



Institut für Schichten und Grenzflächen  
Institut 1: Halbleiterschichten und -bauelemente

***Entwicklung und Charakterisierung vertikaler  
Double-Gate-MOS-Feldeffekttransistoren***

*Stefan Trellenkamp*





# ***Entwicklung und Charakterisierung vertikaler Double-Gate-MOS-Feldeffekttransistoren***

*Stefan Trellenkamp*

**Berichte des Forschungszentrums Jülich ; 4139**

ISSN 0944-2952

Institut für Schichten und Grenzflächen

Institut 1: Halbleiterschichten und -bauelemente JÜl-4139

D 82 (Diss., Aachen, RWTH, 2003)

Zu beziehen durch: Forschungszentrum Jülich GmbH · Zentralbibliothek

D-52425 Jülich · Bundesrepublik Deutschland

☎ 02461 61-5220 · Telefax: 02461 61-6103 · e-mail: [zb-publikation@fz-juelich.de](mailto:zb-publikation@fz-juelich.de)

## Entwicklung und Charakterisierung vertikaler Double-Gate-MOS-Feldeffekttransistoren

**Kurzfassung.** Planare MOS-Feldeffekttransistoren sind derzeit übliche Bauelemente in den meisten Produkten der Computerindustrie. Können diese Bauelemente nicht weiter verkleinert werden, sind neue Konzepte für alternative Transistorentwürfe vonnöten. In der vorliegenden Arbeit wird die Entwicklung vertikaler Double-Gate-Feldeffekttransistoren vorgestellt, ein Transistortyp, bei dem der p-n-p- (bzw. n-p-n-)Übergang vertikal angeordnet ist und somit der Source-Drain-Strom senkrecht zur Waferoberfläche fließt. Double-Gate-Feldeffekttransistoren zeichnen sich dadurch aus, daß ein sehr dünnes Kanalgebiet auf beiden Seiten von Gateelektroden flankiert wird, die parallel geschaltet sind. Die Symmetrie und das verminderte Bulkvolumen dieses Bauelementes versprechen ein verbessertes Skalierpotential sowie eine bessere Gatekontrolle des Kanalgebietes und damit die Unterdrückung von Kurzkanaleffekten.

Die Nanostrukturierung des aktiven Bereichs eines vertikalen Double-Gate-Transistors ist höchst anspruchsvoll, da ca. 300 nm hohe und bis zu 30 nm schmale Siliziumstege reproduzierbar hergestellt werden müssen. Der Einsatz von Hydrogensilsesquioxan (HSQ) als hochauflösender anorganischer Elektronenstrahlresist bildet die Grundlage dieser Nanostrukturierung. In HSQ definierte Strukturen werden mittels ICP-RIE (reaktives Ionenätzen mit induktiv gekoppeltem Plasma) sehr anisotrop und mit hoher Selektivität in Silizium übertragen. 25 nm breite und 330 nm hohe Siliziumstege können so hergestellt werden.

Realisiert werden verschiedene Transistorausführungen, deren Kanallängen vor der Nanostrukturierung durch epitaktisches Wachstum dotierter Siliziumschichten oder nach der Nanostrukturierung mittels Ionenimplantation eingestellt sind.

Die realisierten Transistoren weisen Source-Drain-Ströme bis zu  $380 \mu\text{A}/\mu\text{m}$  und Steilheiten bis zu  $480 \mu\text{S}/\mu\text{m}$  auf. Es wird gezeigt, daß sich bei Abnahme der Siliziumstegbreite das Kurzkanalverhalten verbessert.

## Development and characterization of vertical Double-Gate-MOS-field-effect transistors

**Abstract.** Planar MOS-field-effect transistors are common devices today used by the computer industry. When their miniaturization reaches its limit, alternate transistor concepts become necessary. In this thesis the development of vertical Double-Gate-MOS-field-effect transistors is presented. These types of transistors have a vertically aligned p-n-p junction (or n-p-n junction, respectively). Consequently, the source-drain current flows perpendicular with respect to the surface of the wafer. A Double-Gate-field-effect transistor is characterized by a very thin channel region framed by two parallel gates. Due to the symmetry of the structure and less bulk volume better gate control and hence better short channel behavior is expected, as well as an improved scaling potential.

Nanostructuring of the transistor's active region is very challenging. Approximately 300 nm high and down to 30 nm wide silicon ridges are requisite. They can be realized using hydrogen silsesquioxane (HSQ) as inorganic high resolution resist for electron beam lithography. Structures defined in HSQ are then transferred with high anisotropy and selectivity into silicon using ICP-RIE (reactive ion etching with inductive coupled plasma). 25 nm wide and 330 nm high silicon ridges are achieved.

Different transistor layouts are realized. The channel length is defined by epitaxial growth of doped silicon layers before or by ion implantation after nanostructuring, respectively.

The transistors show source-drain currents up to  $380 \mu\text{A}/\mu\text{m}$  and transconductances up to  $480 \mu\text{S}/\mu\text{m}$ . Improved short channel behavior for decreasing width of the silicon ridges is demonstrated.



# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>1</b>
<b>2</b>	<b>Das X-MOS-Projekt</b>	<b>3</b>
2.1	Der Schottky-Barrieren-MOSFET . . . . .	4
2.2	Der NovaFET . . . . .	5
<b>3</b>	<b>Der MOSFET</b>	<b>9</b>
3.1	Die MOS-Struktur . . . . .	9
3.2	Aufbau und Wirkungsweise . . . . .	12
3.3	Strom-Spannungs-Charakteristik . . . . .	13
3.4	MOSFET-Skalierung . . . . .	16
3.5	Kurzkanaleffekte . . . . .	18
3.5.1	$U_{th}$ -Roll-Off . . . . .	19
3.5.2	Drain-Induced Barrier Lowering . . . . .	19
3.5.3	Punch-Through . . . . .	20
3.5.4	Kanallängenmodulation . . . . .	21
3.5.5	Unterschwelldröme . . . . .	21
3.6	Neue MOSFET-Konzepte . . . . .	23
3.6.1	SOI-MOSFETs . . . . .	23
3.6.2	Double-Gate-MOSFETs . . . . .	24
<b>4</b>	<b>Verwendete Technologie</b>	<b>27</b>
4.1	Technologien zur Strukturierung . . . . .	27
4.1.1	Elektronenstrahl-Lithographie . . . . .	27
4.1.2	Optische Lithographie . . . . .	29
4.1.3	Strukturübertragung durch RIE-Ätzen . . . . .	30
4.2	Ionenimplantation . . . . .	33
4.3	SIMS . . . . .	34
4.4	Chemisch-Mechanisches Polieren (CMP) . . . . .	35
<b>5</b>	<b>Der NovaFET-Prozeß</b>	<b>39</b>
5.1	NovaFET – ursprüngliches Layout . . . . .	39
5.1.1	Prozeßfolge . . . . .	40
5.1.2	Schlüsselprozesse . . . . .	42
5.2	I-NovaFET – die Implantationsvariante . . . . .	44

5.2.1	Prozeßfolge . . . . .	45
5.2.2	Schlüsselprozesse . . . . .	47
5.3	EpiNovaFET . . . . .	53
5.3.1	Prozeßfolge . . . . .	53
5.3.2	Schlüsselprozesse . . . . .	54
<b>6</b>	<b>Nanostrukturierung</b>	<b>57</b>
6.1	Eigenschaften des anorganischen Moleküls HSQ . . . . .	58
6.2	Strukturdefinition mit HSQ . . . . .	58
6.3	Strukturübertragung mit Zweilagelack . . . . .	61
6.4	Strukturübertragung durch Direktätzen . . . . .	62
<b>7</b>	<b>Bauelementcharakterisierung</b>	<b>69</b>
7.1	NovaFET (ursprüngliches Layout) . . . . .	69
7.2	I-NovaFET . . . . .	71
7.3	EpiNovaFET . . . . .	72
7.4	Einordnung . . . . .	73
<b>8</b>	<b>Zusammenfassung und Ausblick</b>	<b>79</b>
<b>A</b>	<b>Prozeßfolgen</b>	<b>81</b>
A.1	Prozeßprotokoll NovaFET (ursprüngliches Layout) . . . . .	81
A.2	Prozeßprotokoll I-NovaFET . . . . .	84
A.3	Prozeßprotokoll EpiNovaFET . . . . .	89
<b>B</b>	<b>Abkürzungen</b>	<b>91</b>

# Kapitel 1

## Einleitung

Viele Jahre schon wird das Ende der erstaunlichen Entwicklung beschworen, die 1965 von G. Moore aus einer kurzen Vergangenheit in die Zukunft projiziert [1] und zur sich selbst erfüllenden Prophezeiung wurde: „Moore’s law“, das für die Miniaturisierung in der Halbleiter-Technologie eine Verdopplung der Transistor-Performance und eine Vervierfachung der Transistoranzahl auf einem Chip alle drei Jahre konstatiert, gilt auch nach über 40 Jahren noch. Lange Zeit galten technologische Grenzen als limitierend, insbesondere die Auflösung der Lithographie, denn für die Skalierung galten einfache Regeln. Während die technologischen Hürden immer weiter nach hinten verschoben werden konnten, treten nun vermehrt ökonomische und physikalische Herausforderungen in den Vordergrund: Die Ökonomie entscheidet letztendlich, ob eine gefundene Lösung umgesetzt werden kann, und die Physik verhindert eine weitere Skalierung nach altem Muster. Quantenmechanische Effekte werden in einigen Jahrzehnten erst erreicht: Die Wellenlänge des Elektrons in Silizium ist 5 nm, Energiequantisierung von einigen meV tritt bei 10 nm-Strukturen auf; die ultimative Grenze ist wohl der Atomabstand im Gitter von 0,4 nm [2].

Derzeit ist das wichtigste Bauelement der Metall-Oxid-Halbleiter-Feldeffekttransistor (MOSFET); beim erfolgreichen CMOS<sup>1</sup>-Design sind die logischen Schaltungen aus p- und n-Kanal-MOSFETs integriert aufgebaut. In Dimensionen, bei denen quantenmechanische Effekte dominieren, wird CMOS wohl schon keine Rolle mehr spielen – neue Arten wie Quantencomputer könnten bis dahin realisiert sein. Für die nächsten Jahrzehnte jedoch müssen Lösungen innerhalb von CMOS gefunden werden. Diese Entwicklung ist derzeit am Scheideweg: Die konventionelle Skalierung konnte sehr lange Zeit mit nur geringen Veränderungen wie z. B. Dotierprofile im Kanalgebiet durchgehalten werden, weiteres bloßes Verkleinern schadet aber der Performance mehr als es der Platzersparnis nutzt – man spricht allgemein von „Kurzkanalverhalten“, dem es zu begegnen gilt. Neuartige, nicht klassische CMOS-Bauelemente werden derzeit untersucht, um für kommende Chipgenerationen bereit zu stehen. Als *non-classical CMOS* bezeichnet die ITRS-Roadmap<sup>2</sup> Transistor-Layouts wie *Ultra-*

---

<sup>1</sup> *Complementary Metal-Oxide-Silicon*

<sup>2</sup> *National Technology Roadmap for Semiconductors*, von der *Semiconductor Industry Association* (SIA) seit 1992 regelmäßig erstellte Analyse der Halbleitertechnologie mit Ausblick auf die Erfordernisse für Chipgenerationen der nächsten 15 Jahre.

