes auf. Dabei nähert sich die Änderung des Ausgangsstromes mit höheren Gate/Source-Spannungen immer mehr dem Idealwert von Null an, ohne ihn jedoch zu erreichen. Zusätzlich stellt die Abbildung noch den Verlauf der Spannungsverstärkung des Bauelementes bei $V_{GS} = 0,7 \text{ V}$ in Abhängigkeit von der Drain/Source-Spannung dar. Dieser Quotient aus der Steilheit g_m und dem Leitwert g_{ds} steigt mit zunehmender Sättigung des Ausgangsstromes steil an und erreicht im Bereich von $V_{DS} \approx 1,8 - 2,0 \text{ V}$ mit dem Steilheitsmaximum seinen Höchstwert. Für höhere Ausgangsspannungen reduziert sich die Steilheit wieder und damit auch die Spannungsverstärkung.

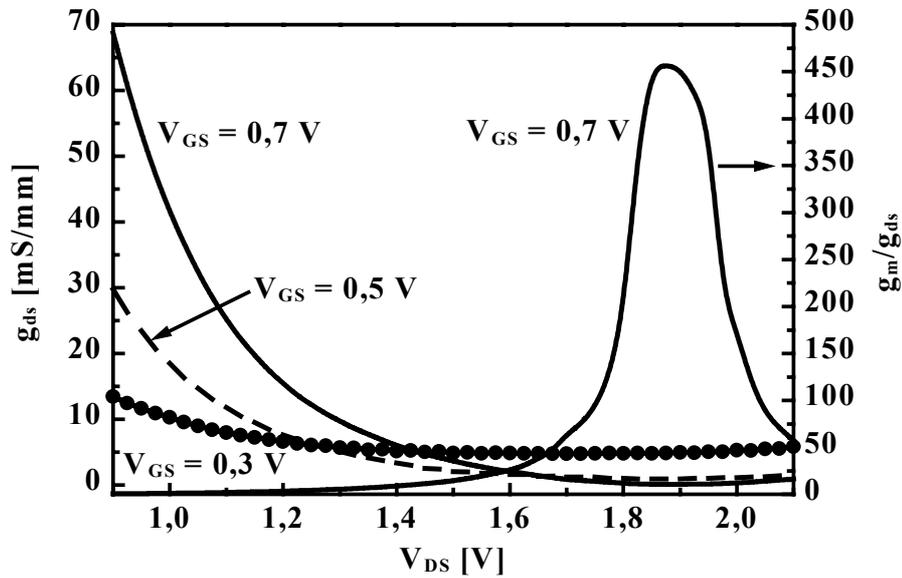


Abbildung 7.6: Verlauf des Ausgangsleitwertes g_{ds} des selbstjustierten Transistors aus Abbildung 7.5 für verschiedene Arbeitspunkte V_{GS} . Abhängigkeit der Spannungsverstärkung g_m/g_{ds} von der Ausgangsspannung des Bauelementes

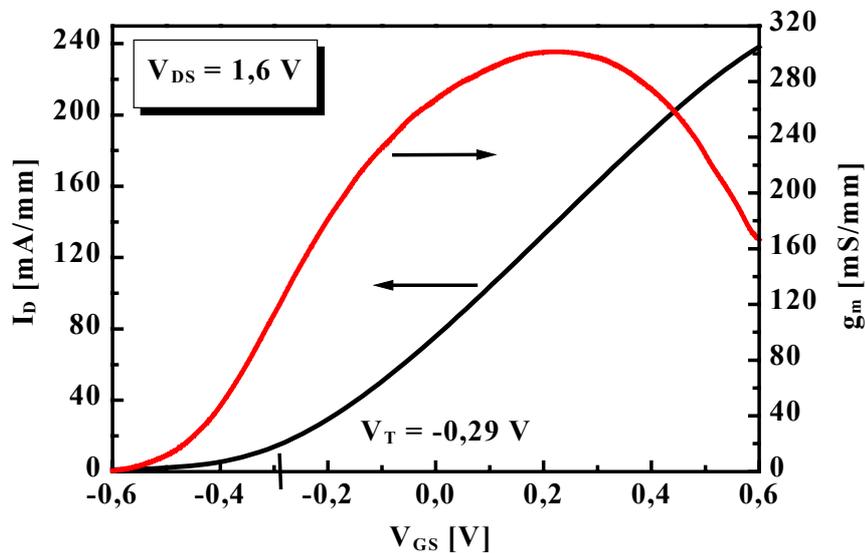


Abbildung 7.7: Typische Transferkennlinie eines SiGe-n-Kanal-HFETs (Verarmungstyp) mit einer Gatelänge von 100 nm, realisiert mit dem konventionellen Integrationskonzept

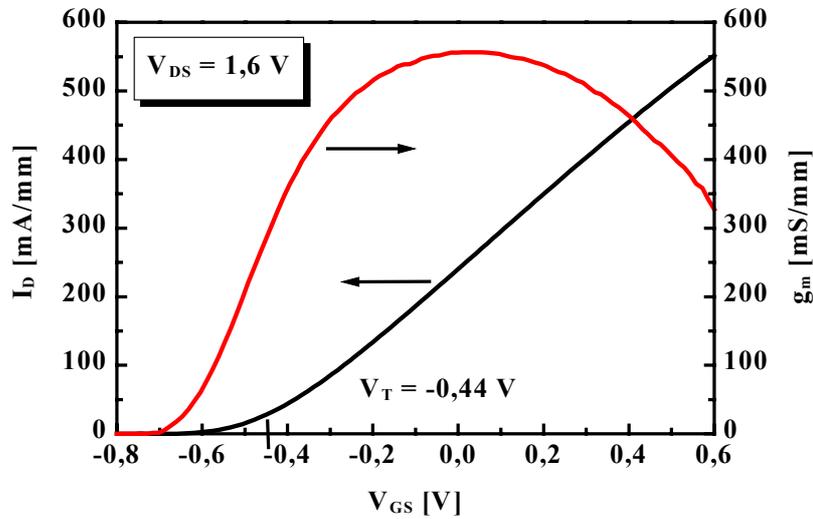


Abbildung 7.8: Typische Transferkennlinie eines selbstjustierten SiGe-HFETs (Verarmungstyp) mit einer Gatelänge von 100 nm

In Abbildung 7.7 ist die Transferkennlinie eines Transistors dargestellt, der mit dem konventionellen Herstellungsablauf realisiert wurde. Zum Vergleich zeigt Abbildung 7.8 die Übertragungskennlinie eines Transistors mit vergleichbarer Schichtkonfiguration und identischer Gatelänge, der jedoch mit dem selbstjustierenden Integrationskonzept entstanden ist. Bei beiden Bauelementen handelt es sich um Verarmungstransistoren, die bei einer Gatespannung von Null Volt bereits einen leitenden Kanal aufweisen. Dadurch ergibt sich ein auffallend breiter Steilheitsverlauf, der sich über den ganzen Steuerbereich am Eingang des Transistors erstreckt. Für Gatespannungen unterhalb der Schwellspannung zeigen beide Bauelemente ein sauberes Abschnürverhalten ohne signifikant hohe Leckströme zwischen Drain und Source. Der Anlaufbereich zeichnet sich durch eine weite Linearität zwischen Drainstrom und Gatespannung aus. Da der Feldeffekttransistor ein nichtlineares Bauelement ist, spielen diese Bereiche eine wichtige Rolle als Kriterium für die Anwendung in analogen Schaltungen.

Trotz eines vergleichbaren Strom- und Steilheitsverlaufes unterscheiden sich die Bauelemente gravierend in ihrer Leistungsfähigkeit. Diese Tatsache spiegelt sich, wie schon bei der Ausgangskennlinie, vor allem im höheren Drainstrom und in der größeren Steilheit wider. Durch den Einsatz des neuen, selbstjustierenden Herstellungs Konzeptes konnte der auf die Gateweite normierte Drainstrom mehr als verdoppelt werden. Zudem kommt es zu einer Steigerung des Steilheitsmaximums von 304 mS/mm auf über 550 mS/mm. Diese signifikante Leistungsver-

besserung basiert primär auf zwei Modifikationen im Herstellungsablauf. Zum einen kommt hier, wie schon angesprochen, die laterale Optimierung der Bauelementstruktur mit der selbstjustierenden Anordnung der Source/Drain-Gebiete zur Gateelektrode zum Tragen, und zum anderen bringt der Einsatz des Nickelsilizides und die bessere Abstimmung der Implantationsparameter eine signifikante Verringerung der parasitären Zuleitungswiderstände.

Die meßtechnische Erfassung der Widerstandsanteile, die das elektrische Verhalten des Bauelementes maßgeblich beeinflussen, ist ein wichtiger Bestandteil der Gleichstromcharakterisierung des Transistors. Die Extraktion dieser Werte dient neben der Kontrolle der Prozeßparameter auch zur objektiven Bewertung unterschiedlicher Technologiekonzepte und neuer Prozeßmodifikationen. Darüber hinaus sind die Widerstände R_s , R_d und R_g wichtige Komponenten des Kleinsignalersatzschaltbildes und bilden so die Grundlage für eine spätere Simulation des Transistorverhaltens bei Schaltungsanwendungen.

Ein bekanntes Verfahren zur Bestimmung der relevanten Widerstandsanteile ist die „End Resistance-Methode“ [7.2], [7.3]. Dabei wird ausgenutzt, daß eine Vorwärtspolung des Schottky-Gates einen signifikanten Stromfluß erzeugt, was, im Gegensatz zum isolierenden MOS-Gate, zusätzliche Freiheitsgrade bei den Meßkonfigurationen eröffnet. Bei einem definierten Strom I_G zwischen Gate und Sourceelektrode und einer stromlosen Spannungsmessung zwischen dem „flo-tenden“ Drainanschluß und Source ergibt sich ein Spannungsabfall über R_s und einem Anteil α des gesteuerten Kanals R_K . Für kleine Spannungen $V_{DS} \ll nkT/q$ ergibt sich α zu 0,5 [7.4]. Dabei steht kT/q für die thermische Spannung und n für den Idealitätsfaktor der Schottky-Diode.

$$\frac{V_{DS}}{I_G} = R_s + \alpha \cdot R_K \quad \text{mit } \alpha = 0,5 \text{ für } V_{DS} \ll \frac{kT}{q} \quad (7.1)$$

Der Widerstandsanteil R_s setzt sich dabei aus dem Metallwiderstand R_M der ersten Kontaktierungsebene, dem Kontaktwiderstand R_C zwischen der ohmschen Metallisierung und dem hochdotierten n⁺-Halbleiter, dem Bahnwiderstand R_{impl} des Implantationsgebietes und dem Widerstand des ungesteuerten Kanalgebietes R_{Si} zusammen. Analoges gilt für die drainseitige Widerstandskomponente.

$$\frac{V_{DS}}{I_G} = R_d + \alpha \cdot R_K \quad \text{mit } \alpha = 0,5 \text{ für } V_{DS} \ll \frac{kT}{q} \quad (7.2)$$

Erfolgt die Spannungsmessung zwischen Gate und Source, bei gleichzeitig eingepprägtem Gatestrom I_G und „floatendem“ Drainanschluß, so ergibt sich Gleichung (7.3) bzw. für die Bestimmung der Widerstandsanteile auf der Drainseite Gleichung (7.4).

$$\frac{dV_{GS}}{dI_G} = R_g + n \cdot \frac{kT}{qI_G} + R_s + \frac{R_K}{3} \quad (7.3)$$

$$\frac{dV_{GS}}{dI_G} = R_g + n \cdot \frac{kT}{qI_G} + R_d + \frac{R_K}{3} \quad (7.4)$$

Der gesamte Widerstandsanteil des Schottky-Gates setzt sich dabei aus der Kleinsignalkomponente $R_{dy} = nkT/qI_G$ und dem metallischen „end-to-end“-Widerstand R_g zusammen, der sich material- und geometrieabhängig zu $R_g = \rho_g \cdot W_G/A_G$ ergibt. Bei der späteren Bestimmung der Kleinsignalersatzschaltbildparameter gilt zu bedenken, daß R_g für hochfrequente Signale eine leerlaufende Leitung darstellt, die bei Modellierung durch ein verteiltes RC-Netzwerk, je nach Anzahl der Gatefinger n , nur einen Teil des metallischen Widerstandes berücksichtigt (Gleichung 7.5) [7.5].

$$R_{g,HF} = \frac{R_g}{3n^2} = \rho_g \cdot \frac{W_G}{3A_G n^2} \quad (7.5)$$

Bei einem quantitativen Vergleich der beiden Widerstandskomponenten R_{dy} und R_g des Gates wird deutlich, daß der metallische Anteil nur einen vernachlässigbaren Beitrag liefert und der dynamische Widerstand den Gesamtwert dominiert. Für die Schottky-Gate Kennlinie in Abbildung 7.3 ergibt sich bei einem eingepprägten Gatestrom $I_G = 1 \text{ mA}$ ein R_{dy} von $54,6 \text{ } \Omega$, während R_g für denselben Transistor bei einer Gateweite von $56 \text{ } \mu\text{m}$ nur $1,5 \text{ } \Omega$ beträgt und damit einen Anteil von $2,7 \text{ } \%$ am Gesamtwiderstand hat.

Eine alternative Methode, um die Widerstandsbeiträge R_s , R_d und R_K zu bestimmen, nutzt neben dem fest eingepprägten Gatestrom I_G einen zusätzlichen Stromfluß I_D zwischen Drain und Source. Dabei ist zu beachten, daß $I_D \ll I_G$ ist, damit keine substantielle Potentialänderung entlang des Kanals erzeugt wird. Unter dieser Randbedingung gilt das Superpositionsprinzip für die Spannungsabfälle der einzelnen Stromanteile, und die Gesamtspannung zwischen Drain und Source ergibt sich aus Gleichung 7.6. Die grafische Darstellung von V_{DS} als Funktion

von I_D liefert dann $R_s+R_d+R_K$ als Steigung der Ausgangsgeraden und $R_s+R_K/2$ als Schnittpunkt mit der y-Achse.

$$V_{DS} = (R_s + R_d + R_K) \cdot I_D + \left(R_s + \frac{R_K}{2}\right) \cdot I_G \quad (7.6)$$

Ein Vergleich der beiden vorgestellten Methoden liefert eine maximale Abweichung von $\pm 5\%$ zwischen den extrahierten Widerstandswerten, so daß beide Ansätze gleichwertig einsetzbar sind. Die lineare Abhängigkeit der ermittelten Gleichungen erlaubt jedoch keine explizite Bestimmung aller Komponenten eines Transistors, so daß eine zusätzliche Meßanordnung zur Auflösung des Problems beitragen muß. Durch die Anwendung der zweiten Methode auf Bauelemente unterschiedlicher Gatelänge läßt sich eine gesteuerte TML Struktur nachahmen. Als Ergebnis liefern diese Messungen den Gesamtwiderstand $R_s+R_d+R_K$ der Transistoren zwischen Source- und Drainanschluß als Funktion der Gatelänge (Abbildung 7.9). Aus der Steigung des Gesamtwiderstandes ergibt sich dabei der Kanalwiderstand R_K pro μm Gatelänge, und der Schnittpunkt mit der y-Achse liefert den Term R_s+R_d .

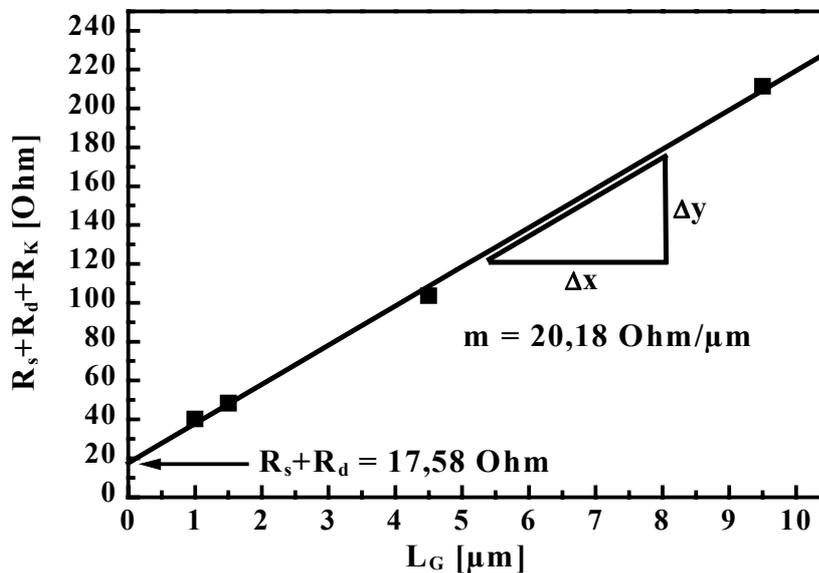


Abbildung 7.9: Ermittlung des Gesamtwiderstandes $R_s+R_d+R_K$ für Transistoren unterschiedlicher Gatelänge und Extraktion des gesteuerten Kanalwiderstandes $R_K/\mu\text{m}$ Gatelänge und der Widerstandsanteile R_s+R_d aus der Steigung m der Ausgleichsgeraden bzw. aus dem Schnittpunkt mit der y-Achse

Mit den dargestellten Analysemethoden sind die Widerstandskomponenten R_s , R_d und R_g von Transistoren unterschiedlicher Technologiekonzepte und verschiedener Entwicklungsstufen der ohmschen Kontakte verglichen worden (Tabelle 7.1). Dabei zeigt sich ganz deutlich, daß eine Reduzierung der lateralen Abmessungen von 2,5 μm Drain/Source-Abstand auf weniger als 0,5 μm bei dem selbstjustierenden Integrationskonzept eine drastische Verringerung der Zugangswiderstände mit sich bringt. Außerdem macht sich auch die niederohmige Kontakttechnologie unter Einsatz des NiSi bemerkbar, die eine zusätzliche Verbesserung von etwa 40 % am realen Bauelement erzielt. Die Kombination beider Prozeßmodifikationen liefert schließlich die bereits in diesem Kapitel analysierten signifikanten Verbesserungen bei der Steilheit g_m und den erzielbaren Stromdichten der Transistoren.

| Technologie | Gatelänge [nm] | R_s [$\Omega \cdot \text{mm}$] | R_d [$\Omega \cdot \text{mm}$] | R_g [Ω/mm] |
|--|-------------------|---------------------------------------|---------------------------------------|---------------------------------|
| konventionelle Technologie, Kontakte nur implantiert Abstand D/S: 2,5 μm | 180 asymmetr. | 1,82 | 2,43 | 59,5 |
| konventionelle Technologie, Kontakte nur implantiert Abstand D/S: 2,0 μm | 130 asymmetr. | 1,34 | 2,21 | 70,1 |
| konventionelle Technologie, Kontakte impl. + 8 nm NiSi Abstand D/S: 2,0 μm | 130 asymmetr. | 1,02 | 1,58 | 68,4 |
| selbstjustierende Technologie Kontakte impl. + 8 nm NiSi Abstand D/S: 0,4 μm | 100 symmetr. | 0,192 | 0,196 | 30,1 |
| selbstjustierende Technologie Kontakte impl. + 10 nm NiSi Abstand D/S: 0,3 μm | 90 symmetr. | 0,181 | 0,175 | 25,7 |

Tabelle 7.1: Vergleich der wirksamen Widerstände R_s , R_d und R_g für das konventionelle und das selbstjustierende Technologiekonzept bei verschiedenen Entwicklungsstufen der ohmschen Kontaktgebiete

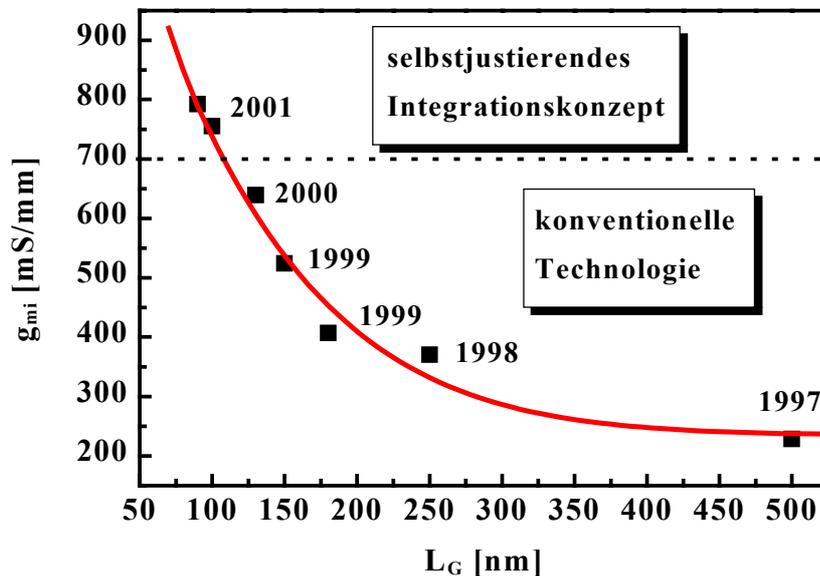


Abbildung 7.10: Verlauf der intrinsischen Steilheit g_{mi} in Abhängigkeit von der effektiven Gatelänge L_G der SiGe-n-Kanal-HFETs

Um die Steilheit g_m der Transistoren auch für verschiedene Technologiekonzepte vergleichbar machen zu können, bietet sich der, um die parasitären Widerstände R_s und R_d bereinigte, intrinsische Wert g_{mi} an. Abbildung 7.10 zeigt dabei die Entwicklung der intrinsischen Steilheit g_{mi} in Abhängigkeit von der Gatelänge L_G der SiGe-HFETs, die im Rahmen dieser Arbeit und bei vorausgehenden Untersuchungen entstanden sind.

7.2 Charakterisierung im Hochfrequenzbereich

Die Hochfrequenzcharakterisierung, der im Rahmen dieser Arbeit entstandenen SiGe-n-Kanal-HFETs, umfaßt die Messung der bauelementspezifischen Streuparameter mit der Ableitung der Verstärkerkenngrößen und Grenzfrequenzen und die Analyse des Rauschverhaltens. Darüber hinaus bilden die an verschiedenen Arbeitspunkten ermittelten S-Parameter die Basis für die Extraktion der Ersatzschaltbildelemente und damit für die Umsetzung eines Kleinsignalmodells, welches das Transistorverhalten für eine große Frequenzbandbreite adäquat nachbildet.

7.2.1 Verstärkungen und Grenzfrequenzen der Transistoren

Bei der Hochfrequenzmessung wird dem gleichstrommässig eingestellten Arbeitspunkt des Transistors ein hochfrequentes, sinusförmiges Signal überlagert. Bei beliebiger Amplitude des Signals ist das Übertragungsverhalten des Transistors im allgemeinen nichtlinear. Eine vollständige Berechnung des Bauelementverhaltens mit den Methoden der Netzwerkanalyse ist deshalb nicht durchführbar. Außerdem führen die nichtlinearen Kennlinien zusätzlich zu einer Verzerrung des verstärkten Signals. Bei einer Aussteuerung mit Signalen kleiner Amplitude kann die Kennlinie im Arbeitspunkt linearisiert werden und die Anwendung der Knoten- und Maschenregeln ist wieder zulässig. Dadurch läßt sich das Transistorverhalten durch ein Ersatzschaltbild beschreiben, dessen Elemente vom Arbeitspunkt abhängig, für kleine Wechselamplituden jedoch von der Aussteuerung unabhängig sind.

Beim Einsatz der SiGe-HFETs in Hochfrequenzschaltungen, wie z. B. rauscharmen Verstärkern (LNA) oder Oszillatoren, werden die Kleinsignaleigenschaften der aktiven Bauelemente bei Frequenzen bis zu einigen 100 MHz durch Y-, Z- oder h-Parameter beschrieben. Die Ermittlung der gesuchten Leitwertparameter beschränkt sich dabei auf einfache Strom- und Spannungsmessungen an den Ein- und Ausgängen des Bauelementes. In höheren Frequenzbereichen stellt sich die Messung dieser Parameter als zunehmend ungenau dar. Ein Kurzschluß ist dann unweigerlich mit der Induktivität der Zuleitung behaftet, und ein Leerlauf hat eine nicht vernachlässigbare kapazitive Wirkung. Darüber hinaus reduziert sich die Wellenlänge der elektromagnetischen Schwingung mit zunehmender Frequenz bis in den Bereich der Schaltungs- und Bauelementabmessungen, so daß sie sich auf die Laufzeit der Ladungsträger in den Transistoren auswirken kann. Außerdem bekommen parasitäre Widerstände, Kapazitäten und Induktivitäten und selbst die Verbindungsleitungen zwischen den Schaltungskomponenten einen nicht vernachlässigbaren Einfluß auf das Bauelementverhalten.