*Thin Body SOI*, *Vertikaler Transistor*, *FinFET* und *Double-Gate-Transistor* [3] (siehe auch 3.6). Welchen Weg CMOS dabei gehen wird, ist noch offen.

Einer dieser möglichen Wege ist der vertikale Double-Gate-MOSFET. Durch seinen symmetrischen Aufbau, senkrechtem Stromfluß und lithographieunabhängiger Kanallänge verspricht er gute Transistoreigenschaften und hohe Packungsdichte. Dieser Transistor war Bestandteil des Projektes „Extended Silicon MOSFET“ am ISG des Forschungszentrums Jülich, seine Entwicklung und Charakterisierung Ziel dieser Arbeit.

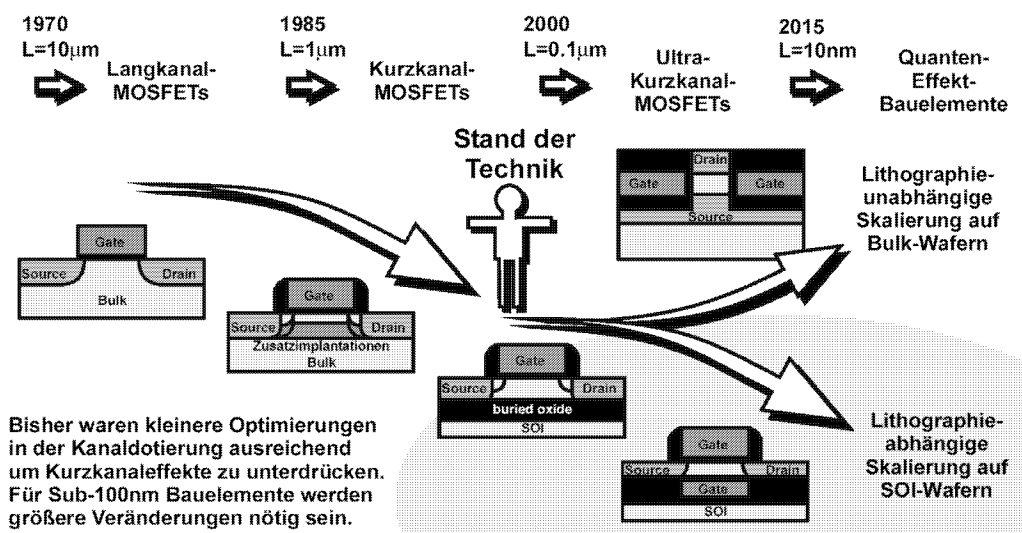


Abb. 1.1: *Evolution und mögliche Weiterentwicklung des MOSFETs (aus [4]).*

# Kapitel 2

## Das Projekt „Extended Silicon MOSFET“

Ausgehend von den Entwicklungen in der Technologie der Silizium-Bauelemente, setzt das Projekt „Extended Silicon MOSFET“ bei den nicht-klassischen Transistormodellen ein. An einem Punkt, an dem ein Ende der bisherigen Skalierung absehbar und die Richtung für neue Konzepte noch nicht entschieden ist, gilt es verschiedene Möglichkeiten zu untersuchen:

*„Übergeordnetes Ziel des Vorhabens „Extended MOSFET“ ist die Erforschung, Realisierung und Charakterisierung von Feldeffekttransistoren mit Kanallängen zwischen 50 und 100 nm. Dabei sollen sowohl die Physik als auch die Technologie zur Herstellung dieser Bauelemente weiter vorangebracht werden. Solche Ultrakurzkanaltransistoren versprechen höhere Geschwindigkeit, höhere Packungsdichte und kleinere Verlustleistungen.“*

(aus dem Projektantrag)

Das Projekt „Extended Silicon MOSFET“ bestand am Institut für Schichten und Grenzflächen (ISG 1) des Forschungszentrums Jülich von Juli 1999 bis Dezember 2002 und wurde vom HGF<sup>1</sup> gefördert.<sup>2</sup> Das Know-how und die Technik sind nicht zuletzt durch die Koordination des vorausgegangenen BMBF-Projektverbundes „NanoMOS“ (1995–1999) durch das ISI (jetzt ISG 1) erworben worden. Im Rahmen dieses Projektes wurden vertikale MOSFETs auf der Basis selektiver Epitaxie entwickelt und realisiert [5]. Das Kanalgebiet wurde dabei mit optischer Lithographie im Mikrometerbereich definiert, die Kanallänge über die Dicke einer Epitaxieschicht eingestellt.

---

<sup>1</sup>„Hermann von Helmholtz-Gemeinschaft Deutscher Forschungszentren“, Dachorganisation der fünfzehn deutschen Helmholtz-Zentren

<sup>2</sup>Förderkennzeichen 01SF9933/6

Das HGF-Strategiefondprojekt eröffnete nun die Möglichkeiten, Schlüsselprozesse der modernen Siliziumtechnologie zu integrieren, wie z. B. verbesserte optische Lithographie, Chemisch-Mechanisches Polieren, Niederenergie-Ionenimplantation, RTP<sup>3</sup> und ICP-RIE.<sup>4</sup> Durch diese Investitionen und die damit verbundene Erweiterung des Reinraums konnten die für dieses Projekt gesteckten anspruchsvollen Ziele weitestgehend erreicht werden.

Das Projekt war in zwei Arbeitsbereiche unterteilt: die Realisierung von vertikalen Double-Gate-MOSFETs („NovaFET“), in deren Rahmen die vorliegende Arbeit entstand, und die Erforschung eines selbst-assemblierten Verfahrens zur Herstellung von Nanostrukturen, die als Bausteine für laterale Schottky-Barrieren-MOSFETs (SB-MOSFETs) auf SOI-Substraten dienen (Abb. 2.1).