Deshalb hat sich bei hohen Frequenzen die Charakterisierung der Transistoren mit Hilfe der S-Parameter durchgesetzt. Durch Messung der hin- und rücklaufenden elektromagnetischen Leistungswellen am Ein- und Ausgang des Zweitorts läßt sich eine störungsfreie Extraktion der gesuchten Vierpolparameter gewährleisten. Dabei erfüllt die S-Parametermessung mit Signalpegeln von -10dBm, was einem Spannungsabfall von 70 mV an einem 50 Ω Widerstand entspricht, die Kleinsignalbedingungen, so daß der nichtlineare HFET als lineares Netzwerk

beschrieben werden kann. Durch die Einbettung des zu messenden Vierpols in ein genormtes 50Ω Leitungssystem, das mit einem 50Ω Lastwiderstand abgeschlossen ist, werden die ehemals notwendigen Leerlauf- und Kurzschlußtests durch einfache Wellenmessungen ersetzt.

Die Hochfrequenzcharakterisierung der SiGe-HFETs erfolgt vorwiegend in Common-Source-Konfiguration. Dabei bilden Gate und Source den Eingang und Drain und Source den Ausgang des Vierpols. Die zugehörige Schaltungsanordnung des Transistors mit den hin- und rücklaufenden Wellen \underline{a}_1 , \underline{a}_2 , \underline{b}_1 und \underline{b}_2 zeigt Abbildung 7.11.

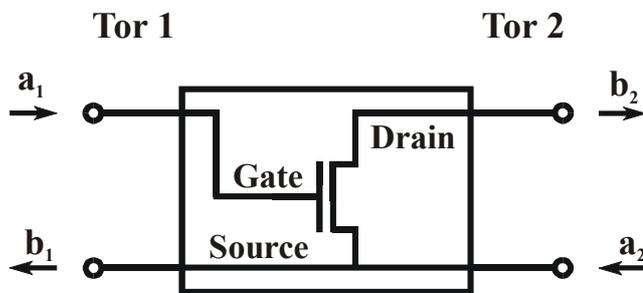


Abbildung 7.11: Feldeffekttransistor in Common-Source-Konfiguration bei der S-Parametermessung

Die formale Beschreibung dieser Anordnung erfolgt über die S-Parameter-Matrix in Gleichung 7.7. Sie verknüpft die hinlaufenden Wellen \underline{a}_n mit den rücklaufenden Wellen \underline{b}_n und liefert gleichzeitig die Definition der einzelnen Parameter bei Abschluß mit einem normierten Wellenwiderstand von 50Ω (Gleichung 7.8 - 7.11).

$$\begin{pmatrix} \underline{b}_1 \\ \underline{b}_2 \end{pmatrix} = \begin{pmatrix} \underline{S}_{11} & \underline{S}_{12} \\ \underline{S}_{21} & \underline{S}_{22} \end{pmatrix} \cdot \begin{pmatrix} \underline{a}_1 \\ \underline{a}_2 \end{pmatrix} \quad (7.7)$$

mit

$$\underline{S}_{11} = \left. \frac{\underline{b}_1}{\underline{a}_1} \right|_{\underline{a}_2=0} \quad \text{Eingangsreflektionsfaktor} \quad (7.8)$$

$$\underline{S}_{12} = \left. \frac{\underline{b}_1}{\underline{a}_2} \right|_{a_1=0} \quad \text{Rückwärtsübertragungsfaktor} \quad (7.9)$$

$$\underline{S}_{21} = \left. \frac{\underline{b}_2}{\underline{a}_1} \right|_{a_2=0} \quad \text{Vorwärtsübertragungsfaktor} \quad (7.10)$$

$$\underline{S}_{22} = \left. \frac{\underline{b}_2}{\underline{a}_2} \right|_{a_1=0} \quad \text{Ausgangsreflektionsfaktor} \quad (7.11)$$

Zur Bewertung der Hochfrequenzeigenschaften von Transistoren werden definierte Verstärkungskennzahlen und Grenzfrequenzen eingesetzt. In Kapitel 3 sind bereits die Transitfrequenz f_T und die maximale Schwingfrequenz f_{\max} in Abhängigkeit von den Elementen des Kleinsignalersatzschaltbildes eingeführt worden. Diese Grenzfrequenzen lassen sich jedoch auch, unabhängig von der Kenntnis dieser Bauelemente, allein mit Hilfe der gemessenen S-Parameter bestimmen. Die für die Ermittlung der Transitfrequenz notwendige Kurzschlußstromverstärkung \underline{h}_{21} ergibt sich unter Einbeziehung der S-Parameter zu:

$$\underline{h}_{21} = \left. \frac{\underline{I}_2}{\underline{I}_1} \right|_{U_2=0} = \frac{-2 \cdot \underline{S}_{21}}{(1 - \underline{S}_{11}) \cdot (1 + \underline{S}_{22}) + \underline{S}_{12} \cdot \underline{S}_{21}} \quad (7.12)$$

Aus dem Verlauf der dynamischen Stromverstärkung über die Frequenz ermittelt sich f_T bei einem Betragswert $|\underline{h}_{21}|$ von 1.

$$|\underline{h}_{21}|(f = f_T) = 1 \quad (7.13)$$

Im Rahmen dieser Arbeit haben diverse Verbesserungen bei der Elektronenstrahlolithografie und die Einführung des selbstjustierenden Integrationskonzeptes zu einer Verringerung der Gatelänge von anfangs $0,5 \mu\text{m}$ auf bis zu 90 nm geführt. Die daraus resultierende Erhöhung der intrinsischen Steilheit g_{mi} nach Abbildung 7.10 hat in ihrem Verlauf direkten Einfluß auf die Transitfrequenz f_T (Abbildung 7.12). Ein leicht steilerer Anstieg der Grenzfrequenz gegenüber g_{mi} beweist, daß es im Laufe der Technologieweiterentwicklung sogar gelungen ist, die Kapazitäten C_{gs} und C_{gd} bzw. die parasitären Widerstände R_s und R_d noch zu senken. Das maximale, bisher für einen SiGe-basierten FET veröffentlichte, f_T von 90 GHz konnte dabei für eine selbstjustierte Transistorvariante mit einer physikalischen Gatelänge von 90 nm erzielt werden.

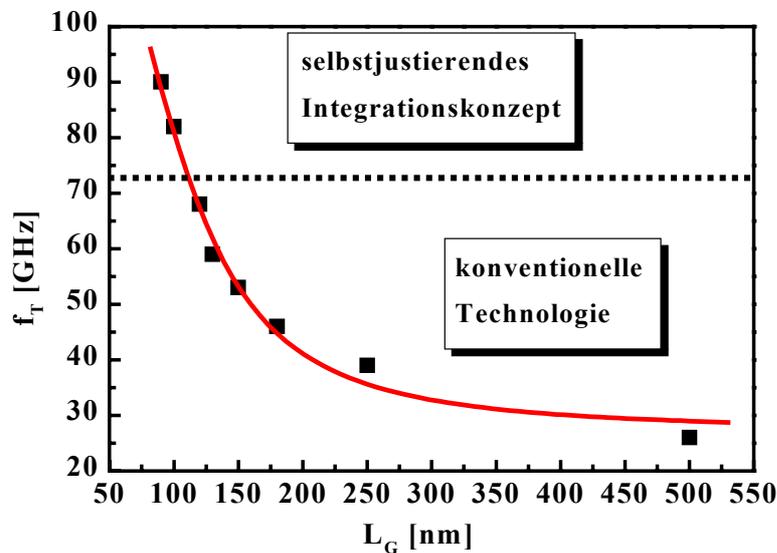


Abbildung 7.12: Verlauf der Transitfrequenz f_T über die Gatelänge L_G für SiGe-Kanal-HFETs verschiedener Integrationskonzepte

Die Beurteilung der Leistungsverstärkung bei Transistoren bedient sich definierter Kennzahlen aus der Schaltungstechnik. Dabei spielt die äußere Beschaltung und das Stabilitätsverhalten des Bauelementes eine entscheidende Rolle. Basierend auf den spezifischen Leistungsverläufen der verschiedenen Modelle ergeben sich unterschiedliche Grenzfrequenzen f_{\max} .

Die Einbettung des zu charakterisierenden Transistors in ein genormtes 50Ω Meßsystem führt meistens zu einer ein- und ausgangsseitigen Fehlanpassung. Um jedoch die maximal verfügbare Leistungsverstärkung MAG (Maximum Available Gain) des Bauelementes zu ermitteln, sind am Ein- und Ausgang konjugiert komplexe Anpassungsnetzwerke nötig, die eine reflektionsfreie Transmission der Signale gewährleisten. Diese Beschaltung liefert allerdings nur sinnvolle Leistungsdaten, wenn das System nicht zur Instabilität durch Schwingungen neigt. Deshalb ist die Leistungsdefinition MAG mit dem Stabilitätsfaktor k verknüpft und nur für stabile Schaltungen mit $k \geq 1$ gültig [7.6].

$$\text{MAG} = \left| \frac{S_{21}}{S_{12}} \right| \cdot \left(k \pm \sqrt{k^2 - 1} \right) \quad (7.14)$$

$$\text{mit } k = \frac{1 - |\underline{S}_{11}|^2 - |\underline{S}_{22}|^2 + |\underline{S}_{11} \cdot \underline{S}_{22} - \underline{S}_{12} \cdot \underline{S}_{21}|^2}{2 \cdot |\underline{S}_{12}| \cdot |\underline{S}_{21}|} \quad (7.15)$$

Der k Faktor überprüft das Kriterium der unbedingten Stabilität einer Schaltung. Für $k > 1$ gilt das betrachtete Netzwerk als unbedingt stabil, so daß bei beliebiger Anpassung am Ein- und Ausgang des Vierpols keine Schwingneigung auftritt. Für den Fall bedingter Stabilität $k < 1$ muß im Einzelfall überprüft werden, ob die gewählte Beschaltung zu unerwünschten Schwingungen führen kann oder nicht.

Liegt keine unbedingte Stabilität vor, so kann der dynamische Leistungsgewinn des Bauelementes trotzdem quantifiziert werden. Der MSG-Wert (Maximum Stable Gain) gibt dabei die maximale Verstärkung an, die unter stabilen Bedingungen ($k = 1$) erreichbar wäre.

$$\text{MSG} = \left| \frac{\underline{S}_{21}}{\underline{S}_{12}} \right| \quad (7.16)$$

Hochfrequenzmessungen an realen Transistoren weisen häufig bei der höchsten Meßfrequenz noch einen Stabilitätsfaktor $k < 1$ auf. Aufgrund kapazitiver Rückkopplungen zwischen Ein- und Ausgang des Bauelementes kommt es dabei zu Schwingungen und Instabilitäten. Eine konkrete Aussage über die maximale Schwingfrequenz f_{\max} ist deshalb unmöglich. Aus diesem Grund wird für die Approximation von f_{\max} häufig auf die Untersuchung der unilateralen Verstärkung U zurückgegriffen. Sie beschreibt die theoretische maximale Leistungsverstärkung eines unangepaßten linearen Vierpols, bei gedachter Neutralisierung seiner Rückwirkung mit Hilfe eines passiven Rückkopplungsnetzwerkes und bei anschließender verlustloser und konjugiert komplexer Anpassung des Ein- und Ausganges an das Gesamtsystem [7.6]. Die Frequenz, bei der die unilaterale Verstärkung U gleich eins wird, kennzeichnet den Wechsel vom aktiven zum passiven Bauelement und bestimmt $f_{\max}(U)$ als Grenze der Leistungsverstärkung.

$$U = \frac{\left| \frac{\underline{S}_{21}}{\underline{S}_{12}} - 1 \right|^2}{2 \cdot k \cdot \left| \frac{\underline{S}_{21}}{\underline{S}_{12}} \right| - 2 \cdot \text{Re} \left(\frac{\underline{S}_{21}}{\underline{S}_{12}} \right)} \quad (7.17)$$

Üblicherweise werden die zu charakterisierenden Einzeltransistoren im Layout in eine Meßstruktur eingebettet, die eine On-Wafer-Charakterisierung des Bauelementes mit speziellen Meßnadeln oder -köpfen ermöglicht. Diese Meßstrukturen haben jedoch, aufgrund ihrer großen Fläche im Vergleich zum eigentlichen Transistor, sowohl kapazitiv als auch induktiv einen großen Einfluß auf die gemessenen S-Parameter, so daß eine Verfälschung des intrinsischen Leistungsbildes des Bauelementes auftritt. Es gilt daher, diese Einflüsse quantitativ zu erfassen und durch ein sogenanntes „Deembedding“ von den reinen Transistordaten zu trennen. Auf den genauen Ablauf dieses Verfahrens wird in Abschnitt 7.2.2 noch detailliert eingegangen. Soweit nicht explizit erwähnt, sind alle im Rahmen dieser Arbeit präsentierten Hochfrequenzverstärkungen und Grenzfrequenzen von den Einflüssen der Meßstruktur bereinigt und damit „deembedded“.

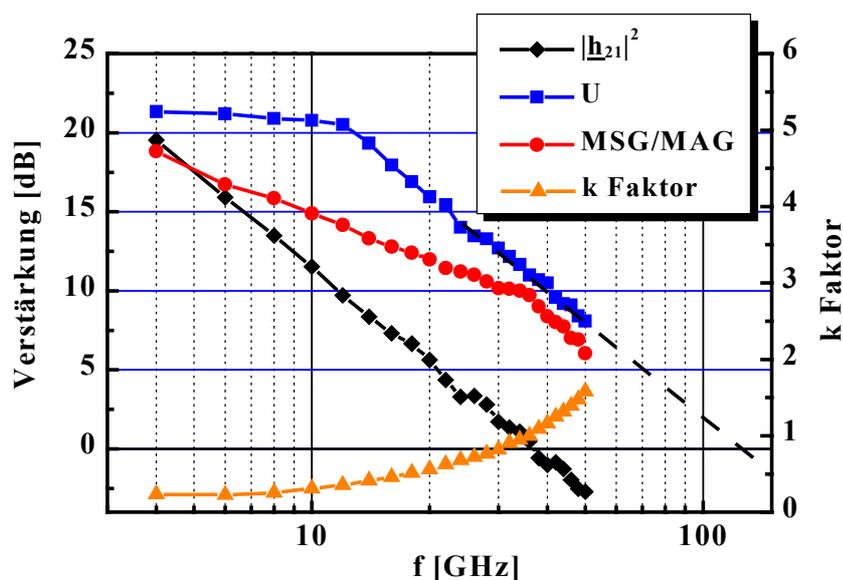


Abbildung 7.13: Verstärkungsverläufe ($|h_{21}|^2$, MSG/MAG und U) eines SiGe-HFETs konventioneller Bauart mit einer Gatelänge von 250 nm für $V_{DS} = 2,0$ V und $V_{GS} = 0$ V

Den typischen Verlauf der aus den gemessenen S-Parametern ermittelten Verstärkungen für einen SiGe-HFET konventioneller Bauart zeigt Abbildung 7.13. Um die Stromverstärkung $|h_{21}|$ mit den Leistungsdaten vergleichbar zu machen, wird das Betragsquadrat $|h_{21}|^2$ als Funktion der Frequenz aufgetragen. Der Schnittpunkt von $|h_{21}|^2$ mit der 0 dB-Achse liefert schließlich die Grenzfrequenz der Kurzschlußstromverstärkung f_T von 39 GHz. Der außerdem dargestellte Sta-

bilitätsfaktor k legt dar, daß der Transistor erst oberhalb einer Frequenz von 38 GHz im stabilen Bereich arbeitet. An dieser Stelle geht das Maximum Stable Gain (MSG) in das Maximum Available Gain (MAG) über. Zur Bestimmung der Leistungsgrenzfrequenzen $f_{\max}(\text{MAG})$ und $f_{\max}(\text{U})$ werden die Verstärkungsverläufe bei einem Wert von 6 bzw. 12 dB mit einer -20 dB/Dekade abfallenden Gerade extrapoliert und der Schnittpunkt mit der 0 dB-Achse als Grenzfrequenz festgelegt. Daraus ergibt sich für den hier betrachteten Transistor mit einer Gatelänge von 250 nm ein $f_{\max}(\text{MAG})$ von 108 GHz und für die unilaterale Verstärkung eine Grenzfrequenz von $f_{\max}(\text{U}) = 120$ GHz.

Für den Einsatz in ICs ist die Kenntnis des Transistorarbeitspunktes ein wichtiges Kriterium für die Dimensionierung der Schaltung. Die Darstellung der Steilheit g_m und der relevanten Grenzfrequenzen f_T und f_{\max} in Abhängigkeit von der Spannung V_{GS} liefert dabei eine wichtige Aussage über die Lage des optimalen Arbeitspunktes am Eingang des Transistors und die maximal tolerierbare Abweichung vom Nominalwert. Für den Transistor aus Abbildung 7.13 ergibt sich aus der Gleichstromcharakterisierung ein Steilheitsverlauf mit einem breiten Maximum, das im Spannungsbereich von -0,2 V bis +0,3 V ein hohes Niveau von über 300 mS/mm halten kann (Abbildung 7.14). Ein ähnlich breitbandiges Verhalten ergibt sich für die Grenzfrequenzen. Sie erreichen ihre Spitzenwerte im Anstieg der Steilheit bei $V_{GS} = -0,2$ V und verfügen über ein breites Plateau mit maximalen $f_{\max}(\text{U})$ - und f_T -Werten.

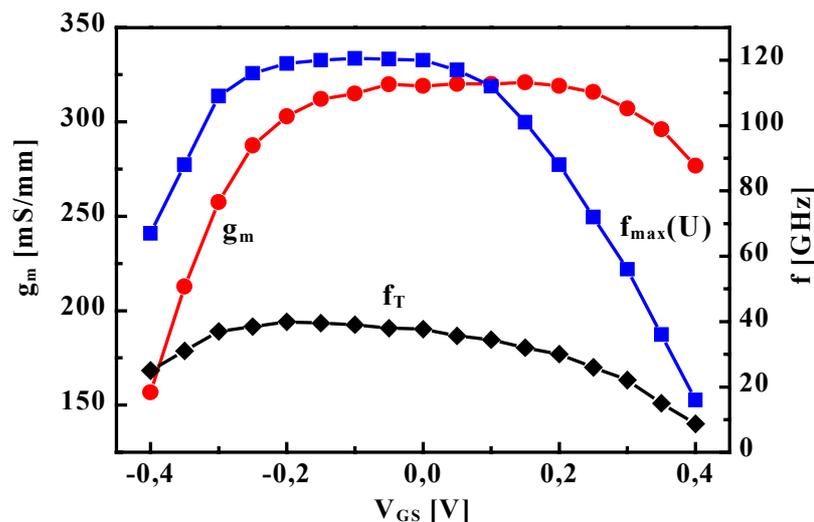


Abbildung 7.14: Abhängigkeit der Grenzfrequenzen f_T u. $f_{\max}(\text{U})$ und der Steilheit g_m vom Arbeitspunkt des Transistors V_{GS} bei $V_{DS} = 2,0$ V

Ein Vergleich zwischen dem typischen frequenzabhängigen Verstärkungsverlauf eines Transistors konventioneller Bauart (Abbildung 7.13) und eines Bauelementes mit selbstjustierender Technologie (Abbildung 7.15) zeigt ein deutlich angehobenes Verstärkungsniveau für $|\underline{h}_{21}|^2$ über den gesamten Frequenzbereich. So liegt die Kurzschlußstromverstärkung des selbstjustierten Transistors bei 5 GHz etwa 30 % über der des Standardbauelementes und bei 20 GHz sogar mehr als doppelt so hoch. Dieses hohe Niveau setzt sich bis zur Grenzfrequenz fort und führt schließlich zu einem f_T von 90 GHz. Bei der unilateralen Leistungsverstärkung dreht sich dieses Bild um, und der konventionelle Transistor erzielt deutlich höhere Verstärkungswerte. Dadurch stellt sich für dieses Bauelement der neuen Technologie nicht das für SiGe-HFETs typische $f_T/f_{\max}(U)$ Verhältnis von 1/2 - 1/3 ein, sondern eine Grenzfrequenz $f_{\max}(U)$ unterhalb von f_T . Außerdem zeigt der Transistor, im Gegensatz zu herkömmlichen Hochfrequenzbauelementen, keine potentielle Instabilität für niedrige Frequenzbereiche, sondern weist über den gesamten Meßbereich von 50 MHz bis 50 GHz einen Stabilitätsfaktor $k > 1$ auf. Um dieses untypische elektrische Verhalten aufzulösen, ist eine detaillierte Analyse der Ersatzschaltbildparameter notwendig, die nachfolgend in Abschnitt 7.2.2 durchgeführt wird.

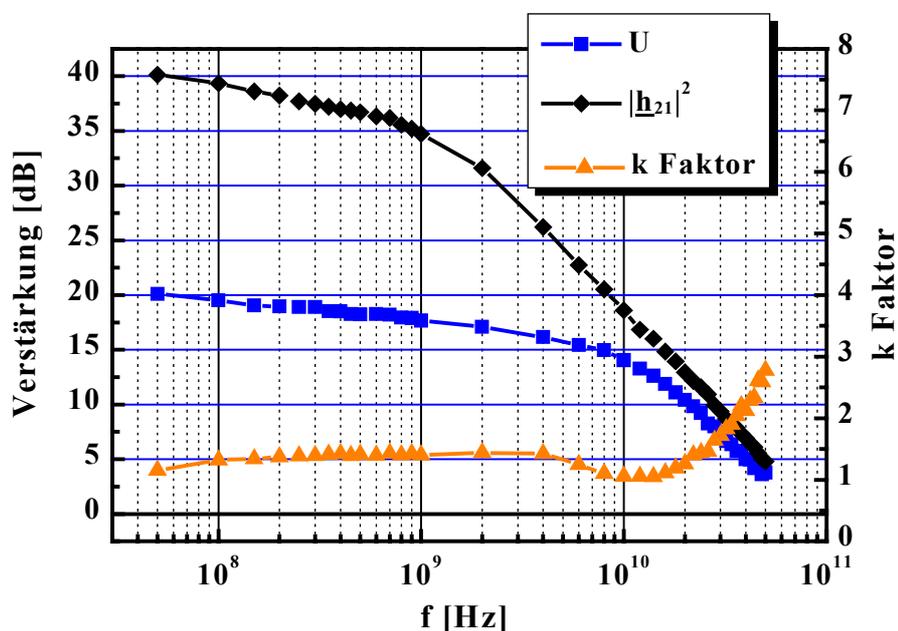


Abbildung 7.15: Frequenzabhängigkeit der Verstärkungskennzahlen U und $|\underline{h}_{21}|^2$ und des k Faktors für einen selbstjustierten SiGe-HFET mit einer Gatelänge von 90 nm für $V_{DS} = 2,0$ V und $V_{GS} = 0,2$ V

Abschließend soll ein quantitativer Vergleich, der im Rahmen dieser Arbeit erzielten Grenzfrequenzen f_T und f_{max} , mit den Bauelementen anderer Hersteller die objektive Einordnung der Ergebnisse erleichtern.

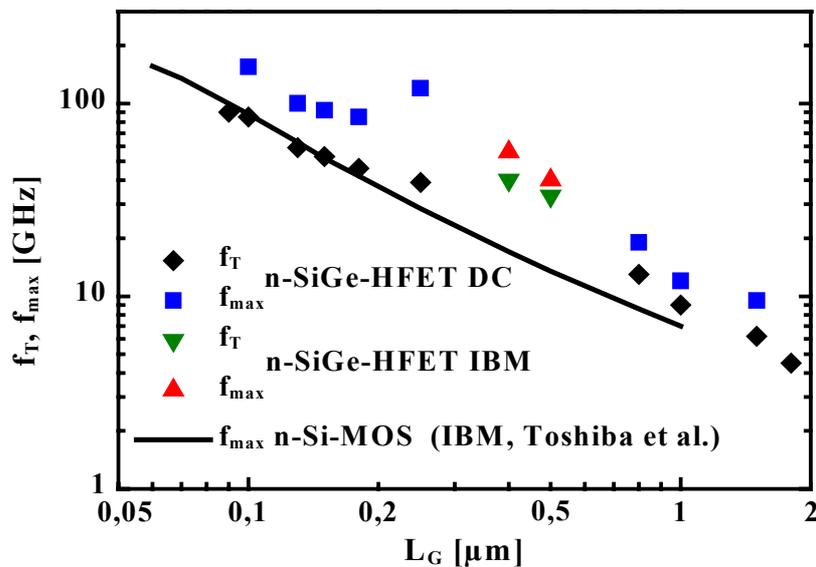


Abbildung 7.16: Entwicklung der Grenzfrequenzen f_T und f_{max} in Abhängigkeit von der physikalischen Gatelänge L_G . Vergleich zwischen den derzeit veröffentlichten Maximalwerten von IBM und DaimlerChrysler bei SiGe-HFETs und n-MOS-Transistoren auf Si

Abbildung 7.16 stellt die zeitliche Entwicklung der bisher veröffentlichten maximalen Grenzfrequenzen von SiGe-n-Kanal-HFETs und Standard Silizium-n-MOS-Transistoren unterschiedlicher Hersteller für verschiedene physikalische Gatelängen gegenüber [7.7, 7.8, 7.9, 7.10]. Dabei wird deutlich, daß sich die f_T -Grenzfrequenzen der SiGe-Heterobauelemente etwa auf dem Niveau der f_{max} -Werte der n-Si-MOS-Bauelemente bewegen. Ein direkter Vergleich der maximalen Oszillationsfrequenzen liefert einen Leistungsvorteil der SiGe-HFETs von ca. 30 - 40 %, was etwa einem Vorsprung von 1 - 2 Technologiegenerationen der SIA-Roadmap entspricht.

7.2.2 Simulation und Extraktion des Kleinsignalersatzschaltbildes

Die erfolgreiche Modellierung des Transistorverhaltens mit Hilfe eines Kleinsignalersatzschaltbildes legt den Grundstein für die Leistungsanalyse des Bauelementes, außerdem bildet sie die Basis für einen späteren Schaltungsentwurf und unterstützt die technologische Prozeßentwicklung. Als Grundlage für die Modellierung und Simulation der Transistorcharakteristik dient dabei das in Kapitel 3 eingeführte Kleinsignalersatzschaltbild (Abbildung 3.11). Dieses läßt sich in einen intrinsischen und einen extrinsischen Bereich unterteilen. Der intrinsische Teil umfaßt die spannungsabhängigen Elemente g_m , τ , g_{ds} , R_i , C_{gs} , C_{gd} und C_{ds} , und der extrinsische Bereich besteht aus den parasitären Elementen R_g , L_g , C_{pgs} , R_d , L_d , C_{pds} , R_s , L_s und C_{pgd} .

Die komplette Bestimmung aller Ersatzschaltbildparameter für einen diskreten Arbeitspunkt ist heute theoretisch ausschließlich mittels mathematischer Optimierungsverfahren möglich. Dabei werden die Größen der Elemente, ausgehend von einem vorgegebenen Startwert, solange variiert, bis die Abweichung zwischen gemessenen und simulierten S-Parametern minimal ist. Der Erfolg dieses Vorgehens ist dabei stark von der Wahl der Startwerte abhängig. Außerdem besteht die Gefahr, daß das Verfahren in einem lokalen Minimum mündet, ohne die optimale Lösung zu erreichen. Daraus ergibt sich dann eine falsche Ersatzschaltbildkonfiguration, die zudem nicht zwingend physikalisch sinnvoll sein muß. Ein erheblich effizienteres Konvergenzverhalten stellt sich ein, wenn die Anzahl der zu bestimmenden Elemente, z. B. durch zusätzliche Messungen, reduziert werden kann. Deshalb haben sich in den letzten Jahren verstärkt Methoden entwickelt, die versuchen, möglichst viele Elemente direkt durch Messungen zu bestimmen und dabei ganz auf Optimierungsverfahren zu verzichten, um statt dessen analytische Extraktionsstrategien zu verfolgen.

Die im Rahmen dieser Arbeit umgesetzte Strategie zur Bestimmung der verschiedenen Ersatzschaltbildelemente bedient sich unterschiedlicher elektrischer Meßkonfigurationen, um zuerst die parasitären, extrinsischen Parameter zu ermitteln und nutzt dann die bereinigten S-Parameter, um die intrinsischen Elemente analytisch aufzulösen.

Dazu müssen zuerst die kapazitiven und induktiven Einflüsse der Meßstruktur auf das eigentliche Bauelement erfaßt werden, um sie anschließend mittels Deembedding von den reinen Transistordaten zu separieren. Zu diesem Zweck

werden spezielle Open- und Short-Teststrukturen in das Layout mit eingebracht, um durch S-Parameter Messungen die parasitären Anteile zu bestimmen. Eine Open-Struktur besteht dabei aus der kompletten Meßstruktur des Transistors mit allen Pads und Zuleitungen, jedoch ohne aktives Bauelement (Abbildung 7.17). Die gemessenen Open-S-Parameter beinhalten alle kapazitiven Kopplungen zwischen den Kontaktpads und den Zuleitungen, berücksichtigen aber auch Leckströme über den nicht ideal isolierenden Puffer und das hochohmige Substrat. Durch die Umwandlung der S-Parameter in Y-Parameter und die Darstellung als passives Netzwerk werden alle diese Effekte in drei Admittanzen Y_1 , Y_2 und Y_3 berücksichtigt. Eine genaue Kenntnis über den detaillierten Aufbau dieser Elemente ist bei dieser Deembedding-Strategie nicht notwendig. Im ursprünglichen Ersatzschaltbild werden die Admittanzen Y_1 , Y_2 und Y_3 durch die Kapazitäten C_{pgs} , C_{pds} und C_{pgd} repräsentiert, da sie üblicherweise den Hauptanteil an der Open-Struktur haben. Im Rahmen der Extraktionsstrategie werden im ersten Schritt die gemessenen S-Parameter des Transistors ebenfalls in Y-Parameter transformiert und anschließend die Y-Parameter der Open-Struktur nach Real- und Imaginärteil subtrahiert (Abbildung 7.19). Damit läßt sich der parasitäre Anteil, der zum intrinsischen FET liegenden parallelen Elemente, eliminieren.

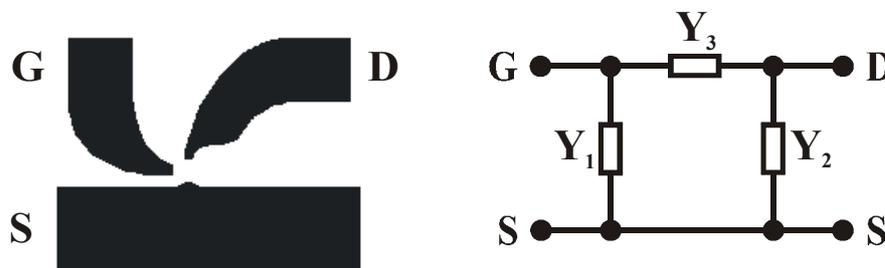


Abbildung 7.17: Layout und Ersatzschaltbild einer Open-Struktur

Im nachfolgenden Schritt sollen die parasitären Elemente, die in Reihe zum intrinsischen Bauelement liegen, erfaßt und entfernt werden. Dazu dient eine spezielle Short-Struktur, die der normalen Meßstruktur entspricht, jedoch an der Stelle des eigentlichen Transistors einen metallischen Kurzschluß der Drain-, Gate- und Source-Zuleitungen aufweist (Abbildung 7.18). Das sternförmige Ersatzschaltbild der Anordnung läßt sich sinnvollerweise durch drei Impedanzen Z_1 , Z_2 und Z_3 beschreiben. Sie berücksichtigen primär die parasitären Padinduk-

tivitäten L_s , L_d und L_g der Meßstruktur, aber auch den Metallwiderstand R_M der einzelnen Zuleitungskomponenten. Z_1 , Z_2 und Z_3 lassen sich aus den gemessenen S-Parametern der Short-Struktur durch Umwandlung in Z-Parameter gewinnen und werden anschließend in einem zweiten Extraktionsschritt von den Z-Parametern, des bereits um die Open-Struktur bereinigten Bauelementes abgezogen (Abbildung 7.19).

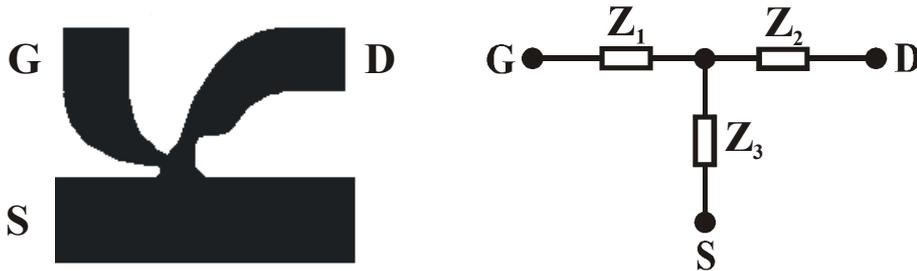


Abbildung 7.18: Layout und Ersatzschaltbild einer Short-Struktur

Um den intrinsischen Kern des Transistors zu erhalten, müssen die in Kapitel 7.1 meßtechnisch bestimmten parasitären Widerstände R_s , R_d und R_g aus dem verbliebenen Ersatzschaltbild eliminiert werden. Dazu bietet sich wieder die Z-Parameter Darstellung an. Nach der Subtraktion der Widerstandsanteile verbleibt das Ersatzschaltbild des inneren FETs in π -Topologie.

Die Y-Parameter der daraus resultierenden Admittanz-Matrix lassen sich durch analytische Ausdrücke in Abhängigkeit der Elemente des intrinsischen Transistors darstellen.

$$\underline{Y}_{11} = \frac{R_i C_{gs}^2 \omega^2}{1 + \omega^2 C_{gs}^2 R_i^2} + j\omega \left(\frac{C_{gs}}{1 + \omega^2 C_{gs}^2 R_i^2} + C_{gd} \right) \quad (7.18)$$

$$\underline{Y}_{12} = -j\omega C_{gd} \quad (7.19)$$

$$\underline{Y}_{21} = \frac{g_m \exp(-j\omega\tau)}{1 + jR_i C_{gs} \omega} - j\omega C_{gd} \quad (7.20)$$

$$\underline{Y}_{22} = g_{ds} + j\omega(C_{ds} + C_{gd}) \quad (7.21)$$

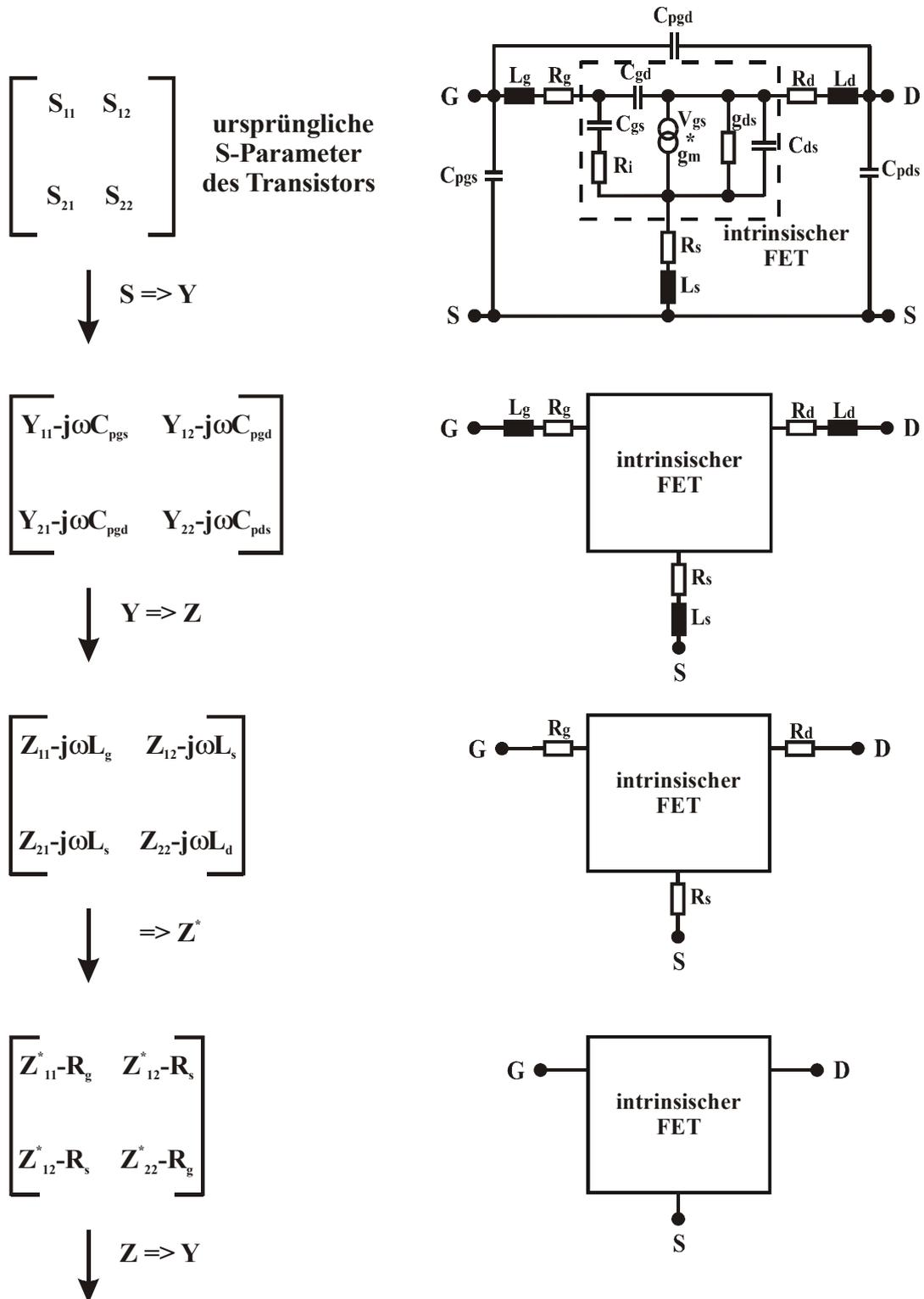


Abbildung 7.19: Angewandte Extraktionsmethode zur Separation der intrinsischen Y-Matrix des SiGe-HFETs

Aus den Real- und Imaginärteilen der Y-Parameter aus (7.18) - (7.21) ergeben sich die folgenden analytischen Zusammenhänge für die einzelnen Ersatzschaltbildelemente des inneren FETs.

$$C_{gd} = -\frac{1}{\omega} \cdot \text{Im}\{\underline{Y}_{12}\} \quad (7.22)$$

$$C_{ds} = \frac{1}{\omega} \cdot \text{Im}\{\underline{Y}_{22}\} - C_{gd} \quad (7.23)$$

$$C_{gs} = \frac{\text{Im}\{\underline{Y}_{11}\} - \omega C_{gd}}{\omega} \cdot \left(1 + \frac{(\text{Re}\{\underline{Y}_{11}\})^2}{(\text{Im}\{\underline{Y}_{11}\} - \omega C_{gd})} \right) \quad (7.24)$$

$$R_i = \frac{\text{Re}\{\underline{Y}_{11}\}}{(\text{Im}\{\underline{Y}_{11}\} - \omega C_{gd})^2 + (\text{Re}\{\underline{Y}_{11}\})^2} \quad (7.25)$$

$$g_m = \sqrt{((\text{Re}\{\underline{Y}_{21}\})^2 + (\text{Im}\{\underline{Y}_{21}\} + \omega C_{gd})^2) \cdot (1 + \omega^2 C_{gs}^2 R_i^2)} \quad (7.26)$$

$$g_d = \text{Re}\{\underline{Y}_{22}\} \quad (7.27)$$

$$\tau = \frac{1}{\omega} \cdot \arcsin\left(\frac{-\omega C_{gd} - \text{Im}\{\underline{Y}_{21}\} - \omega C_{gs} R_i \text{Re}\{\underline{Y}_{21}\}}{g_m}\right) \quad (7.28)$$

Zur Verifizierung der vorgestellten Extraktionsmethode und des verwendeten Ersatzschaltbildes sind in Abbildung 7.20 die experimentell ermittelten S-Parameter eines konventionellen SiGe-HFETs den Ergebnissen der Simulation gegenübergestellt. Die gute Übereinstimmung der Werte über den gesamten Frequenzbereich von 50 MHz bis 40 GHz ist zum einen ein Indiz für die Güte des eingesetzten Kleinsignalmodells und bestätigt zum anderen die gewählte Strategie zur Bestimmung der Ersatzschaltbildparameter.

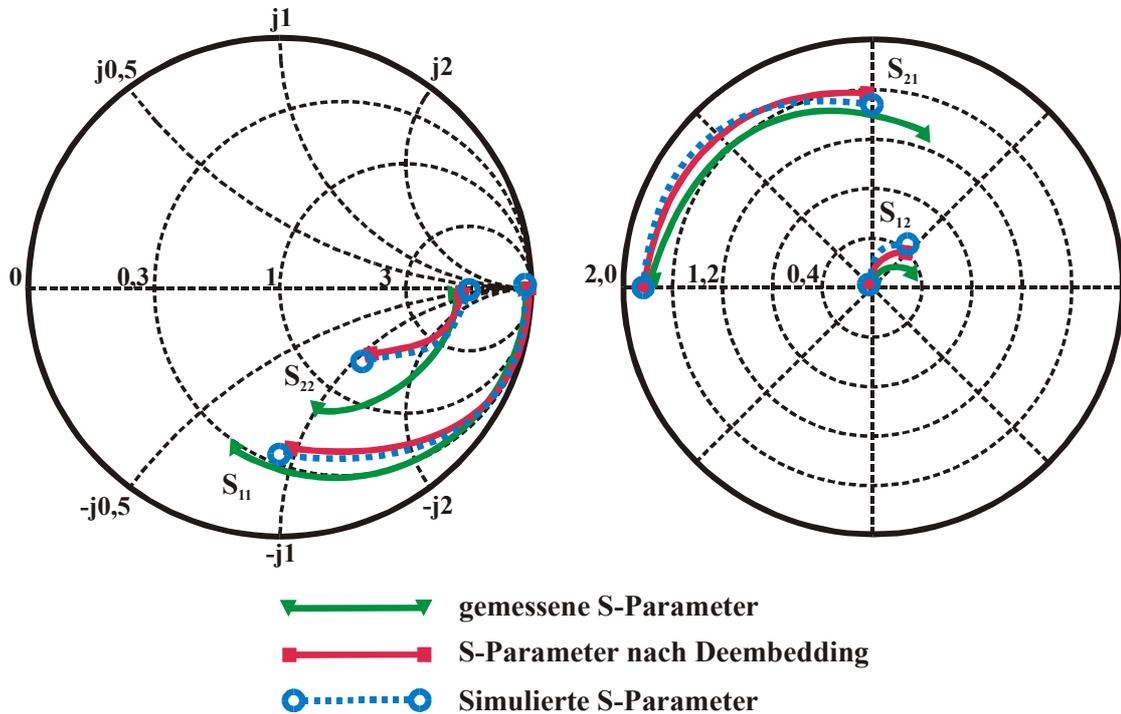


Abbildung 7.20: Vergleich zwischen den experimentell ermittelten und auf der Basis des vorgestellten Kleinsignalersatzschaltbildes simulierten S-Parametern eines konventionellen SiGe-HFETs im Frequenzbereich von 50 MHz bis 40 GHz

Die Anwendung dieser Simulationsstrategie auf den selbstjustierten Transistor aus Abbildung 7.15 scheitert jedoch, da keine zufriedenstellende Modellierung der gemessenen S-Parameter erzielt werden kann. Da die gewählte Extraktionsmethode ihre Gültigkeit bereits für den konventionellen Transistor bewiesen hat, muß das elektrische Verhalten des Bauelementes durch eine Verfeinerung des Ersatzschaltbildes approximiert werden. Schon bei der Betrachtung der experimentell ermittelten S-Parameter zeigt sich, insbesondere für den Ausgangsreflektionsfaktor S_{22} und den Vorwärtsübertragungsfaktor S_{21} , ein untypischer Verlauf (Abbildung 7.21). Dieses läßt darauf schließen, daß bei dem Transistor eine zusätzliche kapazitive Kopplung zwischen Ein- und Ausgang existiert, die sich im Hochfrequenzbetrieb auch auf die Kleinsignalverstärkung auswirkt.

Entsprechend dieser Interpretation der S-Parameter wird das Standardersatzschaltbild im intrinsischen Teil durch zwei Leitungspfade zwischen Drain und Source modifiziert, bestehend aus jeweils einer Kapazität und einem Widerstand (Abbildung 7.22). Durch die Erweiterung des Ersatzschaltbildes ergibt sich eine

verbesserte Modellierung der gemessenen S-Parameter (Abbildung 7.21), entsprechend der Ergebnisse des konventionellen HFETs.

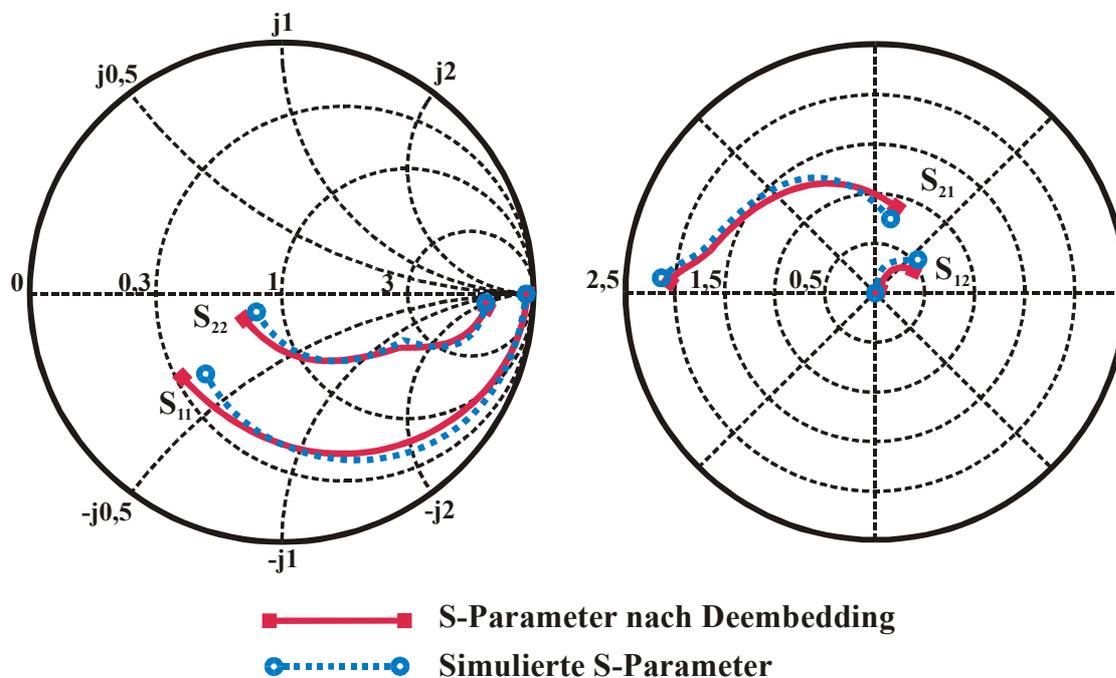


Abbildung 7.21: Gegenüberstellung der gemessenen S-Parameter und der durch Simulation des modifizierten Ersatzschaltbildes erzielten Ergebnisse für einen selbstjustierten SiGe-HFET im Frequenzbereich von 50 MHz bis 40 GHz

Physikalisch spiegeln die beiden RC-Glieder die kapazitive Kopplung zwischen den niederohmigen Source/Drain-Gebieten und der hochdotierten p+-Schicht wider. Diese Schicht wurde beim selbstjustierten SiGe-HFET zusätzlich in den Schichtaufbau mit aufgenommen, um Kurzkanaleffekte aufgrund der kurzen Gatelänge zu unterdrücken. Im Hochfrequenzbetrieb ergeben sich zwischen den Implantationsgebieten und der p+-Schicht zwei n-i-p-Diodenstrukturen, die aufgrund des unterschiedlichen Gleichspannungspotentials an Source und Drain verschieden große Sperrschichtkapazitäten ausbilden. In Abhängigkeit dieser Kapazitäten resultiert daraus ein Strompfad, der als Bypass zum eigentlichen Kanal fungiert und nicht über das Gate steuerbar ist.

Auf die Grenzfrequenz f_T hat dieser parasitäre Kanal keinen Einfluß, weil bei der Bestimmung von \underline{h}_{21} , durch den Kurzschluß am Ausgang des Transistors, dieser

Strompfad unwirksam ist. Für die Leistungsverstärkung jedoch bedeutet dies, daß nur ein reduzierter Anteil der Eingangsleistung über das eigentliche Bauelement verstärkt wird und somit zum Gewinn beiträgt, während ein nicht zu vernachlässigender Beitrag über den Bypass verloren geht. Deshalb weisen alle selbstjustierten SiGe-HFETs, die mit einer p+-Schicht zwischen Puffer und aktivem Schichtstapel entstanden sind, eine signifikant herabgesetzte Leistungsverstärkung und ein f_T/f_{\max} -Verhältnis von >1 auf.