## 2.1 Der Schottky-Barrieren-MOSFET

Das selbst-assemblierte Verfahren für die Herstellung der SB-MOSFETs beruht auf der lokalen Oxidation von Silizidschichten („LOCOSI“, Local Oxidation of Silicides) [7]. Bei diesem Prozeß wird über ein elastisches Spannungsfeld, welches an den Kanten einer  $\text{Si}_3\text{N}_4/\text{SiO}_2$ -Maske entsteht, die Diffusion von Kobaltatomen während der lokalen Oxidation des Silizids gesteuert, so daß sich entlang dieser Kante eine wohldefinierte Aufspaltung formt [6]. Übertragen auf  $\text{CoSi}_2$ -Schichten auf ausgedünnten SOI-Substraten können mit dieser Methode homogene Auftrennungen mit Breiten bis auf 40 nm herunter erzielt werden (Abb. 2.2).

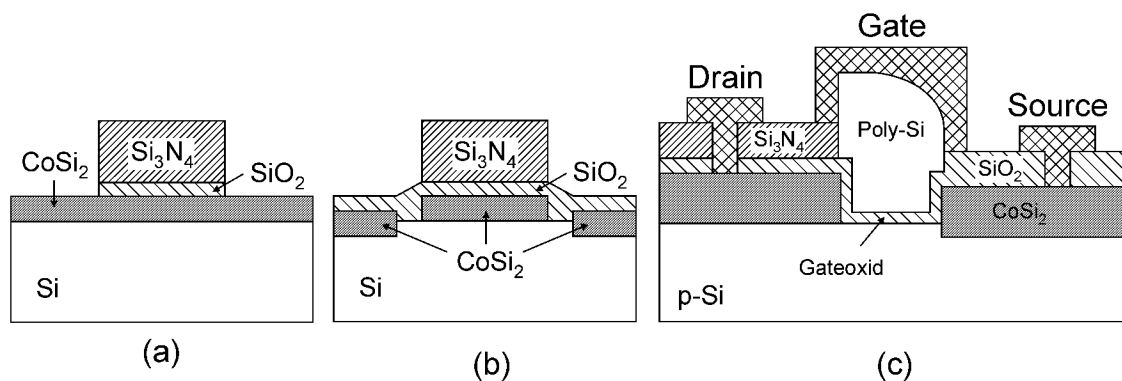


Abb. 2.1: Selbstjustierende Nanostrukturierung für laterale Schottky-Tunnel-MOSFETs. (a) Schichtsystem vor der Nanostrukturierung. (b) Bei der „rapid thermal oxidation“ (RTO) reißt die  $\text{CoSi}_2$ -Schicht an der Kante der Nitridmaske auf. Diese Spaltbreite, die deutlich unter 100 nm realisiert werden kann, bestimmt für den späteren Transistor die Kanallänge. (Die vertikale Verlagerung der  $\text{CoSi}_2$ -Schicht ist überzeichnet.) (c) Layout des Transistors. Der Übergang  $\text{CoSi}_2$  – Si ist hier ein Schottkykontakt. Vom Source zum Substrat ist dieser in Sperrrichtung, nach Anlegen der Gatespannung kann die Barriere jedoch durchtunnelt werden. Der Gatekontakt wird über einen Polysiliziumspacer hergestellt.

<sup>3</sup>„Rapid thermal processing“

<sup>4</sup>„Inductive Coupled Plasma-Reactive Ion Etching“, siehe 4.1.3.

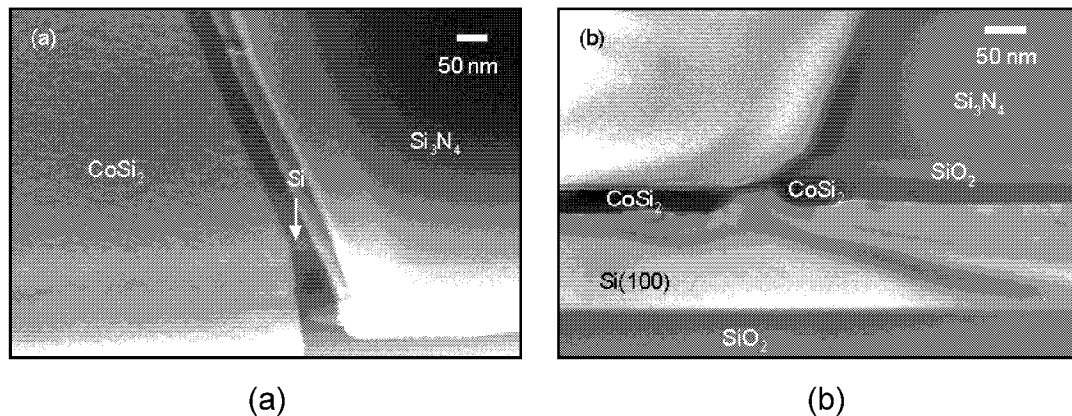


Abb. 2.2: (a) REM- und (b) XTEM-Aufnahmen der Strukturierung von  $\text{CoSi}_2$ -Schichten auf SOI-Substrat. [8]

## 2.2 Der NovaFET

Der in diesem Projekt realisierte NovaFET ist ein vertikaler Double-Gate-MOSFET, wie er in 3.6.2 beschrieben wird. Da die Kanallänge über die Dicke der mittleren Schicht der Dotierfolge definiert wird, erfordert diese Variante des DG-MOSFETs noch eine laterale Strukturdefinition im kritischen sub-100 nm-Bereich. In Abbildung 2.3 ist das Prinzip des NovaFETs dargestellt. (Die genaue Prozeßfolge ist in Kapitel 5 sowie im Anhang A dargestellt.) Die Schlüsselprozesse sind:

### 1. Nanolithographie

Es werden mehrere Mikrometer lange Stege benötigt, die ca. 300 nm hoch und 20–100 nm breit sind, um das Verhalten der Transistoren vergleichen zu können. Strukturen dieser Größenordnung herzustellen, wird auch in Zukunft aufwendig und teuer bleiben, jedoch werden für zukünftige Chipgenerationen neue parallele Lithographiemethoden entwickelt werden. So ist es gerechtfertigt, für Demonstratoren heute die aufwendige und langsame Elektronenstrahl-lithographie zu verwenden, zumal modernste Industrieeinrichtungen für die Forschung meist nicht verfügbar sind. Die Nanostrukturierung, bestehend aus Strukturdefinition und Strukturübertragung, ist in Kapitel 6 dargestellt.