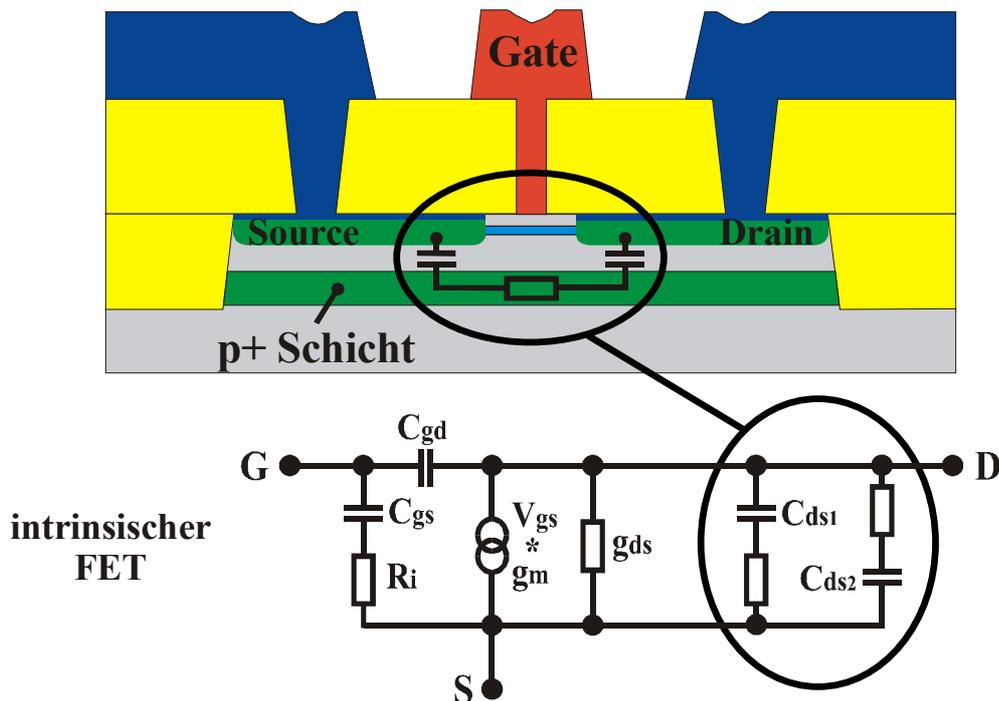


Abbildung 7.22: Modifiziertes intrinsisches Ersatzschaltbild für den selbstjustierten HFET und physikalische Bedeutung der RC-Glieder innerhalb des Bauelementes

Eine Bestätigung dieses Phänomens liefern Erfahrungen aus der III/V-Technologie. Beim Aufbau von GaN-HEMTs auf preisgünstigem Si-Substrat stellten Chumbes et al. ebenfalls eine reduzierte Leistungsverstärkung fest [7.11]. Aufgrund einer höheren Substratleitfähigkeit des Siliziums im Vergleich zu Saphir oder SiC sorgte die kapazitive Kopplung zwischen dem aktiven Bauelement und dem Si-Substrat für niedrigere Grenzfrequenzen f_{\max} als mit Hilfe der Simulation vorausgesagt.

7.2.3 Analyse des Rauschverhaltens

Neben der Forderung nach exzellenten Verstärkungscharakteristiken im Mikrowellenbereich und hohen Grenzfrequenzen spielen Rauscheigenschaften für Hochfrequenztransistoren eine wichtige Rolle. Insbesondere die Anwendung als rauscharmer Verstärker im Empfangsteil analoger Kommunikationssysteme stellt erhöhte Ansprüche an das Übertragungsverhalten der Bauelemente. Um die Linearität der Verstärkerstufe zu gewährleisten, müssen die Nutzsignale am Eingang des Transistors die Kleinsignalbedingungen erfüllen, d. h. kleine Signalamplituden aufweisen. Während des Verstärkungsprozesses überlagert sich das bauelementspezifische elektronische Rauschsignal des HFETs dem einfallenden Nutzsignal und ruft zusätzliche Störungen hervor. Dadurch ergibt sich ein verminderter Störabstand am Ausgang des Verstärkers. Deshalb sind die Rauscheigenschaften eines Transistors ein wichtiges Gütekriterium, insbesondere bei analogen Verstärkeranwendungen.

Das Hochfrequenzrauschen in einem FET wird im wesentlichen von zwei Quellen verursacht, den inneren Rauschquellen des intrinsischen Transistors und den äußeren thermischen Quellen, welche sich aus den parasitären Elementen ergeben. Dabei handelt es sich primär um die Widerstandsanteile von Source, Drain und Gate. Die physikalische Ursache des Widerstandsrauschens liegt in der regellosen Wärmebewegung der im Widerstandsbauelement befindlichen frei verschiebbaren Ladungsträger [7.12]. Die spektrale Leistungsdichte des thermischen Rauschens hängt von der Größe des Wirkwiderstandes R und der Temperatur T ab, ist aber von der Frequenz unabhängig und wird deshalb auch als „weißes“ Rauschen bezeichnet.

Die intrinsischen Anteile des Rauschens haben ihren Ursprung in der regellosen Fluktuation der Ladungsträgeranzahl und -geschwindigkeit und den daraus resultierenden Schwankungen in den Strom- und Spannungsgrößen. Die Hauptquellen des Rauschens ergeben sich dabei durch den Ladungsträgertransport im Kanal und über die nicht ideale Schottkybarriere des Gates. Die lineare Strom-Spannungsbeziehung im Kanal für kleine Drain/Source-Spannungen erlaubt eine Betrachtung als thermische Rauschquelle. Für große longitudinale Feldstärken und Ladungsträgergeschwindigkeiten nahe der Sättigungsgeschwindigkeit tritt Diffusionsrauschen in den Vordergrund [7.13], [7.14]. Zusätzlich kommt es durch die nicht kontinuierliche Ladungsbewegung über Potentialbarrieren zum sogenannten Schrotrauschen [7.15]. Im FET treten diese Bereiche z. B. an sper-

renden pn-Übergängen oder in der Raumladungszone unterhalb des Gates auf. Die kapazitive Kopplung zwischen der Gateelektrode und dem Kanalbereich des Transistors über die Kapazitäten C_{gs} und C_{gd} führt zu Rauschanteilen in der Steuerelektrode. Dabei induzieren die auftretenden Spannungsfluktuationen des Kanals einen frequenzabhängigen Rauschstrom im Gate, der wiederum auf die Ausgangsgrößen des Transistors zurückwirkt. Die Störgrößen im Kanal und im Gate sind miteinander korreliert, weil sie auf denselben Ursprung zurückgehen.

Die Ermittlung der vier relevanten Rauschparameter eines Transistors F_{min} , R_n , G_{ass} und Γ_{opt} kann entweder direkt durch entsprechende Rauschmessungen erfolgen oder unter Zuhilfenahme von empirischen oder physikalischen Modellen. Etablierte empirische Methoden zur Extraktion der Rauschparameter gehen auf Gupta [7.16] und Fukui [7.17], [7.18] zurück. Sie nutzen sowohl Messungen zur Ermittlung bestimmter Rauschkennzahlen als auch verstärkt Informationen über die Elemente des Kleinsignalersatzschaltbildes. Die dabei abgeschätzten Rauschparameter stimmen bis zu einer Frequenz von 20 GHz gut mit den experimentellen Werten überein. Physikalisch basierte Modelle weisen im Gegensatz zu empirischen Modellen eine komplexere Struktur auf und erfordern oftmals detaillierte Informationen über die Bauelementgeometrie und die Materialeigenschaften. Mit Hilfe einer formalen Beschreibung der physikalischen Prinzipien erfolgt die Lösung auf numerischem Weg, mit einer guten Annäherung an die gemessenen Werte über den gesamten Frequenzbereich. Auf das weite Feld der Rauschmodellierung soll im Rahmen dieser Arbeit nicht detaillierter eingegangen werden, einen interessanten Einstieg in diese Thematik bieten jedoch die folgenden Literaturstellen [7.19, 7.20, 7.21, 7.22].

Die Rauschanalyse der SiGe-HFETs wurde im Rahmen des europäischen Verbundprojektes „SIGMUND“ in Zusammenarbeit mit dem Institut d'Electronique Fondamentale der Université Paris Sud in Frankreich durchgeführt. Bei dieser Kooperation entstanden die ersten bisher veröffentlichten, experimentell ermittelten Rauschcharakteristiken für SiGe-n-Kanal-HFETs (Abbildung 7.23) [7.23], [7.24]. Basierend auf einer neuen Meßmethode nach Danelon und Crozat [7.25] erfolgt die Bestimmung der relevanten Rauschparameter für nur zwei feste Eingangsimpedanzen unter Ausnutzung der Variation der Rauschzahl F mit der Frequenz. Als Meßobjekt kam dabei ein modulationsdotierter SiGe-Transistor konventioneller Bauart zum Einsatz, der bereits über ein selbstjustierendes Feldoxid verfügt und ein elektronenstrahldefiniertes T-Gate mit 130 nm Gatelänge aufweist.

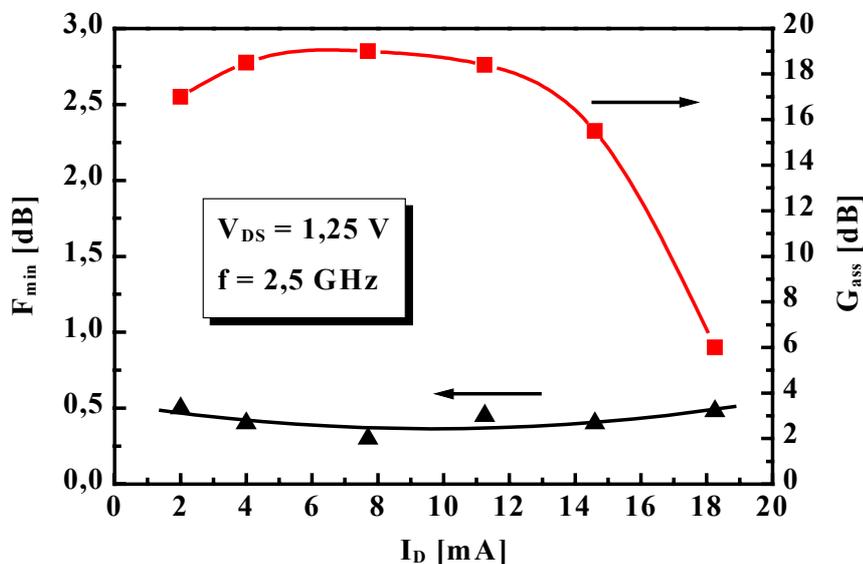


Abbildung 7.23: Minimale Rauschzahl F_{\min} und zugehörige Verstärkung G_{ass} eines SiGe-HFETs konventioneller Bauart in Abhängigkeit des Drainstromes I_D für $V_{\text{DS}} = 1,25$ V und $f = 2,5$ GHz

Abbildung 7.23 zeigt die minimale Rauschzahl F_{\min} und die zugehörige Verstärkung G_{ass} in Abhängigkeit vom Drainstrom I_D im rauschangepaßten Arbeitspunkt bei $V_{\text{DS}} = 1,25$ V und $f = 2,5$ GHz. F_{\min} weicht dabei über den gesamten Arbeitsbereich nur geringfügig von seinem Minimalwert bei 7,77 mA ab, was insbesondere bei analogen Verstärkeranwendungen, mit der Forderung nach hoher Linearität, von großer Bedeutung ist. Das leichte Ansteigen der minimalen Rauschzahl für niedrige und hohe Drainströme korreliert mit der verfügbaren Verstärkung G_{ass} , die in diesen Bereichen abnimmt. Für hohe Ausgangsströme I_D ist dafür der ansteigende Gatestrom verantwortlich, der sich durch eine zunehmend in Vorwärtsrichtung betriebene Schottky-Diode ergibt. Zwischen 3 mA und 12 mA weist die Verstärkung einen nahezu konstanten Bereich mit einem hohen Maximalwert von 19 dB bei 7,77 mA auf.

Ein weiterer wichtiger Parameter zur Charakterisierung der Rauscheigenschaften des Transistors ist der äquivalente Rauschwiderstand R_n . Ziel ist es, einen möglichst geringen R_n -Wert zu realisieren, weil dadurch die Voraussetzungen für ein niedriges F_{\min} und ein hohes G_{ass} gegeben sind. Der Verlauf von R_n in Abhängigkeit von I_D ist dabei primär durch die Steilheit g_m geprägt (Abbildung 7.24). Für kleine Ausgangsströme liegt R_n relativ hoch, da die Steilheit noch sehr gering ist.

Mit zunehmender Steilheit fällt R_n auf seinen Minimalwert von 40Ω ab, um dann aufgrund eines proportional zu I_D ansteigenden Kanalrauschens wieder anzuwachsen.

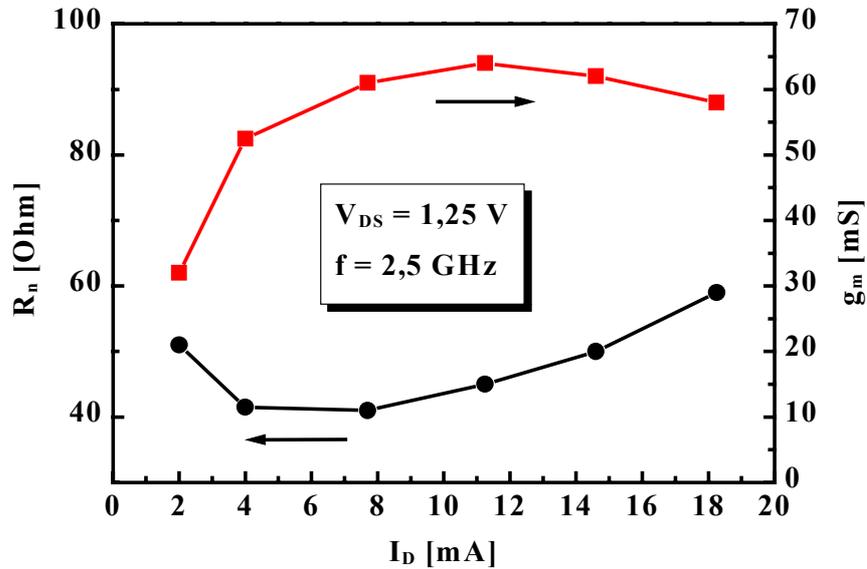


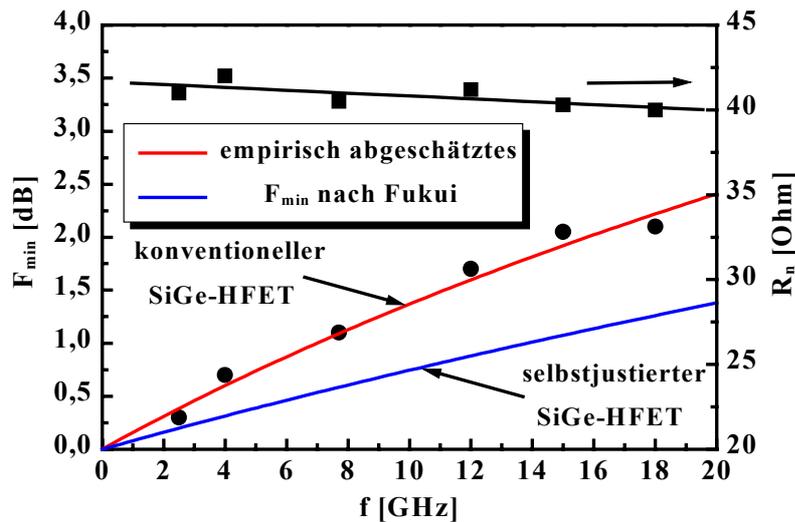
Abbildung 7.24: Abhängigkeit des äquivalenten Widerstandes R_n und der Steilheit g_m vom Drainstrom I_D

Für Hochfrequenzanwendungen ist oft der Frequenzgang der Rauschparameter von großem Interesse. Für den optimalen Gleichstromarbeitspunkt $I_D = 7,77 \text{ mA}$ und $V_{DS} = 1,25 \text{ V}$ zeigt Abbildung 7.25 die gemessene Frequenzabhängigkeit für die minimale Rauschzahl F_{\min} und den äquivalenten Widerstand R_n . Das niedrige und nahezu konstante Niveau von R_n läßt sich auf den niedrigen Gatewiderstand R_g und die hohe intrinsische Steilheit $g_{mi} = 650 \text{ mS/mm}$ des Transistors zurückführen. Die gemessenen F_{\min} -Werte starten bei einem Minimalwert von $0,3 \text{ dB}$ bei $2,5 \text{ GHz}$ und weisen einen typischen Anstieg mit der Frequenz auf. Das untermauert die zusätzlich dargestellte empirische Abschätzung der minimalen Rauschzahl nach Fukui (Gleichung 7.29) [7.18].

$$F_{\min} = 1 + 2\pi f \cdot K \cdot C_{gs} \cdot \sqrt{\frac{R_g + R_s}{g_{mi}}} \quad (7.29)$$

Unter Einbeziehung der extrahierten Ersatzschaltbildparameter C_{gs} , R_g und R_s für den betrachteten Arbeitspunkt ergibt sich eine sehr gute Übereinstimmung mit den gemessenen minimalen Rauschzahlen. K ist dabei ein material- und geome-

trieabhängiger Anpassungsfaktor, der für GaAs-MESFETs im Bereich von 2,2 - 2,5 liegt und für GaAs-HEMTs bei 1,5 - 1,8. Mit einem ermittelten K-Wert von 2,1 positioniert sich der SiGe-HFET zwischen diesen rauscharmen Hochfrequenzbauelementen.



Semi-quantitative Abschätzung des F_{\min} nach Fukui ($K = 2,1$):
 konventioneller SiGe-HFET: $V_{DS} = 1,25$ V; $I_D = 7,77$ mA
 Parameter: $C_{gs} = 130$ fF; $R_g = 6,5$ Ω ; $R_s = 18$ Ω ; $g_{mi} = 0,0525$ S
 selbstjustierter SiGe-HFET: $V_{DS} = 2,0$ V; $I_D = 6,9$ mA
 Parameter: $C_{gs} = 80$ fF; $R_g = 2,5$ Ω ; $R_s = 8,5$ Ω ; $g_{mi} = 0,035$ S

Abbildung 7.25: Minimale Rauschzahl F_{\min} und äquivalenter Rauschwiderstand R_n als Funktion der Frequenz f im Arbeitspunkt $V_{DS} = 1,25$ V, $I_D = 7,77$ mA. Vergleich der gemessenen F_{\min} -Werte mit der empirischen Abschätzung nach Fukui für einen Transistor konventioneller Bauart und einen selbstjustierten SiGe-HFET

Aufgrund der starken Abhängigkeit der minimalen Rauschzahl F_{\min} von der intrinsischen Steilheit g_{mi} und den Widerständen R_s , R_g liegt die Vermutung nah, daß sich die verbesserten elektrischen Eigenschaften des selbstjustierten SiGe-HFETs auch positiv auf das Rauschverhalten des Transistors auswirken. Eine semi-quantitative Abschätzung von F_{\min} nach Fukui für einen SiGe-HFET neuer Bauart ist in Abbildung 7.25, vergleichend zu den gemessenen Werten des konventionellen Transistors dargestellt. Basierend auf den extrahierten Kleinsigna-

lersatzschaltbildparametern und dem ermittelten K-Wert von 2,1 zeigt sich ein deutlich flacherer Verlauf für das lateral optimierte Bauelement. Dabei ist zu erwarten, daß sich die minimale Rauschzahl F_{\min} infolge der signifikant reduzierten Widerstände und der höheren Steilheit um bis zu 40 % verringert.

Um die Qualität der erzielten Ergebnisse einordnen zu können, ist in Abbildung 7.26 ein Vergleich zwischen verschiedenen Bauelementtypen und Materialsystemen dargestellt. Der konventionelle SiGe-n-Kanal-HFET plaziert sich dabei, deutlich abgesetzt von den Standard-Si-Bauelementen, oberhalb der III/V-FETs. Die minimalen Rauschzahlen bis zu einer Frequenz von 2,5 GHz liegen sogar noch unterhalb des GaAs-MESFETs. Das entspricht der qualitativen Einordnung des Bauelementes über den K-Faktor der Fukui-Formel. Bei höheren Frequenzen entwickeln sich die gemessenen Werte jedoch steiler nach oben als bei den Vergleichstransistoren. Hier führt der bereits angesprochene Einfluß der nicht optimierten parasitären Widerstandsanteile von R_s und R_g zu einer signifikanten Verschlechterung der Rauscheigenschaften. Gleichzeitig weisen die niedrigen Rauschzahlen unterhalb von 2,5 GHz auf das hohe Leistungspotential des Bauelementes hin, das in zukünftigen Untersuchungen und Experimenten durch Einsatz des selbstjustierenden Integrationsverfahrens weiter optimiert werden muß.

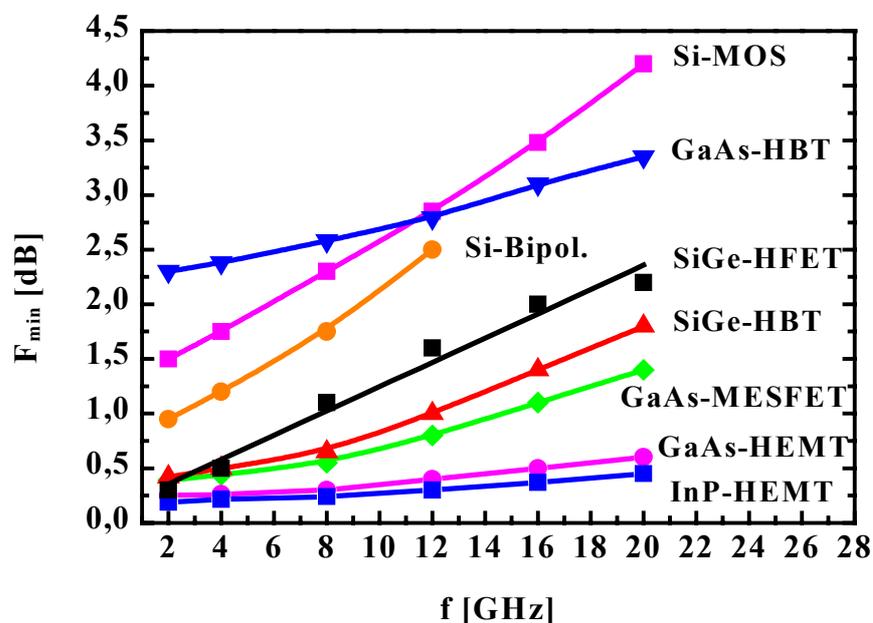


Abbildung 7.26: Vergleich der minimalen Rauschzahlen F_{\min} für verschiedene Transistortypen in den Materialsystemen Si, SiGe, GaAs u. InP

7.3 Das Leistungspotential der Transistoren bei kryogenischen Temperaturen

Die Charakterisierung der SiGe-HFETs bei kryogenischen Temperaturen liefert nicht nur Informationen über das Leistungspotential der Bauelemente bei Tieftemperaturanwendungen, z. B. beim Einsatz im Weltall, sondern trägt primär dazu bei, die Qualität der Si/SiGe-Heterostruktur und des Schichtwachstums zu beurteilen. Abbildung 7.27 zeigt die Temperaturabhängigkeit der Elektronenbeweglichkeit μ im 2DEG eines modulationsdotierten SiGe-Heteroschichtsystems, das sowohl auf einem „constant composition“-Puffer als auch auf einem gradierten Puffer realisiert wurde [7.26].

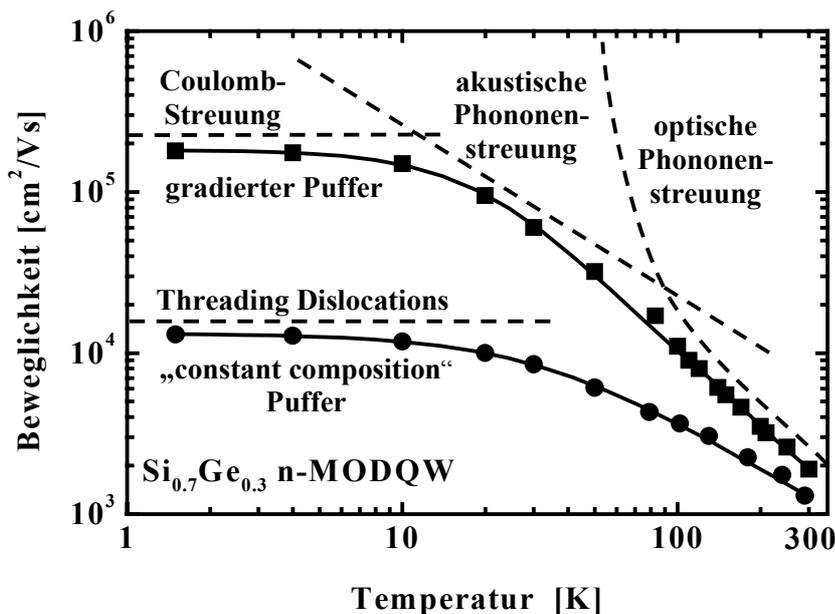


Abbildung 7.27: Temperaturabhängigkeit der Beweglichkeit μ von SiGe-Heterostrukturen auf einem „constant composition“-Puffer und einem gradierten Puffer mit einem Germaniumgehalt $x = 30\%$

Bei Raumtemperatur und damit dem überwiegend relevanten Arbeitsbereich der Transistoren zeigt sich die optische Phononenstreuung als beweglichkeitsdominierender Faktor für den Ladungsträgertransport im Si-Kanal des Bauelementes. Darüber hinaus führen nicht optimal ausgelegte Modulationsdotierungen zu para-

sitären Kanälen in den Dotierschichten und damit zu einer zusätzlichen Einschränkung der Gesamtbeweglichkeit. Im Temperaturbereich von 10 - 80 K limitieren überwiegend akustische Phononenstreuungsprozesse den Ladungsträgertransport. Unterhalb von 10 K tritt für die modulationsdotierte SiGe-Heterostruktur auf dem gradierten Puffer die Coulomb-Wechselwirkung in den Vordergrund, die in Abhängigkeit von der Spacerdicke maßgeblich die Beweglichkeit des 2DEGs bestimmt. Außerdem spielt, insbesondere für Bauelemente mit geringer Ladungsträgerdichte, die Hintergrunddotierung des Kanals eine nicht zu vernachlässigende Rolle. Die auf dem „constant composition“-Puffer gewachsene Heterostruktur weist im Vergleich zum gradierten Puffer eine um zwei Größenordnungen höhere Defektdichte von $2 \cdot 10^9 \text{ cm}^{-2}$ auf. Dabei treten vornehmlich aufsteigende Versetzungslinien auf, die vom Puffer bis in den Kanal reichen und dort den Stromfluß stören. Deshalb limitiert in diesem Fall primär die minderwertige Kristallqualität die Niedertemperaturbeweglichkeit des 2DEGs.

Alle hier dargestellten parasitären Einflüsse auf die Transporteigenschaften der Ladungsträger treten bei tiefen Temperaturen, aufgrund der signifikant reduzierten Phononenstreuung, verstärkt in den Vordergrund. Mit Hilfe der kryogenischen Gleichstrom- und Hochfrequenzeigenschaften des Bauelementes kann deshalb eine Aussage über den Einfluß dieser Effekte auf die Leistungsfähigkeit des Transistors getroffen werden.

Die bisher durchgeführten Betrachtungen bezogen sich ausschließlich auf die Heterostruktur des Bauelementes und betreffen damit primär den intrinsischen Transistor. Die Extraktion der Kleinsignalersatzschaltbildelemente aus S-Parametermessungen bei kryogenischen Temperaturen ermöglicht eine gleichzeitige Analyse der in- und extrinsischen Bauelementfaktoren und verdeutlicht die Auswirkungen auf die Gleichstrom- und Hochfrequenzeigenschaften des HFETs.

Abbildung 7.28 vergleicht den Verlauf der intrinsischen Steilheit g_{mi} und der Zeitkonstante τ eines SiGe-n-Kanal-HFETs bei Raumtemperatur und bei 50 K für verschiedene Arbeitspunkte V_{GS} . Als Meßobjekt diente dabei ein Transistor konventioneller Bauart, mit einer Gatelänge von 250 nm [7.27]. Deutlich zeigt sich für das Maximum der Steilheit ein drastischer Anstieg von 80 % bei 50 K. Gleichzeitig verringert sich τ mit zunehmender Gate/Source-Spannung um bis zu 50 %. Die physikalische Ursache dieser Ergebnisse liegt eindeutig in der erhöhten Niedertemperaturbeweglichkeit der Ladungsträger. Dabei bewirkt die größere

Beweglichkeit zum einen eine höhere Driftgeschwindigkeit der Elektronen und führt dadurch, bei gleicher äußerer Spannung V_{DS} , zu einem zunehmenden Ausgangsstrom I_D , zum anderen verkürzen sich die Neuverteilungszyklen der Ladungsträger in den Raumladungszonen und ermöglichen dadurch schnellere Schaltvorgänge.

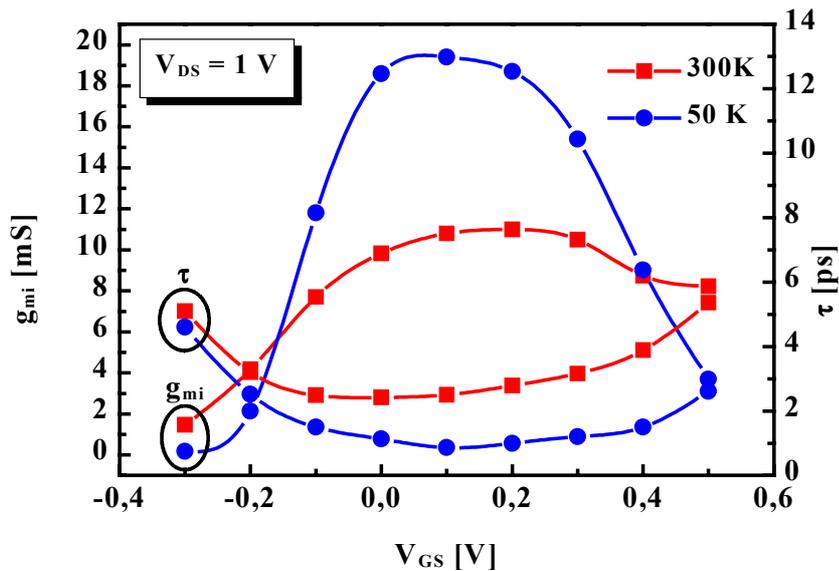


Abbildung 7.28: Verlauf der intrinsischen Steilheit g_{mi} und der Zeitkonstante τ bei 300 K und 50 K für verschiedene Arbeitspunkte V_{GS}

Die extrinsischen Widerstände R_s , R_d und R_g werden, wie schon in Abschnitt 7.2.2, als konzentrierte Elemente angenommen. Die dabei ermittelten temperaturbedingten Veränderungen in den Widerstandswerten sind in Tabelle 7.2 zusammengefasst.

| | R_s [Ω] | R_d [Ω] | R_g [Ω] |
|--------------|--------------------|--------------------|--------------------|
| 300 K | 8 | 30 | 2,5 |
| 50 K | 6,2 | 13,2 | 1,2 |
| Δ [%] | -22,5 | -56 | -52 |

Tabelle 7.2: Extrahierte extrinsische Widerstände R_s , R_d und R_g des SiGe-HFETs bei 300 K und 50 K [7.28]

Die bisher gewonnenen Erkenntnisse über die Ersatzschaltbildparameter bei kryogenischen Temperaturen bestätigen sich im Ausgangskennlinienfeld des Transistors (Abbildung 7.29). Der durch die Niederfeldbeweglichkeit μ_0 und die parasitären Widerstände R_s und R_d bestimmte lineare Anlaufbereich zeigt eine deutlich steilere Steigung als bei Raumtemperatur. Aufgrund der höheren Driftgeschwindigkeit der Ladungsträger verschiebt sich zudem die Kniespannung von +0,7 V auf +0,5 V. Im Sättigungsbereich erhöht sich die maximale Stromdichte um etwa 60 % auf über 500 mA/mm, gleichzeitig ist außerdem die gesteigerte Steilheit im Vergleich zur Messung bei 300 K erkennbar.

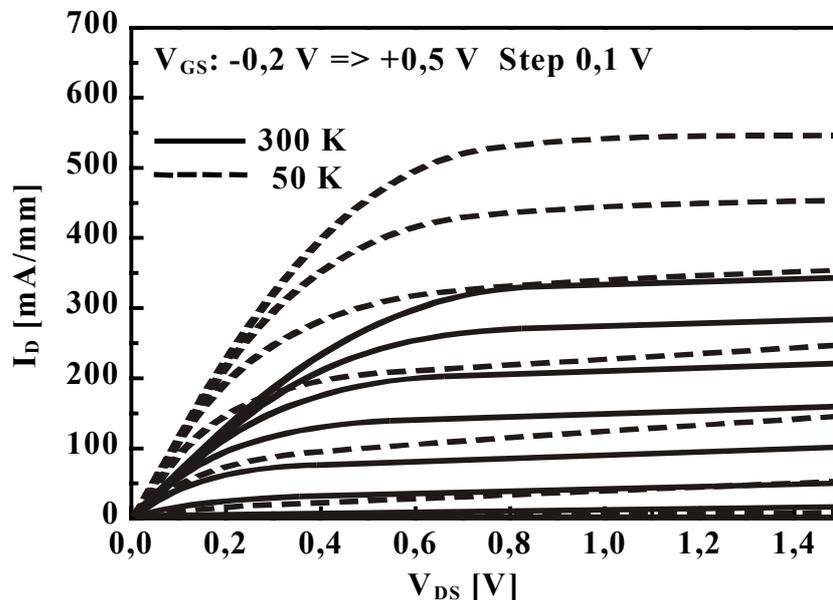


Abbildung 7.29: Ausgangskennlinienfeld eines SiGe-HFETs bei Raumtemperatur (300 K) und bei 50 K

Zur Analyse der Hochfrequenzeigenschaften ist das Verhalten weiterer Ersatzschaltbildparameter von Interesse. Für die Bestimmung der Transitfrequenz f_T spielen neben der Steilheit g_m , die Kapazitäten C_{gs} und C_{gd} eine wichtige Rolle. Unabhängig vom Arbeitspunkt V_{GS} ergibt sich für C_{gd} bei 50 K nur eine marginale Abweichung von den Raumtemperaturwerten. C_{gs} hingegen bewegt sich bei 50 K für Eingangsspannungen kleiner als -0,1 V bis zu 45 % unter den 300 K Werten und für höhere V_{GS} bis zu 20 % darüber (Abbildung 7.30).

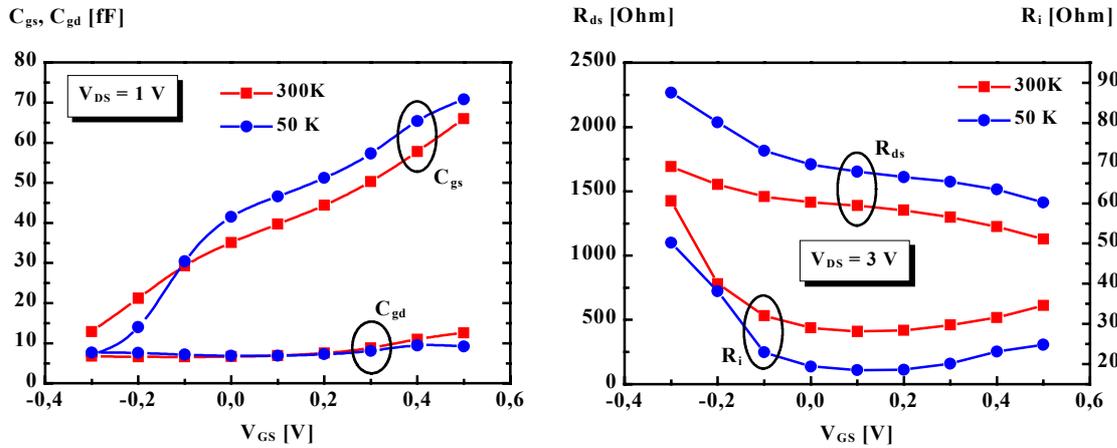


Abbildung 7.30: Verlauf der extrahierten Ersatzschaltbildparameter C_{gs} , C_{gd} , R_{ds} und R_i bei Raumtemperatur und 50 K in Abhängigkeit vom Arbeitspunkt V_{GS} [7.27]

In Verbindung mit der Steilheit g_m resultiert daraus ein Anstieg der Transitfrequenz f_T um über 70 % von 33 GHz bei 300 K auf 57 GHz bei 50 K (Abbildung 7.31). Der Verlauf der Grenzfrequenz wird dabei von der Steilheit dominiert, während die Kapazitäten ihren Einfluß nur wenig geltend machen können.

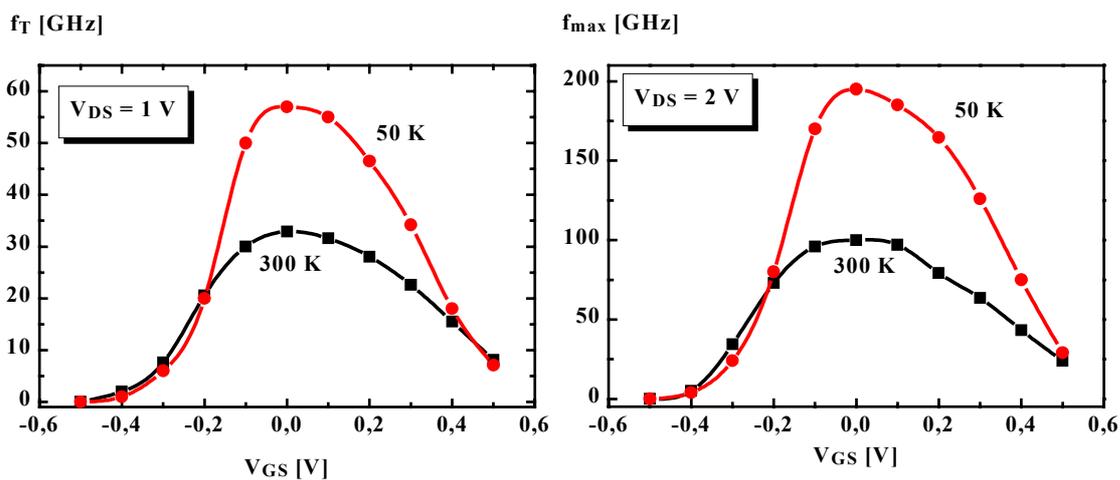


Abbildung 7.31: Einfluß der Temperatur auf die Grenzfrequenzen f_T und $f_{max}(U)$ für verschiedene Arbeitspunkte V_{GS} [7.27]

Die maximale Schwingfrequenz $f_{\max}(U)$ weist zusätzlich noch die Abhängigkeit von den Widerständen des Ersatzschaltbildes auf. Da R_s , R_d und R_g in erster Näherung als arbeitspunktunabhängig angesehen werden können, fällt das Augenmerk hier auf den Verlauf von R_{ds} und R_i . Die Zunahme von R_{ds} spiegelt sich im Sättigungsverhalten der Ausgangskennlinie wider und sorgt für den angestrebten niedrigen Ausgangsleitwert. Der reduzierte Wert von R_i bei niedrigen Temperaturen führt zu einem erhöhten Spannungsabfall über C_{gs} am Eingang des Transistors und resultiert dadurch direkt in einen ansteigenden Kanalstrom. Das Zusammenspiel aller Veränderungen bei 50 K sorgt letztlich für eine Zunahme der maximalen Schwingfrequenz $f_{\max}(U)$ von 100 GHz auf 195 GHz.

Kapitel 8

Zusammenfassung und Ausblick

Die enorme Datenflut heutiger und zukünftiger Telekommunikationssysteme stellt hohe technische Anforderungen an die mikroelektronischen Systemkomponenten und integrierten Schaltungen, die für die Verarbeitung dieser Informationen eingesetzt werden. Die Einzeltransistoren als Herzstück jeder integrierten Schaltung bestimmen dabei maßgeblich die Güte des Systems. Hohe Schaltfrequenzen, verbesserte Leistungsverstärkungen, geringe Rauschanteile, minimaler Leistungsverbrauch und niedrige Stückkosten sind die zentralen Qualitätskriterien an heutige Hochfrequenzbauelemente.

Das ideale Konzept für einen Feldeffekttransistor zur Anwendung im Hochfrequenzbereich vereinigt die brillanten Transporteigenschaften von Heterostruktur-Feldeffekttransistoren auf der Basis von Verbindungshalbleitern, mit den optimierten und hochintegrierten Herstellungsverfahren der Si-CMOS-Technologie. Ein vielversprechendes Potential zur Realisierung dieses Konzeptes bietet der SiGe-Hetero-Feldeffekttransistor. An SiGe-Heterostruktursystemen mit kristallografisch verspannten Si-Kanälen konnten bisher maximale Elektronenbeweglichkeiten von bis zu $2830 \text{ cm}^2/\text{Vs}$ experimentell nachgewiesen werden, und Simulationen gehen sogar von Werten bis zu $3490 \text{ cm}^2/\text{Vs}$ aus. Gleichzeitig lassen sich mit dem Prinzip der Modulationsdotierung hohe Ladungsträgerdichten von bis zu $7 \cdot 10^{12} \text{ cm}^{-2}$ im 2DEG des Kanals erzielen. Zusammen mit einer stetig steigenden Kristallqualität der aktiven Schichten sind damit die Grundlagen für kurze Ladungsträgerlaufzeiten und hohe Grenzfrequenzen gelegt. Auf der prozeßtechnischen Seite weist das SiGe-Materialsystem eine hohe Kompatibilität zur Silizium-Technologie auf, so daß Herstellungsabläufe, Bauelementkonzepte und Integrationsverfahren in großem Rahmen übertragbar sind. Ein Beweis dafür stellt die SiGe-BiCMOS-Technologie dar, die Standard-CMOS-Transistoren mit SiGe-HBTs in einem Herstellungsprozeß vereinigt.

Aus diesen vielversprechenden Fakten resultierten direkt die Ziele dieser Arbeit:

1. Die vertikale Optimierung des modulationsdotierten SiGe-Heteroschichtsystems, mit einer klaren Ausrichtung des Transistors zu Hochfrequenzanwendungen.
2. Die Verbesserung der Lateralstruktur des SiGe-HFETs durch Modifikation der bestehenden Prozeßtechnologie.
3. Entwurf und Entwicklung eines neuen selbstjustierenden Herstellungskonzeptes zur Umsetzung höherer Integrationsdichten, geringerer Produktionskosten und eines gesteigerten Leistungspotentials des Transistors.
4. Realisierung einer Prozeßtechnologie mit einem hohen Maß an Integrationsfähigkeit in spätere Schaltungsanwendungen und möglichst enger Kompatibilität zu einem Standard-CMOS-Verfahren.

Der Auslegung und Dimensionierung des heterogenen Schichtsystems kommt ein bedeutender Schwerpunkt bei der Optimierung des SiGe-HFETs zu. Als Basis für die aktiven Schichten hat die SiGe-Pufferschicht großen Einfluß auf die kristallografische Güte des verspannten Si-Kanals. Im Rahmen dieser Arbeit konnte bestätigt werden, daß sich durch die Einführung des gradierten SiGe-Puffers die Dichte der aufsteigenden Versetzungslinien drastisch, um mehrere Zehnerpotenzen, auf Werte von $2 - 5 \cdot 10^5 \text{ cm}^{-2}$ senken läßt. Randbedingungen dafür sind ein moderater Anstieg des Ge-Gehaltes von $\leq 20 \text{ %}/\mu\text{m}$ und eine entsprechende Anpassung der Substrattemperatur während des Wachstums. Die im Zuge dieser Arbeit realisierten SiGe-Heterostrukturen weisen außerdem einen Trend zu ansteigenden Ge-Gehalten mit Höchstwerten von bis zu 50 % auf. Der dadurch stärker ausgeprägte Energiesprung im Leitungsband des Si-Kanals führt zu einer erhöhten Anzahl an besetzbaren Zuständen im Quantentopf und damit zu einer steigenden Ladungsträgerdichte n_s im 2DEG. Durch die Einführung einer doppelseitigen Modulationsdotierung, auf der Vorder- und Rückseite des Kanals, sind Ladungsträgerdichten von bis zu $7 \cdot 10^{12} \text{ cm}^{-2}$ an realen Transistorstrukturen erzielt worden.

Ebenso bedeutsam für die Auslegung der aktiven Schichten ist das Zusammenspiel von Spacerdicke und Dotierstoffkonzentration auf der Vorder- und Rückseite des Kanals. Hier konnte in Experimenten nachgewiesen werden, daß die optimale Spacerdicke zur Realisierung eines möglichst hohen $\mu \cdot n_s$ -Produktes zwischen 3,5 und 4,5 nm liegt. Gleichzeitig sollte das Verhältnis zwischen Vor-

der- und Rückseitendotierung 4/1 bis 5/1 nicht überschreiten. Die gezielte Ausrichtung der Schichtsysteme auf hohe Ladungsträgerdichten liefert für die elektrische Charakteristik des SiGe-HFETs, neben einer Steigerung der Sättigungsströme und der Steilheiten, auch eine Reduzierung der parasitären Widerstände R_s und R_d , die ihrerseits entscheidend das Hochfrequenzverhalten des Transistors beeinflussen. Da das $\mu \cdot n_s$ -Produkt auch für Ge-Gehalte über 50 % noch ansteigt, ist dieser Entwicklungstrend im Rahmen einer Hochfrequenzoptimierung des Schichtsystems auf jeden Fall gerechtfertigt.

Eine zentrale Rolle in dieser Arbeit spielt die Verbesserung des bestehenden Technologieprozesses und die Entwicklung und Implementierung eines neuen selbstjustierenden Integrationskonzeptes. Durch die fortschreitende Miniaturisierung des SiGe-HFETs mit einer Skalierung der Gatelänge auf unter 100 nm bekommen parasitäre Elemente, wie Widerstände (R_s , R_d , R_g) und Kapazitäten (C_{gs} , C_{gd}), zunehmend stärkeren Einfluß auf das elektrische Leistungspotential des Transistors. Das aus der III/V-Technologie abgeleitete konventionelle Integrationskonzept des SiGe-HFETs ist in seinen lateralen Abmessungen durch das vorgegebene Maskenlayout begrenzt, und die Justage der Maskenebenen erfolgt mit optischer Genauigkeit. Dieses Vorgehen resultiert zwangsläufig in einer relaxierten Lateralstruktur des Transistors, mit hohen parasitären Einflüssen und geringer Reproduzierbarkeit im Herstellungsverfahren.

Das im Rahmen dieser Arbeit komplett neu entwickelte selbstjustierende Integrationsverfahren für SiGe-HFETs orientiert sich, mit seiner selbstorganisierten Ausrichtung der Source/Drain-Gebiete zum Gate des Transistors, an den hochoptimierten Konzepten der CMOS-Technologie. Dabei berücksichtigt die Anwendung einer speziellen Replacement-Gate-Struktur das reduzierte Temperaturbudget des SiGe-Herstellungsprozesses. Außerdem umgeht die rein lichtoptische Realisierung des T-Gates, im Vergleich zum konventionellen Integrationsverfahren, den Einsatz der zeit- und kostenintensiven Elektronenstrahlithografie, die für jeden Produktionsablauf ein „Bottle-Neck“ darstellt. Durch die Bereitstellung dreier unterschiedlicher Ausführungsformen des Bauelementes als Schottky-Gate, als MOS- und als pn-Junction-FET beweist das neue Integrationsverfahren eine hohe Flexibilität und eröffnet sich damit einen breiten Anwendungsbereich.

Im Zuge der Prozeßentwicklung wurde das Temperaturbudget der SiGe-Integrationsverfahren analysiert. Um eine permanente und nachhaltige Degrada-

tion des modulationsdotierten Schichtsystems zu vermeiden, sollte die maximale Prozeßtemperatur bei Heterostrukturen mit einem Ge-Gehalt von $\geq 40\%$ $600\text{ }^{\circ}\text{C}$ nicht übersteigen. Unter dieser Prämisse erfolgte anschließend die Optimierung der ohmschen Kontakte. Dazu war es notwendig, die Reichweitenparameter R_p und ΔR_p bei der Ionenimplantation an die speziellen Bedingungen des SiGe-Heteroschichtsystems anzupassen. Durch die gezielte Abstimmung der Implantationsparameter und die Einführung eines selbstjustierenden NiSi-Prozesses ließ sich der Schichtwiderstand der ohmschen Kontakte um 90% auf $10\ \Omega/\square$ senken, gleichzeitig reduzierte sich der spezifische Kontaktwiderstand auf einen Minimalwert von $0,01\ \Omega\text{mm}$. Ein weiteres temperatursensitives Element in der Bauelementstruktur stellt der Pt/Si-Schottky-Übergang zwischen dem Metallgate und der Halbleiteroberfläche der aktiven Schichtstruktur dar. Hier konnte durch die Implementierung einer Schottky-Diode auf PtSi-Basis die Qualität des Gatemo- duls maßgeblich verbessert werden. Durch die gezielte und kontrollierte Silizidierung dünner Platinschichten steigert sich die Temperaturstabilität des Schottky-Gates auf $450\text{ }^{\circ}\text{C}$, gleichzeitig bleibt die hohe Schottky-Barriere von $0,82\text{ eV}$ erhalten.