### 2. Herstellung der Dotierfolge

Ausgehend von einer in einer LPCVD epitaktisch gewachsenen Dotierfolge für die erste Version des NovaFETs, wurden die Transistoreigenschaften bei den weiteren Versionen durch Implantation verbessert. Die Motivationen und Maßnahmen dazu werden in 5 erläutert.

### 3. Gateoxidation

Während zunächst die Gateoxidation durch RTP angewendet wurde, ist später eine Niedrigtemperatur-Feuchtoxidation etabliert worden. Diese Methode, die in 5.2.2 beschrieben ist, schon das thermische Budget.

#### 4. Topkontakt

Damit der Topkontakt direkt auf dem Steg angelegt werden kann, müssen verschiedene Voraussetzungen erfüllt sein:

- (a) Die Gatespacer müssen selektiv zurückgeätzt werden, damit sie nicht ebenfalls kontaktiert werden. Dies ist mit der ICP-RIE möglich geworden. Der Prozeß ist in 5.2.2 beschrieben; die Ätzprozesse sind die gleichen, wie sie für die Nanolithographie in 6 erläutert sind.
- (b) Eine Deckoxidschicht muß planarisiert werden. Dafür wurde eine konforme, dicke TEOS-Schicht durch Chemisch-Mechanisches Polieren planarisiert (siehe 5.2.2).

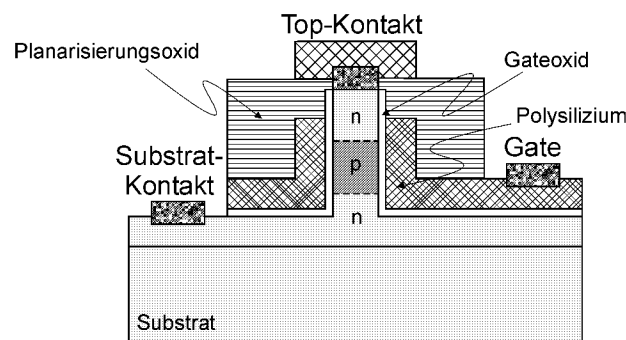


Abb. 2.3: Schema des vertikalen DG-MOSFETs.

Insgesamt sind durch die Auswertungen der Ergebnisse und das Potential der im Rahmen des Projektes neu installierten Technik drei verschiedene Variationen des NovaFETs entstanden:

##### 1. NovaFET (ursprüngliches Layout)

Im ersten Entwurf wird die Dotierfolge durch LPCVD-Epitaxie erstellt. Der obere Kontakt wird durch den dünnen Steg herausgeführt und außerhalb des aktiven Bereichs über ein Pad abgenommen. Diese Geometrie und die Top-Dotierung im Bereich  $4\text{--}5 \cdot 10^{18} \text{ at cm}^{-3}$  bewirken hohe Zuleitungswiderstände, wodurch die Transistoreigenschaften stark verschlechtert werden. Die Herstellung dieses Transistors ist in 5.1 dokumentiert.

##### 2. I-NovaFET

Flache Implantationen können Dotierungen im Bereich  $10^{20} \text{ at cm}^{-3}$  erzeugen. In der Implantationsvariante wird die Dotierfolge durch Implantation in die fertigen Nanostege und spätere Diffusion definiert. Der Einsatz von ICP-RIE und CMP ermöglicht die direkte Kontaktierung des Steges von oben. In 5.2 ist die Herstellung dieses Transistors beschrieben.

##### 3. EpiNovaFET

Der I-NovaFET hat flache Substratzuleitungen, die nun einen Engpaß bilden. Die Epitaxie in (niedriger) Kanaldotierung auf hochdotiertem Substrat und

Implantation nur der oberen Schicht (jetzt vor der Steg-Ätzung) verringern diese hohen Zuleitungswiderstände im Substrat. Diese ergänzenden Maßnahmen sind in 5.3 erläutert.

Kapitel 7 beinhaltet die Charakterisierung und den Vergleich der Transistoren, und im Anhang A sind detaillierte Protokolle zu den Prozessen dokumentiert.



# Kapitel 3

## Metall-Oxid-Halbleiter-Feldeffekttransistoren

Die auf Silizium basierende Transistortechnologie dominiert seit über 30 Jahren die Mikroelektronikindustrie. Metall-Oxid-Halbleiter-Feldeffekttransistoren (MOSFETs) sind hierbei die zentralen Bauelemente. Kern eines MOSFETs ist eine Schichtfolge aus Metall (M), Oxid (O) und Halbleiter (S, semiconductor). Am Beispiel eines p-dotierten Halbleiters soll zunächst diese sogenannte MOS-Struktur diskutiert werden.

### 3.1 Die MOS-Struktur

Im idealisierten Fall ist das Bänderschema einer MOS-Struktur ohne angelegte Spannung flach (Abb. 3.1, nach [9]).