Die Optimierung der Lateralstruktur des SiGe-HFETs erfordert zwangsläufig die Einführung selbstjustierender Prozeßmodule. Die Umsetzung der T-förmigen Gatestruktur des Transistors mit Hilfe eines Replacement-Gates ermöglicht die selbstjustierende Anordnung der Source/Drain-Gebiete und die Einführung eines Ni-Silizidierungsprozesses unter Einhaltung des limitierten Temperaturbudgets. Des weiteren gewährleistet die Entwicklung eines kontrollierbaren Skalierungsverfahrens die reproduzierbare Realisierung von Gate-längen bis zu 75 nm , ohne den Einsatz der Elektronenstrahlolithografie. Zusätzlich bietet der selbstjustierende Prozeß eine größere Flexibilität bei der Dimensionierung des T-Gates. Aufgrund eines zweiteiligen Herstellungsverfahrens sind die Abmessungen von Gatefuß und Gatekopf unabhängig voneinander einstellbar, was eine systematische Optimierung der Hochfrequenzcharakteristik des Bauelementes erlaubt. Um das Abschnürverhalten des SiGe-HFETs zu verbessern, wurden zusätzlich zwei Prozeßsequenzen für eine selbstjustierende, planare Anordnung des Feldoxides entwickelt. Zusammen mit einem neuen Prozeßmodul zur Planarisierung der Bauelementpassivierung sorgen diese Verfahren dafür, daß die Waferoberfläche nach der Transistorherstellung eine minimale Topographie aufweist und damit die Integrierbarkeit in das mehrlagige Metallisierungssystem einer Schaltung erleichtert wird.

Die umfassende Charakterisierung der DC- und HF-Eigenschaften ermöglicht den detaillierten Vergleich der Integrationsvarianten und belegt das elektrische Leistungspotential des SiGe-HFETs. Dabei zeigt sich, daß durch die Einführung des neuen selbstjustierenden Integrationskonzeptes der Einfluß der parasitären Widerstände und Kapazitäten auf das Bauelementverhalten stark gemindert bzw. besser kontrollierbar wird. Eindeutiger Beweis dafür sind ein Anstieg des Sättigungsstromes um mehr als 40 % auf über 500 mA/mm und eine Erhöhung der Steilheit um 80 % auf Maximalwerte von 685 mS/mm. Zusätzlich führt die selbstjustierende Anordnung des Feldoxides zu einem verbesserten Sättigungsverhalten des Transistors bei gleichzeitig reduzierten Leckströmen. Dieser Trend setzt sich auch im Hochfrequenzbereich fort, hier stellt ein erzieltes f_T von 90 GHz den derzeitigen Höchstwert für SiGe-basierte FETs dar. Darüber hinaus unterstreichen maximale Schwingfrequenzen f_{max} von bis zu 120 GHz und exzellente Rauschkennzahlen mit $F_{min} = 0,3$ dB und $G_{ass} = 19$ dB bei einer Frequenz von 2,5 GHz die Konkurrenzfähigkeit des SiGe-HFETs zu anderen Hochfrequenzbauelementen.

Alle im Rahmen dieser Arbeit durchgeführten Optimierungen und Entwicklungen wurden am Beispiel des SiGe-HFETs mit Schottky-Gate realisiert. Für eine weiterführende Verbesserung des Bauelementverhaltens und eine noch effektivere Ausschöpfung des Leistungspotentials ist die Implementierung eines Gateoxides und die Weiterentwicklung zum SiGe-Hetero-MOSFET unumgänglich. Mit dem hier neu vorgestellten selbstjustierenden Integrationsverfahren wurde gleichzeitig ein Konzept für die Realisierung eines MOS-Transistors eingeführt, dessen technologische Umsetzung Gegenstand zukünftiger Arbeiten sein sollte. Ein weiteres Ziel stellt die Entwicklung einer komplementären Schaltungstechnologie dar. Durch die Kopplung eines p- und eines n-Typ-SiGe-Hetero-MOSFETs läßt sich eine SiGe-CMOS-Technologie realisieren, die analoge, digitale und Hochfrequenzkomponenten auf einem IC monolithisch integriert. Eine technologische Voraussetzung für dieses Konzept ist die Reduzierung der Pufferdicke. Alternative Techniken zur Umsetzung dünnerer Puffer sind im Laufe dieser Arbeit bereits vorgestellt worden, jedoch besitzt keines dieser Verfahren derzeit die Reife und Qualität, den gradierten Puffer komplett zu ersetzen.

Literaturverzeichnis

- [1.1] R. Istvan
„The Future Potential“, GaAs Mantech 1999
- [1.2] Y. Mochida, T. Takano, H. Gambe
„Future directions and technology requirements of wireless communications“, IEDM 2001, Washington, Dec. 2001
- [1.3] J. Knorr
„The Information Society's Demands on Microelectronics: A European Perspective“, interner Vortrag, Siemens AG, Februar 1996
- [1.4] FlashComm Inc.
„SiGe Technology“, Market Overview, FlashComm, <http://www.flashcomm.com/products/market/sld001.html>
- [1.5] K. A. Jackson
„Compound Semiconductor Devices“, Wiley-VCH, pp.48-50, 1998
- [1.6] D. C. Ahlgren
„IBM's SiGe Technology for Telecommunications and Mixed Signal Applications“, IBM Micro News, vol. 3, no. 1, pp. 22-25, 1997
- [1.7] D. Nguyen-Ngoc et al.
„A 200 nm SiGe HBT BiCMOS Technology for Mixed-Signal Applications“, BCTM 1995 Proceedings, p. 89, 1995
- [1.8] T. Mimura, S. Hiyamizu, T. Fujii
„A new Field-Effect Transistor with Selectively Doped GaAs/n-Al_xGa_{1-x}As Heterojunctions“, Japanese Journal of Applied Physics, vol. 19, no. 5, pp. L225-L227, May 1980
- [1.9] H. Dämbkes, H.-J. Herzog, H. Jorke, H. Kibbel, E. Kasper
„The n-Channel SiGe/Si Modulation-Doped Field-Effect Transistor“, IEEE Transactions on Electron Devices, vol. 33, no. 5, pp. 633-638, May 1986
- [1.10] G. Höck, T. Hackbarth, N. Käß, H.-J. Herzog, M. Enciso, F. Aniel, P. Crozat, R. Adde, E. Kohn, U. König
„0.1µm gate length p-type Ge/Si_{0.4}Ge_{0.6} MODFET with 135 GHz f_{max} “, Electronics Letters, vol. 36, no. 16, pp. 1428-1429, Aug. 2000

- [1.11] E. A. Fitzgerald, Y.-H. Xie, M. L. Green, D. Brasen, A. R. Kortab, J. Michel, Y.-J. Mii, B. E. Weir
Appl. Phys. Lett. 59, p. 811, 1991
- [1.12] K. Ismail, S. F. Nelson, J. O. Chu, B. S. Meyerson
„Electron transport properties of Si/SiGe heterostructures: Measurements and device implications“, Appl. Phys. Lett. 63 (5), pp. 660-662, Aug. 1993
- [1.13] M. Glück, T. Hackbarth, M. Birk, A. Haas, E. Kohn, U. König
„Design and fabrication of Si/SiGe n-type MODFETs“, Physica E, E 2, pp. 763-767, 1998
- [2.1] G. E. Moore
„Cramming more components onto integrated circuits“, Electronics, vol. 38, no. 8, April 1965
- [2.2] International Technology Roadmap for Semiconductors, 2001 Edition, Semiconductor Industry Association, 2001
- [2.3] J. Knorr
„The Information Society's Demands on Microelectronics: A European Perspective“, interner Vortrag, Siemens AG, Februar 1996
- [2.4] „DRAMs are drivers below half-micron“, World News aus Solid State Technology, p. 54, Oct. 1997
- [2.5] „Product and technology distribution, 2000“, World News aus Solid State Technology, p. 20, June 2000
- [2.6] Dr. Thoreg
Konzernplanung und Controlling: Halbleitermarkt (interner Vortrag), Daimler Benz AG 1997
- [2.7] Stuart McIntosh
„Integration for Multimedia“, Vortrag, Philips Semiconductors, 1996
- [2.8] Gartner Dataquest
„Worldwide Semiconductor Revenues and Forecast, 1999-2005“, Semiconductor Magazine, vol. 3, no. 1, Jan. 2002
- [2.9] A. Fischer, H. Richter
„Recherche IHP-CMOS I“, Institut für Halbleiterphysik Frankfurt (Oder), S. 1-25, 1996
- [2.10] D. Behammer
„Produktionstechniken der Halbleitertechnologie“, Vorlesungsskript, Universität Ulm, Institut für elektronische Bauelemente und Schaltungen, WS 00/01

- [2.11] A. Brand, A. Haranahalli, N. Hsieh, et al.
„Intel's 0.25 Micron, 2.0 Volts Logic Process Technology“, Intel Technology Journal 3rd quarter '98,
<http://developer.intel.com/technology/itj/q31998.htm>
- [2.12] R. Paul
„MOS-Feldeffekttransistoren“, Reihe Halbleiter-Elektronik, Band 21,
Springer Verlag, S. 332-340, 1994
- [2.13] S. Thompson, P. Pachkan, M. Bohr
„MOS Scaling: Transistor Challenges for the 21st Century“, Intel Technology Journal 3rd quarter '98,
<http://developer.intel.com/technology/itj/q31998.htm>
- [2.14] G. G. Shahidi, J. Warnock, S. Fischer et al.
„High- Performance Devices for a 0,15- μm CMOS Technology“, IEEE Electron Device Letters, vol. 14, no. 10 Oct. 1993
- [2.15] P. H. Woerlee, C. A. H. Juffermans, H. Lifka et al.
„A Low Power 0.25 μm CMOS Technology“, IEEE IEDM April 1992
- [2.16] W.-H. Chang, B. Davari, M. R. Wordemann et al.
„A High-Performance 0.25- μm CMOS Technology: I-Design and Characterization“, IEEE Transactions on Electron Devices, vol. 39, no. 4, April 1992
- [2.17] D. J. Frank
„Design Considerations for CMOS Near the Limits of Scaling“, ULIS 2000 Workshop, Grenoble, Frankreich, 20./21. Januar 2000
- [2.18] T. Sugii
„Continued Growth in CMOS beyond 0.1 μm “, ULIS 2000 Workshop, Grenoble, Frankreich, 20./21. Januar 2000
- [2.19] G. G. Shahidi, J. D. Warnock, J. Comfort et al.
„CMOS scaling in the 0.1- μm , 1.X-volt regime for high-performance applications“, IBM Journal Research & Development, vol. 39, no. 1/2, January/March 1995
- [2.20] Y. Taur, Y.-J. Mii, D. J. Frank et al.
„CMOS scaling into the 21st century: 0.1 μm and beyond“, IBM Journal Research & Development, vol. 39, no. 1/2, January/March 1995
- [2.21] „1999 Roadmap: Solutions and caveats“, Executive millennium report, Solid State Technology, pp. 76-114, May 2000

- [2.22] P. Burggraaf
„Optical lithography to 2000 and beyond“, Millennium Series: Lithography, Solid State Technology, February 1999
- [2.23] J. E. Bjorkholm
„EUV Lithography – The Successor to Optical Lithography?“, Intel Technology Journal 3rd quarter '98,
http://developer.intel.com/technology/itj/q31998/articles/art_4a.htm
- [2.24] S. Deleonibus
„CMOS: How far can it go?“, Solid State Technology, pp. S14-S21, March 2000
- [2.25] T. Skotnicki
„Heading for decananometer CMOS – Is navigation among icebergs still a viable strategy?“, ESSDERC 2000, Cork Ireland, Sep. 2000
- [2.26] Y. Taur, H. Wann, D. J. Frank
„25 nm CMOS Design Considerations“, IEEE IEDM 1998, San Francisco, Dec. 1998
- [2.27] B. Metzger
„CMOS devices steal show“, Compound Semiconductor, pp. 31-33, Jan./Feb. 2002
- [2.28] IBM's Technology Showcase
„IBM's SOI approach“, IBM's Technology Showcase,
<http://www.chips.ibm.com/bluelogic/showcase/soi/appr1.html>
- [2.29] E. Korczynski
„Low-*k* dielectric integration cost modelling“, Solid State Technology, pp. 123-128, Oct. 1997
- [2.30] D. Seligson
„Planning for the 300 mm Transition“, Intel Technology Journal 4th quarter '98, <http://developer.intel.com/technology/itj/q41998.htm>
- [2.31] S. Strähle
„Entwurf und Technologie von Heterostrukturfeldeffekttransistoren auf InP-Basis mit hohem Verstärkungsbandbreiteprodukt bei hoher Leistungsverstärkung“, Dissertation, Fortschrittberichte des VDI, Reihe 9, Nr. 234, S. 1-4, 1996
- [2.32] K. Shinohara, Y. Yamashita, A. Endoh et al.
„InP-Based HEMTs with a cutoff frequency higher than 450 GHz“, Device Research Conference 2002, conference digest, pp. 163-166, Santa Barbara, 2002

- [2.33] K. A. Jackson
„Compound Semiconductor Devices“, Wiley-VCH, pp.48-50, 1998
- [2.34] R. Dingle, H. L. Störmer, A. C. Gossard, W. Wiegmann
„Electron mobilities in modulation-doped semiconductor heterojunction superlattices“, Applied Physics Letters, vol. 33, no. 7, pp.665-667, 1978
- [2.35] T. Mimura, S. Hiyamizu, T. Fujii
„A new Field-Effect Transistor with Selectively Doped GaAs/n-Al_xGa_{1-x}As Heterojunctions“, Japanese Journal of Applied Physics, vol. 19, no. 5, pp. L225-L227, May 1980
- [2.36] D. Delagebeaudeuf, P. Delescluse et al.
„Two-Dimensional Electron Gas M.E.S.F.E.T. Structure“, Electronics Letters, vol. 16, no. 17, pp. 667-668, August 1980
- [2.37] F. Ali, A. Gupta
„HEMTs and HBTs: Devices, Fabrication, and Circuits“, Artech House, Chapter 3, pp. 81-85, 1991
- [2.38] K. K. Ng
„Complete Guide to Semiconductor Devices“, McGraw-Hill, Chapter 22, pp. 197-202, 1995
- [2.39] K. A. Jackson
„Compound Semiconductor Devices“, Wiley-VCH, pp. 54-56, 1998
- [2.40] F. Ali, A. Gupta
„HEMTs and HBTs: Devices, Fabrication, and Circuits“, Artech House, Chapter 3, pp. 91-100, 1991
- [2.41] J. J. Rosenberg, M. Benlarmi, P. D. Kirchner et al.
„An In_{0.15}Ga_{0.85}As/GaAs Pseudomorphic Single Quantum Well HEMT“, IEEE Electron Device Letters, vol. 6, no. 10, pp. 491-493, Oct. 1985
- [2.42] K. Heime
„III-V-Halbleiter“, Vorlesungsskript RWTH Aachen, Institut für Halbleitertechnik, S. 253-257, 1992
- [2.43] L. F. Lester, P. M. Smith, P. Ho et al.
„0.15 μm Gate-Length Double Recess Pseudomorphic HEMT with f_{\max} of 350 GHz“, IEEE IEDM 1988, Washington, p. 172, 1998
- [2.44] K. Heime
„III-V-Halbleiter“, Vorlesungsskript RWTH Aachen, Institut für Halbleitertechnik, S. 257-265, 1992

- [2.45] P. M. Smith, S.-M. J. Liu, M. Y. Kao et al.
„W-Band High Efficiency InP-Based Power HEMT with 600 GHz f_{\max} “, IEEE Microwave and Guided Wave Letters, vol. 5, no. 7, pp. 230-232, 1995
- [2.46] U. K. Mishra, A. S. Brown, S. E. Rosenbaum, et al.
IEEE Electron Device Letters, vol. 9, no. 12, pp. 647-649, 1998
- [2.47] R. Szweda
„Markets and technologies for advanced semiconductors“, Compound Semiconductor Industry Directory, Elsevier Science Ltd., pp. 14-30, 1997
- [2.48] D. C. Dumka, W. E. Hoke, P. J. Lemonias et al.
„Metamorphic $\text{In}_{0.52}\text{Al}_{0.48}\text{As}/\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ HEMT on GaAs Substrate with f_T over 200 GHz“, IEEE IEDM 1999
- [2.49] K. A. Jackson
„Compound Semiconductor Devices“, Wiley-VCH, pp. 84-94, 1998
- [2.50] A. N. LePore, H. M. Levy, R. C. Tiberio, et al.
„0.1 μm gate length MODFETs with unity current gain cutoff frequency above 110 GHz“, Electron. Lett., vol. 24, p. 364, 1988
- [2.51] P. M. Smith, P. C. Chao, K. H. G. Duh, L. F. Lester, J. M. Ballingall
„Advances in HEMT technology and applications“, IEEE MTT-S Digest, Las Vegas, p. 749, 1987
- [2.52] P. C. Chao, M. S. Shur, R. C. Tiberio, et al.
„DC and microwave characteristics of sub-0.1 μm gate-length planar-doped pseudomorphic HEMT's“, IEEE Trans. Electron Devices, vol. 36, p. 461, 1989
- [2.53] L. D. Nguyen, P. J. Tasker, D. C. Radulescu, L. F. Eastman
„Characterization of ultra-high-speed pseudomorphic AlGaAs/InGaAs (on GaAs) MODFETs“, IEEE Trans. Electron Devices, vol. 36, p. 2243, 1989
- [2.54] F. Ali, A. Gupta
„HEMTs and HBTs: Devices, Fabrication, and Circuits“, Artech House, Chapter 3, pp. 103-121, 1991
- [2.55] C. Gässler, V. Ziegler, C. Wölk, R. Deufel, F. J. Berlec
„Metamorphic HFETs on GaAs with InP-subchannels for device performance improvements“, IEEE IEDM 2000, San Francisco, Dec. 2000

- [2.56] D. Xu, T. Suemitsu, J. Osaka
„Ultrahigh-Speed InP-Based D- and E-Mode MODFETs with Ultra-Short Electrochemically-Recessed Gate Contacts“, IEEE DRC 1999, Santa Barbara, Conference Digest, pp. 150-151, 1999
- [2.57] F. Ali, A. Gupta
„HEMTs and HBTs: Devices, Fabrication, and Circuits“, Artech House, Chapter 3, pp. 123-145, 1991
- [2.58] L. D. Nguyen, A. S. Brown, M. A. Thompson, L. M. Jelloian
„50nm InP High Electron Mobility Transistor“, Microwave Journal, Technical Note, pp. 96-101, June 1993
- [2.59] M. Y. Kao, S. M. J. Liu, K. H. G. Duh, P. C. Chao, P. Ho, A. W. Swanson
„A 0.1 μm T-shaped gate InP-based HEMT for low noise application“, 4th International Conference on InP and Related Materials, paper TUA3, April 1992
- [2.60] P. Ho, P. M. Smith, K. C. Hwang, S. C. Wang, M. Y. Kao, P. C. Chao, S. M. J. Liu
„60 GHz power performance of 0.1 μm gate-length InAlAs/InGaAs HEMTs“, 6th International Conference on InP and Related Materials, pp. 18-20, 1994
- [2.61] S. S. Iyer, G. L. Patton, S. S. Delage, S. Tiwari, J. M. C. Stork
„Silicon-Germanium Heterojunction Bipolar Transistor by Molecular Beam Epitaxy“, IEDM 1987, pp. 874-876, 1987
- [2.62] D. C. Ahlgren
„IBM's SiGe Technology for Telecommunications and Mixed Signal Applications“, IBM Micro News, vol. 3, no. 1, pp. 22-25, 1997
- [2.63] B. S. Meyerson
„Silicon:germanium-based mixed-signal technology for optimization of wired and wireless telecommunications“, IBM Journal Research & Development, vol. 44, no. 3, May 2000
- [2.64] M. Carroll, T. Ivanov, S. Kuehne, et al.
„COM2 SiGe Modular BiCMOS Technology for Digital, Mixed-Signal, and RF Applications“, IEDM 2000, San Francisco, Dec. 2000
- [2.65] R. A. Metzger
„Is Silicon Germanium the New "Material of the Future"?“, Compound Semiconductor, pp.21-28, Nov/Dec 1995

- [2.66] A. Gruhle
„Prospects for 200 GHz on Silicon with SiGe Heterojunction Bipolar Transistors“, Proceedings of the IEEE 2001 BCTM, Minneapolis, Sept. 2001
- [2.67] U. König
„SiGe RF-Electronic: Devices, Circuits, Competitors, Markets“, Materials Research Symposium Proceedings, vol. 533, pp. 1-17, 1998
- [2.68] FlashComm Inc.
„SiGe Technology“, Market Overview, FlashComm, <http://www.flashcomm.com/products/market/sld001.html>
- [2.69] D. C. Ahlgren, J. Dunn
„SiGe comes of age“, IBM Micro News, vol.6, no. 1, 2000
- [2.70] D. Nguyen-Ngoc et al.
„A 200 mm SiGe HBT BiCMOS Technology for Mixed-Signal Applications“, BCTM 1995 Proceedings, p. 89, 1995
- [2.71] D. Harame
„Status and Trends of SiGe BiCMOS Technology“, US-Asia Technology Management Center Stanford University, Spring 2001 Seminar, <http://asia.stanford.edu/events/spring01/slides/HarameSlides.pdf>
- [2.72] S. A. St. Onge, D. Harame, J. S. Dunn et al.
„A 0.24 μm SiGe BiCMOS Mixed-Signal Production Technology Featuring a 47 GHz f_T HBT and 0,18 μm L_{eff} CMOS“, Proceedings of the 1999 Bipolar Circuits and Technology Meeting, pp. 117-120, Oct. 1999
- [2.73] T. Hackbarth, G. Höck, H.-J. Herzog, M. Zeuner
„Strain relieved SiGe buffers for Si-based heterostructure field-effect transistors“, Journal of Crystal Growth, vol. 201/202, pp. 734-738, 1999
- [2.74] K. Ismail, S. F. Nelson, J. O. Chu, B. S. Meyerson
„Electron transport properties of Si/SiGe heterostructures: Measurements and device implications“, Appl. Phys. Letters, vol. 63, no. 5, pp. 660-662, Aug. 1993
- [2.75] S. Takagi, A. Toriumi, M. Iwase, H. Tango
„On the universality of inversion layer mobility in Si MOSFETs: Part I effects of substrate impurity concentration“, IEEE Transactions on Electron Devices, 41 (12), pp. 2357-2369, 1994

- [2.76] B. Fischer, K. R. Hofmann
„Full-band Monte Carlo model of electron and hole transport in strained Si including inelastic acoustic phonon scattering“, Appl. Phys. Letters, vol. 74, no. 15, 1999
- [2.77] H. Miyata, T. Yamada, D. K. Ferry
„Electron transport properties of a strained Si layer on a relaxed Si_{1-x}Ge_x substrate by Monte Carlo simulation“, Appl. Phys. Lett., vol. 62, no. 21, pp. 2661-2663, May 1993
- [2.78] A. Cappy, B. Carnez, R. Fauquembergues, G. Salmer, E. Constant
„Comparative Potential Performance of Si, GaAs, GaInAs, InAs Submicrometer-Gate FETs“, IEEE Trans. on Electron Dev., vol. 27, no. 11, pp. 2158-2160, Nov. 1980
- [2.79] D. J. Paul
„Silicon Germanium Heterostructures in Electronics: The Present and the Future“, Thin Solid Films, 321, pp. 172, 1998
- [2.80] K. Ismail
„Si/SiGe High-Speed Field-Effect Transistors“, IEEE Tech. Digest IEDM 1995, pp. 509-512, 1995
- [2.81] M. Enciso, F. Aniel, P. Crozat, L. Giguere, R. Adde, M. Zeuner, G. Höck, A. Fox
„Ultra-low noise strained Si/SiGe n- and Ge/SiGe p-MODFETs“, ESSDERC 2001, Nürnberg, September 2001
- [2.82] U. König
„N- and P- Type Si/SiGe Hetero FETs“, Proceedings of 8th IEEE Int. Symposium EDMO, pp. 1, 2000
- [2.83] E. H. C. Parker, T. E. Whall
persönliche Mitteilung
- [2.84] M. Arafa, K. Ismail, J. O. Chu, B. S. Meyerson, I. Adesia
„A 70 GHz f_T low operating bias self aligned p-type SiGe MODFET“, IEEE Electron Device Lett., vol. 17, no. 12, pp. 586-588, 1996
- [2.85] G. Höck, T. Hackbarth, N. Käß, H.-J. Herzog, M. Enciso, F. Aniel, P. Crozat, R. Adde, E. Kohn, U. König
„0,1 μm gate length p-type Ge/Si_{0.4}Ge_{0.6} MODFET with 135 GHz f_{max} “, Electronics Letters, vol. 36, no. 16, pp. 1428-1429, 2000
- [2.86] U. König, H. Dämbkes
„SiGe HBTs and HFETs“, Solid-State Electronics, vol. 38, no. 9, pp. 1595-1602, 1995