Wird am Metall eine Spannung  $U_G$  gegen das geerdete Substrat angelegt, so entsteht ein Potentialverlauf. Wählt man das Potential weit im Halbleiterinneren zum Nullpunkt, so können für das Oberflächenpotential  $\varphi_s$  am Isolator-Halbleiter-Übergang verschiedene Fälle auftreten [10]:

1.  $U_G < 0$ : Es ist dann das Potential  $\varphi_s < 0$ , so daß sich die Bänder des Halbleiters nach oben verbiegen. Es reichern sich in diesem Gebiet Löcher an, die hier Majoritätsladungsträger sind. Dieser Fall wird „Anreicherung“ genannt. Es entsteht eine Raumladungszone positiver Ladung.
2.  $U_G = 0$ : Dieser Fall heißt bei einer idealisierten MOS-Struktur „Flachbandfall“, da  $\varphi_s = 0$  ist und somit keine Bandverbiegung und keine Umverteilung von Ladung stattfindet. (Abb. 3.1)
3.  $U_G > 0$ : Es ist  $\varphi_s > 0$ ; Die Bänder werden nach unten verbogen. Je nach Größe des Potentials in Relation zum Volumenpotential  $\Psi_B$  treten verschiedene Effekte auf. (Bei einem p-dotierten Halbleiter ist  $\Psi_B < 0$ .)
  - (a)  $0 < \varphi_s < -\Psi_B$ : Bewegliche Löcher werden verdrängt, so daß das Potential über die durch die negativ geladenen Akzeptoren des p-Halbleiters do-

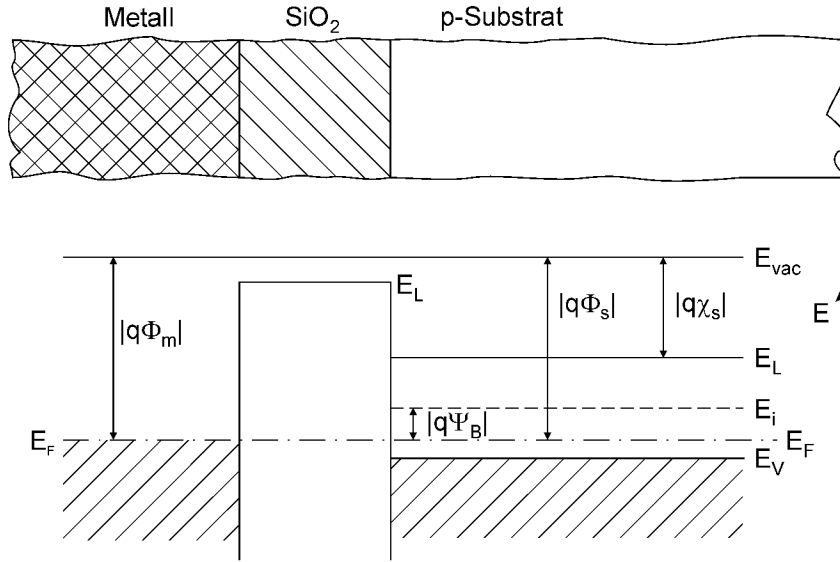


Abb. 3.1: Aufbau und Bänderschema einer idealisierten MOS-Struktur.  $E_F$  bezeichnet die Fermienergie,  $E_L$  und  $E_V$  Leitungs- und Valenzbandkanten,  $E_i$  intrinsisches Fermienergielevel im Halbleiter.  $|q\Phi_m|$  und  $|q\Phi_s|$  sind die Austrittsarbeiten von Metall und Halbleiter,  $|q\chi_s|$  ist die Elektronenaffinität des Halbleiters,  $\Psi_B = (E_F - E_i)/q$  dessen Volumenpotential.

minierte Raumladungszone, die Verarmungszone, abfällt. Die Dicke dieser Verarmungszone vergrößert sich mit zunehmender Spannung. (Abb. 3.2 (a))

- (b)  $\varphi_s = -\Psi_B$ : Das Fermienergielevel fällt an der Oberfläche mit dem intrinsischen Fermienergielevel zusammen. Die Konzentrationen der positiven und negativen freien Ladungsträger sind gleich. Dieser Punkt wird „Einsatzpunkt der schwachen Inversion“ genannt. (Abb. 3.2 (b))
- (c)  $-\Psi_B < \varphi_s < -2\Psi_B$ : Der Abstand zwischen Fermienergie und Leitungsband ist klein genug, daß neue Elektron-Loch-Paare erzeugt und Elektronen im Leitungsband bereitgestellt werden können. Die Löcher werden ins Volumen gezogen und haben keinen Einfluß mehr. Die Konzentration der Minoritätsladungsträger (hier Elektronen) ist größer als die der Majoritätsladungsträger (hier Löcher). Dieser Zustand heißt „Inversion“, die sich dabei ausbildende Minoritätsladungsträgerschicht „Inversionsschicht“.
- (d)  $\varphi_s > -2\Psi_B$ : Die Verarmungszone hat ihre maximale Dicke erreicht. Höhere Spannungen fallen nur noch über die Inversionsschicht ab, da dort eine sehr hohe Elektronendichte erzeugt werden kann. Die Raumladungszone ist somit gegen höhere Felder abgeschirmt und die Bandverbiegung ist nur noch schwach von der angelegten Spannung abhängig. Dieser Zustand heißt „starke Inversion“ (Abb. 3.2 (c), (d)).

Für n-dotierte Halbleiter gelten ähnliche Überlegungen; die Inversionsschicht wird dann nicht aus Elektronen, sondern aus beweglichen Löchern gebildet.

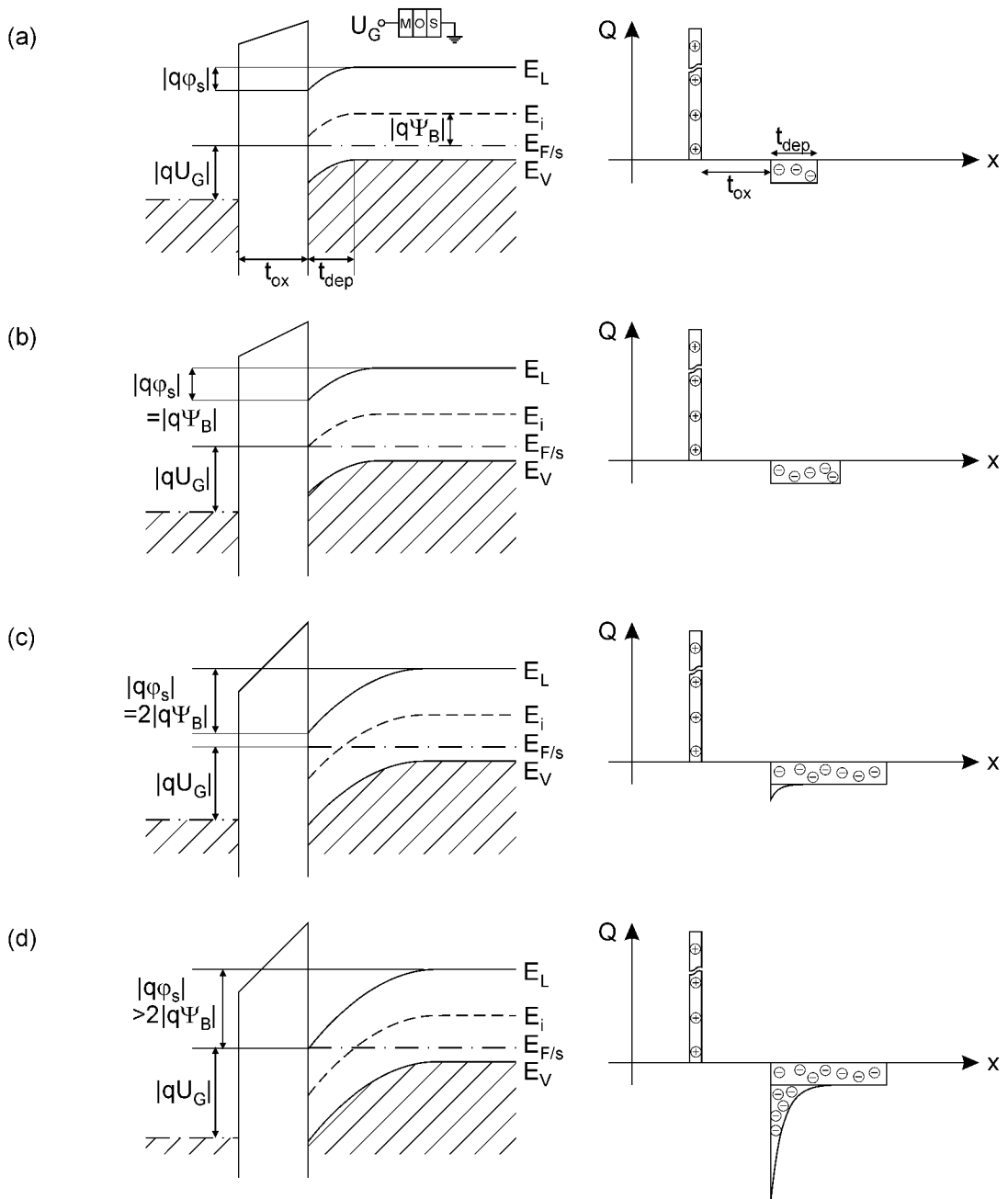


Abb. 3.2: Bandverbiegung und Ladungsverteilung an einer MOS-Struktur bei angelegter positiver Spannung  $U_G$ . (a) Verarmung, (b) Einsatzpunkt der schwachen Inversion, (c) Beginn der starken Inversion, (d) starke Inversion.

Mit einer MOS-Struktur lässt sich also über eine angelegte Spannung die Ladungsträgerdichte an einem Isolator-Halbleiterübergang kontrollieren. Mit Hilfe dieses „Feldeffektes“ kann der Widerstand in einem Halbleiter in einem Kanal nahe der Oberfläche gesteuert werden. Dieses Prinzip wird beim Feldeffekttransistor genutzt.

Für die Betrachtung der MOS-Struktur wurden einige Vereinfachungen gemacht: Beim der „idealisierten MOS-Struktur“ wird davon ausgegangen [11],

- ... daß Metall und Halbleiter die gleichen Austrittsarbeiten haben. Sind diese verschieden, so verbiegen sich die Halbleiterbänder, bis die Fermi-niveaus von Metall und Halbleiter angepaßt sind. Um den Flachbandfall zu erhalten, muß also eine Spannung  $U_G \neq 0$  angelegt werden. Diese Spannung heißt „Flachbandspannung“  $U_{FB}$  und geht als Offset in die bisherigen Überlegungen ein.
- ... daß am Isolator-Halbleiter-Übergang keine Grenzflächenladungen existieren. Werden beim Anlegen einer Spannung unbesetzte Grenzflächenzustände unter das Fermi-niveau abgesenkt, so werden sie umgeladen und binden Ladungen. Die Inversionsladungsträgerdichte kann also nicht mehr so effektiv gesteuert werden. Eine geringe Grenzflächenzustandsdichte ist also angebracht, um ein Bauelement mit einer steuerbaren MOS-Struktur herstellen zu können ([9] S. 197 f).
- ... daß der Isolator ideal ist, also kein Strom durch ihn fließen kann.

### 3.2 Aufbau und Wirkungsweise eines MOSFETs

Der Aufbau eines planaren n-Kanal-MOSFETs ist in Abbildung 3.3 gezeigt.<sup>1</sup> Die Oberfläche des p-dotierten Substrats ist mit einer dünnen  $\text{SiO}_2$ -Schicht, dem Gateoxid, überzogen. Durch Öffnungen im  $\text{SiO}_2$  werden hoch n-dotierte Wannen diffundiert oder implantiert, die Source und Drain bilden. Für die Anschlüsse werden Metallkontakte an Source und Drain angebracht. Eine Metall- oder eine kontaktierte Polysiliziumschicht auf dem Isolator bildet die Gateelektrode.