- [2.87] A. Sadek, K. Ismail, M. A. Armstrong, D. A. Antoniadis, F. Stern
„Design of Si/SiGe Heterojunction Complementary Metal-Oxide-Semiconductor Transistors“, IEEE Transactions on Electron Devices, vol. 43, no. 8, pp. 1224-1232, 1996
- [2.88] T. Hackbarth
persönliche Mitteilung
- [3.1] S. M. Sze
„Physics of Semiconductor Devices“, John Wiley & Sons, New York, 2.Auflage, pp. 7-16, 1981
- [3.2] S. C. Jain, W. Hayes
„Structure properties and applications of $\text{Ge}_x\text{Si}_{1-x}$ strained layers and superlattices“, Semiconductor Sci. Technology 6, pp. 547-576, 1991
- [3.3] J. Weber, M. I. Alonso
„Near-band-gap photoluminescence of Si-Ge alloys“, Phys. Rev. B 40 (8), pp. 5683-5693, 1989
- [3.4] M. Wachter
„Optische Untersuchungen an Potentialtopfstrukturen aus Si/SiGe“, Diplomarbeit in der Abteilung Halbleiterphysik der Universität Ulm, S. 3-22, 1992
- [3.5] S. Lüthin
„Simulation von Si/SiGe Hetero-Feldeffekttransistoren“, Diplomarbeit im Studiengang Elektrotechnik an der Universität Ulm, S. 19-31, 1995
- [3.6] E. Kasper, H.-J. Herzog
„Elastic strain and misfit dislocation density in $\text{Ge}_{0.92}\text{Si}_{0.08}$ films on silicon substrates“, Thin Solid Films 44, pp. 357-370, 1977
- [3.7] C. G. van de Walle
„Band lineups and deformation potentials in the model solid theory“, Phys. Rev. B 39 (3), pp. 1871-1883, 1989
- [3.8] C. Herring, E. Vogt
„Transport and deformation-potential theory for many-valley semiconductors with anisotropic scattering“, Phys. Rev. 101 (3), p. 944, 1955
- [3.9] B. Steck
„Photolumineszenzuntersuchungen an pseudomorphen Si/SiGe-Heterostrukturen unter externem uniaxialen Druck“ Diplomarbeit in der Abteilung Halbleiterphysik der Universität Ulm, S. 4-22, 1995

- [3.10] M. M. Rieger, P. Vogl
„Electronic-band parameters in strained $\text{Si}_{1-x}\text{Ge}_x$ alloys on $\text{Si}_{1-y}\text{Ge}_y$ substrates“, Phys. Rev. B 48 (19), pp. 14276-14287, 1993
- [3.11] C. G van der Walle, R. M. Martin
„Theoretical study of Si/Ge interfaces“, Phys. Rev. B 34 (3), pp. 5621-5629, 1986
- [3.12] C. G. van der Walle, R. M. Martin
„Theoretical calculations of heterojunction discontinuities in Si/Ge systems“, Phys. Rev. B 34 (3), pp. 5621-5629, 1986
- [3.13] L. Colombo, R. Resta, S. Baroni
„Valance band offsets at strained Si/Ge interfaces“, Phys. Rev. B 44, p. 5572, 1991
- [3.14] F. Schäffler
„High-mobility Si and Ge structures“, Semicond. Sci. Technology, no. 12, pp. 1515-1549, 1997
- [3.15] G. Höck, T. Hackbarth, N. Käß, H.-J. Herzog, M. Enciso, F. Aniel, P. Crozat, R. Adde, E. Kohn, U. König
„0.1 μm gate length p-type Ge/ $\text{Si}_{0.4}\text{Ge}_{0.6}$ MODFET with 135 GHz f_{max} “, Electronics Letters, vol. 36, no. 16, pp. 1428-1429, Aug. 2000
- [3.16] M. Arafa, P. Fay, K. Ismail, J. O. Chu, B. S. Meyerson, I. Adesida
„DC and RF Performance of 0.25 μm p-Type SiGe MODFET“, IEEE Electron Device Letters, vol. 17, no. 9, September 1996
- [3.17] R. People, J. C. Bean, D. V. Lang, A. M. Segent, H. L. Störmer, K. W. Wecht, R. T. Lynch, K. Baldwin
„Modulation doping in $\text{Ge}_x\text{Si}_{1-x}/\text{Si}$ strained layer heterostructures“, Appl. Phys. Lett. 45 (11), pp. 1231-1233, Dec. 1984
- [3.18] R. People, J. C. Bean, D. V. Lang
„Modulation doping in $\text{Ge}_x\text{Si}_{1-x}/\text{Si}$ strained layer heterostructures: Effects of alloy layer thickness, doping setback, and cladding layer dopant concentration“, Journal. Vac. Sci. Technology A 3 (3), pp. 846-850, May/June 1985
- [3.19] H. Jorke, H.-J. Herzog
„Mobility Enhancement in Modulation-Doped Si- $\text{Si}_{1-x}\text{Ge}_x$ Superlattice Grown by Molecular Beam Epitaxy“, Journal of the Electrochemical Society, vol. 133, no. 5, pp. 998-1001, May 1986

- [3.20] H. Dämbkes, H.-J. Herzog, H. Jorke, H. Kibbel, E. Kasper
„The n-Channel SiGe/Si Modulation-Doped Field-Effect Transistor“,
IEEE Transactions on Electron Devices, vol. 33, no. 5, pp. 633-638,
May 1986
- [3.21] K. Ismail, M. Arafa, K. L. Saenger, J. O. Chu, B. S. Meyerson
„Extremely high electron mobility in Si/SiGe modulation-doped hete-
rostructures“, Appl. Phys. Lett. 66 (9), pp. 1077-1079, Feb. 1995
- [3.22] G. S. Boebinger, A. M. Chang, H. L. Störmer, D. C. Tsui
Phys. Rev. Lett. 55, p. 1606, 1985
- [3.23] K. Ismail, S. F. Nelson, J. O. Chu, B. S. Meyerson
„Electron transport properties of Si/SiGe heterostructures: Measure-
ments and device implications“, Appl. Phys. Lett. 63 (5), pp. 660-662,
Aug. 1993
- [3.24] F. Schäffler
„High-mobility Si and Ge structures“, Semicond. Sci. Technol. 12,
pp. 1515-1549, 1977
- [3.25] P. Dollfus
„Si/Si_{1-x}Ge_x heterostructures: Electron transport and field-effect transi-
stor operation using Monte Carlo simulation“, Journal Appl. Phys. 82
(8), pp. 3911-3916, 1997
- [3.26] D. W. Winston
SimWindows Version 1.5.0, University of Colorado, Boulder, 1999
- [3.27] E. B. Johnson, A. Rose
„Simple general analysis of amplifier devices with emitter, control and
collector function“, Proc. Inst. Radio Engrs., vol. 47, pp. 407-418, 1959
- [3.28] D. Delagebeaudeuf, N. T. Linh
„Metal-(n) AlGaAs-GaAs Two-Dimensional Electron Gas FET“, IEEE
Transactions on Electron Devices, vol. 29, no. 6, pp. 955-960, 1982
- [3.29] T. J. Drummond, H. Morkoc, K. Lee, M. Shur
„Model for Modulation Doped Field Effect Transistors“, IEEE Electron
Device Letters, vol. 3, no. 11, pp. 338-341, 1982
- [3.30] M. B. Das, M. L. Roszak
„Design calculations for submicron gate-length AlGaAs/GaAs modula-
tion-doped FET structures using carrier saturation velocity/charge con-
trol model“, Solid-State Electronics, vol. 28, no. 10, pp.997-1005, 1985

- [3.31] D. M. Caughey, R. E. Thomas
„Carrier Mobilities in Silicon Empirically Related to Doping and Field“, Proceedings of the IEEE, vol. 55, pp. 2192-2193, 1967
- [3.32] M. B. Das, W. Kopp, H. Morkoc
„Determination of Carrier Saturation Velocity in Short-Gate-Length Modulation-Doped FETs“, IEEE Electron Device Letters, vol. 5, no. 11, pp. 446-449, 1984
- [3.33] M. Glück
„Herstellung von sub- μm Si/SiGe Heterostruktur Feldeffekttransistoren (HFETs) und ersten Testschaltungen für Hochfrequenzanwendungen“, Dissertation an der Universität Ulm, Fakultät für Ingenieurwissenschaften, 1999
- [3.34] B. Fischer, K. R. Hofmann
„Full-band Monte Carlo model of electron and hole transport in strained Si including inelastic acoustic phonon scattering“, Applied Physics Letters, vol. 74, no. 15, 1999
- [3.35] H. Miyata, T. Yamada, D. K. Ferry
„Electron transport properties of a strained Si layer on a relaxed $\text{Si}_{1-x}\text{Ge}_x$ substrate by Monte Carlo simulation“, Appl. Phys. Lett. 62 (21), pp. 2661-2663, 1993
- [3.36] T. Ostermann
„Untersuchungen des Si/Si $_{1-x}$ Ge $_x$ -MODFET für den Einsatz in integrierten Schaltungen“, Dissertation an der Universität Linz, Abteilung für Mikroelektronische Systeme, 1999
- [3.37] B. Hughes, P. J. Tasker
„Bias Dependence of the MODFET Intrinsic Model Elements Values at Microwave Frequencies“, IEEE Transactions on Electron Devices, vol. 36, no. 10, pp. 2267-2273, 1989
- [3.38] P. J. Tasker, B. Hughes
„Importance of Source and Drain Resistance to the Maximum f_T of Millimeter-Wave MODFET's“, IEEE Electron Device Letters, vol. 10, no. 7, pp. 291-293, 1989
- [4.1] T. Hackbarth, H.-J. Herzog, M. Zeuner, G. Höck, E. A. Fitzgerald, M. Bulsara, C. Rosenblad, H. von Känel
„Alternatives to thick MBE-grown relaxed SiGe buffers“, Thin Solid Films, vol. 369, pp. 148-151, 2000

- [4.2] F. Schäffler
„High-mobility Si and Ge structures“, *Semicond. Sci. Technology*,
vol. 12, pp. 1515-1549, 1997
- [4.3] E. A. Fitzgerald, Y.-H. Xie, M. L. Green, D. Brasen, A. R. Kortab, J.
Michel, Y.-J. Mii, B. E. Weir
Appl. Phys. Lett. 59, p. 811, 1991
- [4.4] T. Hackbarth, H. Kibbel, M. Glück, G. Höck, H.-J. Herzog
„Artificial substrates for n- and p-type SiGe heterostructure field-effect
transistors“, *Thin Solid Films*, no. 321, pp. 136-140, 1998
- [4.5] M. Hohnisch
„Strukturuntersuchungen an relaxierten SiGe-Schichten“, Diplomarbeit
an der Technischen Universität München, 1995
- [4.6] C. S. Peng, H. Chen, Z. Y. Zhao, J. H. Li, D. Y. Dai, Q. Huang, J. M.
Zhou, Y. H. Zhang, C. H. Tung, T. T. Sheng, J. Wang
Journal of Crystal Growth 201/202, p. 530, 1999
- [4.7] S. Mantl, B. Holländer, R. Liedtke, S. Mesters, H.-J. Herzog, H. Kibbel,
T. Hackbarth
Nucl. Instr. and Meth. in Phys. Res. B 148, p. 200, 1999
- [4.8] H. Trinkaus, B. Holländer, S. Rongen, S. Mantl, H.-J. Herzog, J. Ku-
chenbecker, T. Hackbarth
Zur Veröffentlichung angenommen bei *Appl. Phys. Lett.*
- [4.9] Z. Yang, J. Alperin, W. I. Wang, S. S. Iyer, T. S. Kuang, F. Semendy
Journal of Vac. Sci. Technology B 16, p.1489, 1998
- [4.10] J. L. Liu, C. D. Moore, G. D. U'Ren, Y. H. Luo, Y. Lu, G. Jin, S. G.
Thomas, M. S. Goorsky, K. L. Wang
Appl. Phys. Lett. 75, p. 1586, 1999
- [4.11] D. Reinking, M. Kammler, M. Horn van Högen, K. Hofmann
Appl. Phys. Lett. 71, p. 2545, 1997
- [4.12] H. Dämbkes, H.-J. Herzog, H. Jorke, H. Kibbel, E. Kasper
„The n-Channel SiGe/Si Modulation-Doped Field-Effect Transistor“,
IEEE Transactions on Electron Devices, vol. 33, no. 5, pp. 633-638,
May 1986
- [4.13] K. Ismail, B. S. Meyerson, S. Rishton, J. Chu, S. Nelson, J. Nocera
„High-Transconductance n-Type Si/SiGe Modulation-Doped Field-
Effect Transistor“, *IEEE Electron Device Letters*, vol. 12, no. 5,
pp. 229-231, 1992

- [4.14] U. König, F. Schäffler
„Modulation-Doped High-Mobility Si/SiGe Heterostructures for Device Application“, International Conference on Solid State Devices and Materials (SSDM), Japan, 1993
- [4.15] J. W. Matthews, A. E. Blakeslee
„Defects in epitaxial multilayers I: Misfit dislocations“, J. Cryst. Growth, vol. 27, p. 118, 1974
- [4.16] J. C. Bean, L. C. Feldman, A. T. Fiory, S. Nakahara, J. K. Robinson
„Ge_xSi_{1-x}/Si strained layer superlattice grown by molecular beam epitaxy“, Journ. of Vac. Science and Technology, A2 (2), pp. 436-440, 1984
- [4.17] M. Glück
„Herstellung von sub- μm Si/SiGe Heterostruktur-Feldeffekttransistoren (HFETs) und ersten Testschaltungen für Hochfrequenzanwendungen“, Dissertation der Fakultät für Ingenieurwissenschaften an der Universität Ulm, S. 43 – 46, 1998
- [4.18] S. C. Jain et al.
Solid-State Electronics, vol. 35, p. 1073, 1992
- [4.19] T. Hackbarth
„SiGe MODFETs grown by MBE on CVD strain relieved buffers“, Vortrag auf dem 11th European Workshop on MBE, 6. Feb. 2001
- [5.1] M. Glück
„Herstellung von sub- μm Si/SiGe Heterostruktur-Feldeffekttransistoren (HFETs) und ersten Testschaltungen für Hochfrequenzanwendungen“, Dissertation der Fakultät für Ingenieurwissenschaften an der Universität Ulm, S. 130 – 149, 1998
- [5.2] M. Zeuner, T. Hackbarth, U. König, A. Gruhle, F. Aniel
„High Performance 0.25 μm T-Gate SiGe-n-MODFET“, Proceedings of the DRC, pp. 177-178, 1999
- [5.3] M. Zeuner, A. Fox, T. Hackbarth, D. Behammer, U. König
„90 GHz f_T HFET with fully optical self-aligned sub 100 nm gate“, Proceedings of the DRC 2002, pp. 53-54, 2002
- [5.4] M. Zeuner, U. König
„SiGe HFET Technology“, Invited Paper, Proc. of the MWW, 2001
- [6.1] F. Schäffler, H.-J. Herzog, H. Jorke, E. Kasper
„Influence of thermal annealing on the electron mobility in modulation doped Si/SiGe heterostructures“, J. Vac. Sci. Technol. B 9 (4), pp. 2039-2044, July/Aug. 1991

- [6.2] H. Klauk, T. N. Jackson, S. F. Nelson, J. O. Chu
„Thermal stability of undoped strained Si Channel SiGe heterostructures“, Appl. Phys. Lett. 68 (14), pp. 1975-1977, April 1996
- [6.3] M. Griglione, T. J. Anderson, Y. M. Haddara, M. E. Law, K. S. Jones, A. van den Bogaard
„Diffusion of Ge in Si_{1-x}Ge_x/Si single quantum wells in inert and oxidizing ambients“, J. Appl. Phys. 88 (3), pp. 1366-1372, Aug. 2000
- [6.4] B. Holländer, S. Mantl, B. Stritzker, R. Butz
„Ion scattering studies of diffusion and strain relaxation in Si/Si_{1-x}Ge_x superlattices“, IBMM 1990, Knoxville, 1990
- [6.5] P. Zaumseil, G. G. Fischer, K. Brunner, K. Eberl
„Comparison of the thermal stability of Si_{0.603}Ge_{0.397}/Si and Si_{0.597}Ge_{0.391}C_{0.012}/Si superlattice structures“, J. Appl. Phys. 81 (9), pp. 6134-6140, May 1997
- [6.6] U. König, A. J. Boers, F. Schäffler
„N-Channel Si/SiGe MODFET's: Effects of Rapid Thermal Activation on the DC Performance“, IEEE Electron Device Letters, vol. 14, no. 3, pp. 97-99, March 1993
- [6.7] A. N. Larsen, P. Kringhoj
„Diffusion of Sb in relaxed Si_{1-x}Ge_x“, Appl. Phys. Lett. 68 (19), pp. 2684-2686, May 1996
- [6.8] J. Lindhard, M. Scharff, H. E. Schiott
Phys. Rev. 124, p. 128, 1961
- [6.9] H. Ryssel, I. Ruge
„Ionenimplantation“, Teubner Verlag Stuttgart, 1. Auflage, pp. 15-39, 1978
- [6.10] J. F. Gibbons, W. S. Johnson, S. W. Mylroie
„Projected Range Statistics“, Halsted Press Stroudsburg, 2nd Edition, p. 29, 1975
- [6.11] H. Ryssel, I. Ruge
„Ionenimplantation“, Teubner Verlag Stuttgart, 1. Auflage, pp.313-343, 1978
- [6.12] U. König
persönliche Mitteilung
- [6.13] International Technology Roadmap for Semiconductors, Update 2000 der Roadmap von 1999, Semiconductor Industry Association, 2000

- [6.14] L. J. Chen, S. L. Cheng, S. M. Chang, Y. C. Peng, H. Y. Huang, L. W. Cheng
„Silicide contacts for sub-0.25 μm devices“, Mat. Res. Symp. Proc.,
vol. 564, pp. 123-134, 1999
- [6.15] A. Fischer, H. Richter
„Recherche IHP-CMOS I“, Institut für Halbleiterphysik Frankfurt
(Oder), S. 1-25, 1996
- [6.16] G. Schumicki, P. Seegebrecht
„Prozesttechnologie“, Springer-Verlag, Berlin, pp. 241-260, 1991
- [6.17] J. Winnerl
„Silicides for High Density Memory and Logic Circuits“, Semicon-
ductor International, vol. 8, pp. 81-86, August 1994
- [6.18] K. Maex
„CoSi₂: An Attractive Alternative to TiSi₂“, Semiconductor Internatio-
nal, vol. 3, pp. 75-80, March 1995
- [6.19] J. Chen, J.-P. Colinge, D. Flandre, R. Gillon, J. P. Raskin, D. Vanhoen-
nacker
„Comparison of TiSi₂, CoSi₂ and NiSi for Thin-Film Silicon-on-
Insulator Applications“, J. Electrochem. Soc. ,vol. 144, no. 7, pp. 2437-
2441, July 1997
- [6.20] G. R. Hilbrandie, S. J. M. Bakker, E. van der Drift, B. A. C. Rous-
seeuw, T. M. Klapwijk, S. Radelaar
„Mesoscopic silicon coupled superconducting junctions of CoSi₂ for-
med in a self-aligned process“, Microelectronic Engineering 23,
pp. 445-448, 1994
- [6.21] M. C. Poon, F. Deng, M. Chan, W. Y. Chan, S. S. Lau
„Resistivity and thermal stability of nickel mono-silicide“, Applied Sur-
face Science, vol. 157, pp. 29-34, 2000
- [6.22] B. Meyer, U. Gottlieb, O. Laborde, H. Yang, J. C. Lasjaunias, A. Sul-
pice, R. Madar
„Intrinsic properties of NiSi“, Journal of Alloys and Compounds, 262-
263, pp. 235-237, 1997
- [6.23] A. Lauwers, P. Besser, T. Gutt, A. Satta, M. de Potter et al.
„Comparative study of Ni-silicide and Co-silicide for sub 0.25- μm
technologies“, Microelectronic Engineering, vol. 50, pp. 103-116, 2000

- [6.24] A. E. Morgan, E. K. Broadbent, M. Delfino, B. Coulman, D. K. Sadana
„Characterization of a Self-Aligned Cobalt Silicide Process“, J. Electrochem. Soc.: Solid-State Science and Technology, vol. 134, no. 4, pp. 925-935, April 1987
- [6.25] S. P. Murarka
„Self-aligned Silicides or metals for very large integrated circuit applications“, J. Vac. Sci. Technol. B 4 (6), pp. 1325-1331, Nov./Dec. 1986
- [6.26] J. F. Luy, R. Russer
„Silicon-Based Millimeter-Wave Devices“, Springer-Verlag, Berlin, pp. 89-140, 1994
- [6.27] M. Glück
„Herstellung von sub- μm Si/SiGe Feldeffekttransistoren (HFETs) und ersten Testschaltungen für Hochfrequenzanwendungen“, DDD-Verlag Darmstadt, pp. 66-76, 1999
- [6.28] S. C. H. Ho, M. C. Poon, M. Chan, H. Wong
„Thermal Stability of Nickel Silicides in Different Silicon Substrates“, IEEE HKEDM, Hong Kong, pp. 105-108, May 1998
- [6.29] R. D. Thompson, K. N. Tu, J. Angillelo, S. Delage, S. S. Iyer
„Interfacial reaction between Ni and MBE grown SiGe alloy“, Electrochemical Society, New York, pp. 235-242, 1988
- [6.30] G. Sarcona, S. K. Saha, M. K. Hatalis
„Nickel Silicides Grown on Amorphous Silicon and Silicon-Germanium Thin Film“, Electrochemical and Solid-State Letters, 1 (5), pp. 233-234, 1998
- [6.31] D. K. Schroder
„Semiconductor Material and Device Characterization“, John Wiley and Sons, New York, pp. 238-248, 1998
- [6.32] I. I. Amirov, V. A. Fedorov
„Fabrication of Submicron Structures by Anisotropically Etching Resists in Oxygen Plasma Produced by a Radio-Frequency Induction Discharge“, Russian Microelectronics, vol. 29, no. 1, pp. 35-42, 2000
- [6.33] J. A. G. Baggerman, R. J. Visser, E. J. H. Collart
„Ion-induced etching of organic polymers in argon and oxygen radio-frequency plasmas“, J. Appl. Phys. 75 (2), pp. 758-769, 15. January 1994

- [6.34] D. Behammer
„Niedertemperaturtechnologien zur Herstellung von skalierfähigen Si/SiGe/Si-Heterobipolartransistoren“, Dissertation, Ruhr-Universität Bochum, pp. 41-67, 1996
- [6.35] M. Tawata, H. Kato, S. Morita, S. Hattori
„Totally Dried Three-Layer Resist Process Using Evaporated Fatty Acid“, J. Electrochem. Soc. vol. 140, no. 3, pp. 861-865, March 1993
- [6.36] M. A. Hartney, D. W. Hess, D. S. Soane
„Critical Review: Oxygen plasma etching for resist stripping and multilayer lithography“, J. Vac. Sci. Technol. B 7 (1), pp. 1-13, Jan./Feb. 1989
- [6.37] W. H. Juan, S. W. Pang
„High aspect ratio polyimide etching using an oxygen plasma generated by electron cyclotron resonance source“, J. Vac. Sci. Technol. B 12 (1), pp. 422-426, Jan./Feb. 1994
- [6.38] H. H. Richter, A. Wolff, B. Tillack, T. Skaloud
„Optical emission end point detection for reactive ion etching of Si/SiGe structures“, Materials Science and Engineering, B 27, pp. 39-45, 1994
- [6.39] T. J. Dalton, W. T. Conner, H. H. Sawin
„Interferometric Real-Time Measurement of Uniformity for Plasma Etching“, J. Electrochem. Soc. vol. 141, no. 7, pp. 1893-1900, July 1994
- [6.40] F. Deng, R. A. Johnson, P. M. Asbeck, S. S. Lau, W. B. Dobbelday, T. Hsiao, J. Woo
„Salicidation process using NiSi and its device application“, J. Appl. Phys. 81 (12), pp. 8047-8051, 15 June 1997
- [6.41] D. Behammer
„Produktionstechniken der Halbleitertechnologie“, Vorlesungsskript, Universität Ulm, Institut für elektronische Bauelemente und Schaltungen, WS 00/01
- [6.42] D. Widmann, H. Mader, H. Friedrich
„Technologie hochintegrierter Schaltungen“, Springer Verlag Berlin, pp. 75-80, 1988
- [6.43] G. Schumicki, P. Seegebrecht
„Prozeßtechnologie“, Springer-Verlag, Berlin, pp. 223-236, 1991

- [6.44] S. Fujii, M. Fukumoto, G. Fuse, T. Ohzone
„A Planarization Technology Using a Bias-Deposited Dielectric Film and an Etch-Back Process“, IEEE Transactions on Electron Devices, vol. 35, no. 11, pp. 1829-1833, Nov. 1988
- [6.45] Y. Kuo
„Planarization of Multilevel Metalization Processes: A Critical Review“, SPIE, vol. 797, Advanced Processing of Semiconductor Devices, pp. 49-60, 1987
- [6.46] L. K. White
„Planarization Properties of Resist and Polyimide“, Journal of the Electrochemical Society: Solid-State Science and Technology, pp. 1542-1548, July 1983
- [6.47] L. E. Stillwagon, R. G. Larson, G. N. Taylor
„Planarization of Substrate Topography by Spin Coating“, Electrochemical Society: Solid-State Science and Technology, pp. 2030-2037, August 1987
- [6.48] L. E. Stillwagon
„Planarization of Substrate Topography by Spin-Coated Films: A Review“, Solid State Technology, pp. 67-71, June 1987
- [6.49] G. Grivna, R. Goodner
„A New Planarization Technique Using In Situ Isotropic Photoresist Mask Erosion“, J. Electrochem. Soc., vol. 141, no. 1, pp. 251-254, Jan. 1994
- [6.50] International Technology Roadmap for Semiconductors, 1999 Edition, Semiconductor Industry Association, 1999
- [6.51] D. S. DeBear, J. A. Levert, S. P. Mukherjee
„Spin-etch planarization for dual damascene interconnect structures“, Solid State Technology, pp. 53-60, March 2000
- [6.52] S. M. Sze
„Physics of Semiconductor Devices“, John Wiley & Sons, Inc., New York, pp. 245-311, 1981
- [6.53] H. Morkoc, H. Unglu, G. Ji
„Principles and Technology of MODFETs“, vol. 1, John Wiley & Sons, Inc., New York, pp. 65-75, 1991
- [6.54] M. Wittmer
„Conduction mechanism in PtSi/Si Schottky diodes“, Physical Review B, vol. 43, no. 5, pp. 4385-4395, Feb. 1991

- [6.55] J. F. Luy, P. Russer
„Silicon-Based Millimeter-Wave Devices“, Springer Verlag, Berlin,
pp. 89-148, 1994
- [6.56] R. Matz, R. Purstel, Y. Yokata, G. W. Rubloff, P. S. Ho
„Chemical reaction and silicideformation at the Pt/Si interface“, J. Vac.
Sci. Technol. A2 (2), p. 253, 1984
- [6.57] G. Rossi, I. Abbati, L. Braichovich, I. Lindau, W. E. Spicer
„Si(111)-Pt interface at room temperature: A synchrotron radiation
photoemission study“, Physical Review B 25 (6), p. 4307, 1982
- [6.58] C. A. Chang
„PtSi contact metallurgy: Effect of silicide formation process“, J. Appl.
Phys. 58, p.3258, 1985
- [6.59] M. Wittmer
„Growth kinetics of platinum silicide“, J. Appl. Phys. 54, p.5081, 1983
- [6.60] M. Wittmer
„Current transport in high-barrier IrSi/Si Schottky diodes“, Physical
Review B, vol. 42, no. 8, pp. 5249-5259, Sep. 1990
- [6.61] H. Beneking
„Halbleitertechnologie“, Teubner Verlag, Stuttgart, S. 234-243, 1991
- [6.62] D. K. Ferry, R. O. Grondin
„Physics of submicron devices“, Plenum Press, New York, pp. 52-68,
1991
- [6.63] B. E. Maile
„Herstellung und optische Analyse niederdimensionaler Verbindungs-
halbleiterstrukturen“, Dissertation, Physikal. Institut der Universität
Stuttgart, S. 20-28, 1990
- [6.64] N. Saitou
„Electron-Beam Lithography“ in „Handbook of Semiconductor Manu-
facturing Technology“, Marcel Dekker Inc., New York, pp. 571-587,
2000
- [6.65] B. E. Maile
„Fabrication limits of nanometer T and Γ gates: Theory and experi-
ment“, J. Vac. Sci. Technology B 11 (6), pp. 2502-2508, Nov. 1993
- [6.66] B. Raynor
Fraunhofer Institut für Angewandte Festkörperphysik, persönliche
Mitteilung

- [6.67] W. Prost
„Technologie der III/V Halbleiter“, Springer Verlag, S. 144-150, 1997
- [6.68] N. G. Einspruch, S. S. Cohen, G. S. Goldenblatt
„VLSI Metallization“, vol. 15 of VLSI Electronics Microstructure Science, Academic Press, Inc., London, pp. 43-68, 1987
- [6.69] A. E. Braun
„Aluminum Persists as Copper Age Draws“, Semiconductor International, August 1999
- [6.70] M. E. Thomas, R. H. Havemann
„Overview of Interconnect“ in „Handbook of Semiconductor Manufacturing Technology“, Marcel Dekker Inc., New York, pp. 287-308, 2000
- [6.71] P. C. Andricacos, C. Uzoh, J. O. Dukovic, J. Horkans, H. Deligianni
„Damascene copper electroplating for chip interconnections“, IBM Journal of Research & Development, vol. 42, No. 5, 1998
- [6.72] A. E. Braun
„Copper Electroplating Enters Mainstream Processing“, Semiconductor International, Cover Story, April 1999
- [6.73] S. M. Rossnagel
„Physical Vapor Deposition“ in „Handbook of Semiconductor Manufacturing Technology“, Marcel Dekker Inc., New York, pp. 395-414, 2000
- [6.74] T. Smy, L. Tan, K. Chan, R. N. Tait, J. N. Broughton, S. K. Dew, M. J. Brett
„A Simulation Study of Long Throw Sputtering of Diffusion Barrier Deposition into High Aspect Vias and Contacts“, IEEE Transactions on Electron Devices, vol. 45, no. 7, pp. 1414-1425, July 1998
- [6.75] G. Engelmann, O. Ehrmann, R. Leutenbauer, H. Schmitz, H. Reichl
„Selectively released microstructures electroplated into thick positive photoresists“, J. Micromech. Microeng. 3, pp. 152-154, 1993
- [7.1] M. Zeuner, A. Fox, T. Hackbarth, D. Behammer, U. König
„90 GHz f_T HFET with fully optical self-aligned sub 100 nm gate“, IEEE DRC 2002, Santa Barbara, pp. 53-54, 2002
- [7.2] M. Shur
„GaAs Devices and Circuits“, Plenum Press, New York, pp. 372-381, 1989

- [7.3] D. K. Schroder
„Semiconductor Material and Device Characterization“, John Wiley & Sons, Inc., New York, pp. 238-241, 1998
- [7.4] K. Lee, M. Shur, K. W. Lee, T. Vu, P. Roberts, M. Helix
„A New Interpretation of "End" Resistance Measurements“, IEEE Electron Device Letters, vol. 5, no. 1, pp. 5-7, January 1984
- [7.5] H. Rohdin, N. Moll, C.Y Su, G. S. Lee
„Interfacial Gate Resistance in Schottky-Barrier-Gate Field-Effect Transistors“, IEEE Transactions on Electron Devices, vol. 45, no. 12, pp. 2407-2416, Dec. 1998
- [7.6] M. S. Gupta
„Power Gain in Feedback Amplifiers, a Classic Revisited“, IEEE Transactions on Microwave Theory and Techniques, vol. 40, no. 5, pp. 864-879, 1992
- [7.7] K. Ismail
„Si/SiGe High-Speed Field-Effect Transistors“, Proceedings of the IEDM 1995, Washington, pp. 509-512, 1995
- [7.8] K. Ismail, S. Rishton, J. O. Chu, K. Chan, B. S. Meyerson
„High-Performance Si/SiGe n-Type Modulation-Doped Transistors“, IEEE Electron Device Letters, vol. 14, no. 7, pp. 348-350, 1993
- [7.9] M. Glück
„Herstellung von sub- μm Si/SiGe Heterostruktur-Feldeffekttransistoren und ersten Testschaltungen für Hochfrequenzanwendungen“, Dissertation, Fakultät für Ingenieurwissenschaften, Uni Ulm, S. 130-149, 1999
- [7.10] B. Metzger
„CMOS devices steal show“, Compound Semiconductor, pp. 31-33, Jan./Feb. 2002
- [7.11] E. M. Chumbes, A. T. Schremer, J. A. Smart, D. Hogue, J. Komiak, J. R. Shealy
„Microwave Performance of AlGaIn/GaN High Electron Mobility Transistors on Si (111) Substrat“, Proceedings of the IEDM 1999, Washington, Dec. 1999
- [7.12] M. H. W. Hoffmann
„Hochfrequenztechnik“, Springer Verlag, Berlin, S. 497-503, 1997
- [7.13] H. Morkoc, H. Unlu, G. Ji
„Principles and Technology of MODFET“, vol. 2, John Wiley and Sons Ltd., Chichester, pp. 414-429, 1991

- [7.14] D. K. Ferry, R. O. Grondin
„Physics of Submicron Devices“, Plenum Press, New York, pp. 363-380, 1991
- [7.15] D. K. Ferry, R. O. Grondin
„Physics of Submicron Devices“, Plenum Press, New York, pp. 385-391, 1991
- [7.16] M. S. Gupta, O. Pitzalis, S. E. Rosenbaum, P.T. Greiling
„Microwave Noise Characterization of GaAs MESFET's: Evaluation by On-Wafer Low-Frequency Output Noise Current Measurement“, IEEE Trans. Microwave Theory Tech., vol. MTT-35, pp. 1208-1218, 1987
- [7.17] H. Fukui
„Design of Microwave GaAs MESFET's for Broad-Band Low-Noise Amplifiers“, IEEE Trans. Microwave Theory Tech., vol. MTT-27, pp. 643-650, July 1979
- [7.18] H. Fukui
„Optimal Noise Figure of Microwave GaAs MESFET's“, IEEE Trans. Electron Devices, vol. ED-26, pp. 1032-1037, July 1979
- [7.19] R. A. Pucel, H. A. Haus, H. Statz
„Signal and Noise Properties of Gallium Arsenide Microwave Field-Effect Transistors“, Advances in Electronics and Electron Physics, vol. 38, Academic Press, pp. 195-265, 1981
- [7.20] M. W. Pospieszalski
„Modeling of Noise Parameters of MESFET's and MODFET's and Their Frequency and Temperature Dependence“, IEEE Transactions on Microwave Theory and Techniques, vol. 37, no. 9, 1340-1350, 1989
- [7.21] G. Dambrine, J. P. Raskin, F. Danneville, D. Vanhoenacker-Janvier, J. P. Colinge, A. Cappy
„High-Frequency Four Noise Parameters of Silicon-on-Insulator-Based Technology MOSFET for the Design of Low-Noise RF Integrated Circuits“, IEEE Trans. on Electron Devices, vol. 46, no. 8, 1733-1741, Aug. 1999
- [7.22] A. Cappy, A. Vanoverschelde, M. Schortgen, C. Versnaeyen, G. Salmer
„Noise Modeling in Submicrometer-Gate Two-Dimensional Electron-Gas Field-Effect Transistors“, IEEE Transactions on Electron Devices, vol. ED-32, no. 12, pp. 2787-2795, Dec. 1985

- [7.23] M. Enciso, F. Aniel, P. Crozat, R. Adde, M. Zeuner, A. Fox, T. Hackbarth
„0.3 dB Minimum Noise Figure of 0.13 μm Gate Length Strained Si/Si_{0.58}Ge_{0.42} n-MODFETs“, IEEE Electronics Letters, vol. 37, no. 17, pp. 1089-1090, August 2001
- [7.24] M. Enciso, F. Aniel, P. Crozat, L. Giguerre, R. Adde, M. Zeuner, G. Höck, A. Fox
„Ultra-low noise strained Si/SiGe n- and Ge/SiGe p-MODFETs“, ESSDERC 2001, Nürnberg, September 2001
- [7.25] V. Danelon, P. Crozat, F. Aniel, G. Vernet
„Four noise parameter determination method for transistors based on the frequency dependence of the noise figure“, Electronics Letters, vol. 34, no. 16, pp. 1612-1613, Aug. 1998
- [7.26] F. Schäffler
„High-mobility Si and Ge structures“, Semicond. Sci. Technol., vol. 12, pp. 1515-1549, 1997
- [7.27] F. Aniel, N. Zerounian, R. Adde, M. Zeuner, T. Hackbarth, U. König
„Low Temperature Analysis of 0.25 μm T-Gate Strained Si/Si_{0.55}Ge_{0.45} n-MODFETs“, IEEE Trans. on Electron Devices, vol. 47, no. 7, pp. 1477-1484, July 2000
- [7.28] M. Zeuner, U. König
„SiGe HFET Technology“, Proceedings of the European Microwave Week 2001, London, September 2001

Verzeichnis der im Rahmen dieser Arbeit entstandenen Publikationen und Patentanmeldungen

D. Behammer, M. Zeuner, T. Hackbarth, H.-J. Herzog, M. Schäfer, T. Grabolla
„Comparison of lateral and vertical Si-MOSFETs with ultra short channels“;
Thin Solid Films, no. 336, pp. 313-318, 1998

T. Hackbarth, G. Höck, H.-J. Herzog, M. Zeuner
„Strain relieved SiGe buffers for Si-based heterostructure field-effect transistors“,
Journal of Crystal Growth, no. 201/202, pp. 734-738, 1999

M. Zeuner, T. Hackbarth, U. König, A. Gruhle, F. Aniel
„High Performance 0.25 μm T-gate SiGe n-MODFET“, Device Research Conference,
Santa Barbara, pp. 177-178, Juni 1999

T. Hackbarth, H.-J. Herzog, M. Zeuner, H. Kibbel, G. Höck, E. A. Fitzgerald, M. Bulsara,
C. Rosenblad, H. v. Känel, S. Mantl, B. Holländer
„Alternatives to thick MBE-grown relaxed SiGe buffers“, Int. Joint Conf. On Epitaxy
and Heterostructures, Miyagi, Japan, September 1999

U. König, M. Zeuner, G. Höck, T. Hackbarth, M. Glück, T. Ostermann, M. Saxarra
„n- and p-Type SiGe HFETs and circuits“, Solid State Electronics, no. 43,
pp. 1383-1388, 1999

M. Zeuner, T. Hackbarth, G. Höck, D. Behammer, U. König
„High-Frequency SiGe n-MODFET for Microwave Applications“, IEEE Microwave
and Guided Wave Letters, vol. 9, no. 10, Oct. 1999

M. Zeuner, A. Gruhle, U. König
„Kryotemperatur 200 GHz SiGe Bauelemente“, Kryoelektrische Bauelemente
1999, Köln, Oktober 1999

F. Aniel, N. Zerounian, R. Adde, M. Zeuner, T. Hackbarth, U. König
„Low Temperature Analysis of 0.25 μm T-Gate Strained Si/Si_{0.55}Ge_{0.45} n-MODFETs“,
IEEE Trans. on Electron Devices, vol. 47, no. 7, pp. 1477-1484, July 2000

M. Enciso, L. Giguere, F. Aniel, P. Crozat, R. Adde, M. Zeuner, H.-J. Herzog, U. König

„Optimization of HF and noise performance of Si/SiGe n-MODFETs at 77 K“, 3^o Journées nationales Hétérostructures à semiconducteurs IV-IV, Paris, Juli 2000

G. Höck, M. Zeuner, T. Hackbarth

„RF-performance of n- and p-type SiGe MODFETs with respect to complementary integration“, 1. European Workshop on Ultimate Integration of Si, ULIS Grenoble, 2000

H.-J. Herzog, T. Hackbarth, G. Höck, M. Zeuner, U. König

„SiGe based FETs: Buffer issues and device results“, Thin Solid Films, no. 380, pp. 36-41, 2000

M. Enciso, F. Aniel, P. Crozat, R. Adde, M. Zeuner, A. Fox, T. Hackbarth

„0.3 dB Minimum Noise Figure of 0.13 μm Gate length Strained Si/Si_{0.58}Ge_{0.42} n-MODFETs“, IEEE Electronics Letters, vol. 37, no. 17, pp. 1089-1090, August 2001

M. Enciso, F. Aniel, L. Giguere, P. Crozat, R. Adde, M. Zeuner, A. Fox, T. Hackbarth

„High frequency properties of Si/SiGe n-MODFETs: dependence on gate-length and temperature“, European Microwave Week 2001, GAAS 2001 Conference Proceedings, pp. 187-190, London, September 2001

M. Zeuner, U. König

„SiGe HFET Technology“, invited talk, European Microwave Week 2001, London, September 2001

M. Enciso, F. Aniel, P. Crozat, L. Giguere, R. Adde, M. Zeuner, G. Höck, A. Fox

„Ultra-low noise strained Si/SiGe n- and Ge/SiGe p-MODFETs“, ESSDERC 2001, Nürnberg, September 2001

M. Zeuner, A. Fox, T. Hackbarth, D. Behammer, U. König

„90 GHz f_T SiGe HFET with fully optical self-aligned sub 100 nm gate“, Device Research Conference, Santa Barbara, pp. 53-54, Juni 2002

I. Kallfass, M. Zeuner, Hermann Schumacher, Thomas J. Brazil
„A DC to 40 GHz Large Signal Model for n-channel SiGe HFET Transistors“,
zur Veröffentlichung eingereicht bei IEEE Electron Devices

D. Behammer, M. Zeuner
„Selbstjustierendes Verfahren zur Herstellung von Feldeffekttransistoren“,
Amtl. Aktenzeichen 19936005.7

D. Behammer, M. Zeuner
„Verfahren zur Herstellung von FET mit Dummy-Gate und Schrumpfätzung“,
Amtl. Aktenzeichen 10053671.9

Nachwort

Die vorliegende Arbeit entstand im Rahmen meiner Tätigkeit als wissenschaftlicher Mitarbeiter in der Abteilung REM/CS am Forschungszentrum der DaimlerChrysler AG in Ulm, unter Betreuung von Herrn Prof. Dr. Karl Hofmann vom Institut für Halbleiterbauelemente und Werkstoffe der Universität Hannover.

Mein besonderer Dank für die Unterstützung bei dieser Arbeit gilt:

Herrn Dr. Ulf König, dem Leiter der Abteilung REM/CS, der mir die Möglichkeit zur Durchführung dieser Arbeit gegeben hat, mir stets mit seiner Diskussionsbereitschaft zur Seite stand und mich durch seine vorbildliche Begeisterungsfähigkeit motiviert hat.

Herrn Prof. Dr. Karl Hofmann für die wohlwollende Förderung meiner wissenschaftlichen Tätigkeit, sein entgegengebrachtes Vertrauen und die Übernahme des Erstgutachtens.

Herrn Prof. Dr.-Ing. Karl-Heinz Löcherer danke ich für das große Interesse an meiner Arbeit und die Übernahme des Zweitgutachtens.

Eine besondere Freundschaft verbindet mich mit Herrn Dr. Dag Behammer, der meine bisherige berufliche Laufbahn hilfreich und motivierend unterstützt hat und mir durch seine unschätzbare Fachkompetenz bei wissenschaftlichen Problemen stets neue Perspektiven eröffnen konnte.

Des weiteren bin ich den Kolleginnen und Kollegen der Abteilung REM/CS am Forschungszentrum der DaimlerChrysler AG für das freundliche und harmonische Arbeitsklima und für die vielfältige Unterstützung zu Dank verpflichtet. Besonders namentlich erwähnen möchte ich die Herren Dr. Thomas Hackbarth, Horst Kibbel, Peter Muthsam und Frau Jessica Kuchenbecker für die mir zahlreich zur Verfügung gestellten Epitaxieschichten. Außerdem danke ich Frau Ursula Riek für die Analyse der unzähligen Proben und Wafer am Rasterelektronenmikroskop und Frau Andrea Häussler für die kreative Mithilfe bei diversen

Präsentationen und Dokumentationen. Weiterer Dank gilt Frau Sonja König und den Herren Alexander Klaiber, Karl-Heinz Hieber, Günther Kohn und Dieter Eisbrenner für die Unterstützung bei vielen technologischen Problemen und Arbeiten und meinen Kollegen den Herren Hans-Joest Herzog, Dr. Hartmut Presting und Georg Höck für die wertvollen wissenschaftlichen Diskussionen und Anregungen.

Auf universitärer Seite bedanke ich mich bei den Mitarbeitern des Institutes für Halbleiterbauelemente und Werkstoffe der Universität Hannover für die harmonische und anregende wissenschaftliche Kooperation. Besonders erwähnen möchte hier ich die Herren Dr. Björn Fischer und Niels Hoffmann, die mir immer kompetent und mit voller Einsatzfreude zur Seite standen.

An der Universität Ulm entwickelte sich eine intensive Zusammenarbeit mit der Abteilung für Elektronische Bauelemente und Schaltungen unter der Leitung von Herrn Prof. Dr. Hermann Schumacher. Hier möchte ich mich ganz herzlich bei Frau Ursula Spitzberg für die zahlreichen Hochfrequenzmessungen bedanken, ebenso wie bei Herrn Norbert Käß für die Durchführung der Elektronenstrahlithografie. Den Herren Kai-Boris Schad, Ertugrul Sönmez, Ingmar Kallfass und Peter Abele verdanke ich wertvolle wissenschaftliche Erkenntnisse im Bereich der Transistormodellierung und des Schaltungsdesigns.

Mit der Abteilung für Mikroelektronische Systeme der Universität Linz, geleitet von Herrn Prof. Dr. R. Hagelauer, verbindet mich eine enge Kooperation auf den Gebieten des Schaltungsentwurfs und der Erstellung neuer Maskenlayouts. Besonders anführen möchte ich die angenehme und erfolgreiche Zusammenarbeit mit Herrn Dr. Timm Ostermann.

Wesentliche Bestandteile meiner Arbeit entstanden im Rahmen der EU-Förderprojekte VAHMOS 2000 und SIGMUND in Kooperation mit dem Interuniversity Microelectronics Centre (IMEC) in Leuven, Belgien, dem Institut d' Electronique Fondamentale (IEF/CNRS) in Paris Süd, Frankreich und der Universität Stuttgart. Nur durch intensive gemeinsame Forschungsaktivitäten mit den oben aufgeführten Partnern ist es mir gelungen, die an mich gestellte Aufgabe zu erfüllen und damit meinen Anteil zu diesen Projekten zu leisten. Ich möchte deshalb diese Gelegenheit nutzen, allen Beteiligten für die kreative und fruchtbare Zusammenarbeit, ihre wissenschaftliche Fachkenntnis und ihre Einsatzfreude zu danken.

Ein weiterer Dank gilt allen Studenten, Praktikanten und Diplomanden, die ich im Laufe meiner wissenschaftlichen Tätigkeit betreuen durfte. Sie haben mich durch ihr Engagement und ihre Beiträge wesentlich in meiner Arbeit unterstützt. Namentlich erwähnen möchte ich hier die Herren Alexander Fox, Andreas Pribil, Dietmar Stolz, Rene Maßmann, Hezi Rahamim, Thomas Hager, Torsten Mack und Tom Zimmermann.

Meiner Familie danke ich für die Förderung meines Bildungsweges und die Unterstützung meines Werdeganges.

Ganz besonders weiß ich die Geduld, die Rücksichtnahme und die Motivation für diese Arbeit seitens meiner Freundin Juliane und meiner Freunde zu schätzen, bei ihnen habe ich stets Verständnis und Rückhalt auch in schwierigen Situationen erhalten, um diese Aufgabe zu bewältigen.

Marco Zeuner

Ulm, im Juli 2002

Lebenslauf

- 04.06.1970 Geburt in Lemgo/Ostwestfalen
- 1976 - 1980 Grundschule V in Lemgo
- 1980 - 1981 Hauptschule II in Lemgo
- 1981 - 1987 Bürgermeister-Gräfer-Realschule in Lemgo
- 1987 - 1990 Marianne-Weber-Gymnasium in Lemgo
- 1990 - 1991 Grundwehrdienst
- 1991 - 1997 Studium der Elektrotechnik an der Universität Hannover
Studienschwerpunkt: Mikroelektronik
- 1997 Diplomarbeit am Forschungszentrum der Daimler-Benz AG
in Ulm, Abteilung FT2/HS für Si/SiGe-Höchstfrequenz-
elektronik (Dr. Ulf König) in Zusammenarbeit mit dem In-
stitut für Halbleiterbauelemente und Werkstoffe an der Uni-
versität Hannover (Prof. Dr. K. Hofmann)
- 1998 - 2001 Wissenschaftlicher Mitarbeiter am Forschungszentrum der
DaimlerChrysler AG in Ulm, Abteilung REM/CS, externe
Doktorarbeit in Zusammenarbeit mit dem Institut für Halb-
leiterbauelemente und Werkstoffe an der Universität Hanno-
ver (Prof. Dr. K. Hofmann)
- 2001 - 2002 Mitglied der Austauschgruppe der DaimlerChrysler AG
- seit 2002 Mitarbeiter des Bereiches Produktstrategie in der PKW-
Entwicklung der DaimlerChrysler AG
Schwerpunkt: Innovationsmanagement