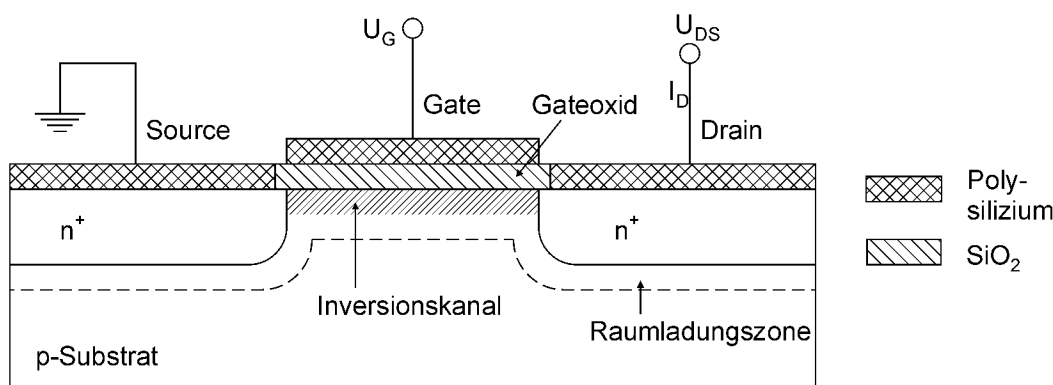


Abb. 3.3: Schematischer Aufbau eines planaren n-Kanal-MOSFETs.

Legt man die idealisierte MOS-Struktur zugrunde, so ist bei der Gatespannung  $U_G = 0$  kein n-leitender Kanal vorhanden und es kann, unabhängig von einer Source-Drain-Spannung  $U_{DS}$ , kein Drainstrom  $I_D$  fließen, weil einer der beiden p-n-Übergänge stets gesperrt ist. Da bei einer realen MOS-Struktur jedoch bereits eine Bandverbiegung auftritt, unterscheidet man nun, ob bei  $U_G = 0$  ein Inversionskanal vorhanden ist oder nicht. Ist dies der Fall, so spricht man von einem „normally on“-

<sup>1</sup>Für einen p-Kanal-MOSFET gilt der analoge Aufbau mit einem n-dotierten Substrat.

andernfalls von einem „normally off“-Transistor ([9] S. 208 f). Im Folgenden wird der normally-off-Transistor betrachtet, der bei  $U_G = 0$  keinen Kanal ausbildet. Dieser Typ ist technologisch besonders relevant, da er im Ruhezustand nur Verlustleistung über Restströme verbraucht.

Die Gatespannung, bei der die starke Inversion beginnt und somit ein Kanal beweglicher Ladungsträger entsteht, wird mit „Einsatzspannung“ oder „Schwellenspannung“  $U_{th}$  des MOSFETs bezeichnet (engl.: *threshold voltage*). Ist  $U_G > U_{th}$ , existiert also ein Inversionskanal zwischen Source und Drain, so kann bei  $U_{DS} \neq 0$  Strom fließen.

### 3.3 Strom-Spannungs-Charakteristik

Die Ausgangskennlinie eines Transistors zeigt den Drainstrom als Funktion der Drainspannung. Bei angelegter Gatespannung  $U_G > U_{th}$  ergeben sich zwei Bereiche der  $I_D$ - $U_{DS}$ -Abhängigkeit ([9] S. 202 f) (vgl. Abb. 3.4):

1. Linearer Bereich. Für kleine Drainspannungen  $U_{DS}$  steigt die Stromstärke  $I_D$  nahezu linear mit der Spannung. Der MOSFET verhält sich wie ein Ohmscher Widerstand. Es gilt [11]:

$$I_D = 2k (U_G - U_{th}) U_{DS} , \quad \text{für } U_{DS} \ll U_G - U_{th} ,$$

wobei  $k$ , der sogenannte „Entwurfsparameter“, eine von den Material- und Geometrieeigenschaften des MOSFETs abhängige Konstante ist.

Der Inversionskanal wird vom Source zum Drain schmaler, da zwischen diesen beiden Elektroden eine Spannung – im idealen Fall linear – abfällt und sich das lokale Potential als Differenz der vom Gate erzeugten, konstanten und dieser ortsabhängigen Komponenten ergibt.

2. Sättigungsbereich.<sup>2</sup> Wenn die Drainspannung  $U_{DSsat} = U_G - U_{th}$  anliegt, hat die Inversionsschicht am Drain die Dicke Null erreicht. Der Drainstrom beträgt  $I_{Dsat} = k \cdot U_{DSsat}^2$  und ändert sich auch für größere Spannungen  $U_{DS}$  nicht mehr. Gründe für dieses Sättigungsverhalten sind die Kanalabschnürung und -verkürzung. Diese haben zur Folge, daß ein Teil des Kanals nicht mehr im Zustand der Inversion, sondern nur noch in dem der Verarmung ist. Aufgrund des hohen Feldes erreichen Elektronen, die aus der Inversionsschicht in diesen Bereich injiziert werden, dennoch die Drainelektrode, so daß der Strom nicht absinkt. Da über den Inversionskanal bis zum Abschnürpunkt immer dieselbe Spannung abfällt, steigt der Strom jedoch auch nicht an.

Das Ausgangskennlinienfeld eines Transistors zeigt die beschriebene Strom-Spannungs-Abhängigkeit mit der Gatespannung als Parameter. Abbildung 3.5 zeigt das Feld eines MOSFETs und veranschaulicht die quadratische Beziehung zwischen Sättigungsstrom und -spannung in Abhängigkeit von der Gatespannung.

<sup>2</sup>Sättigung, engl. *saturation* (Abk. „sat“)

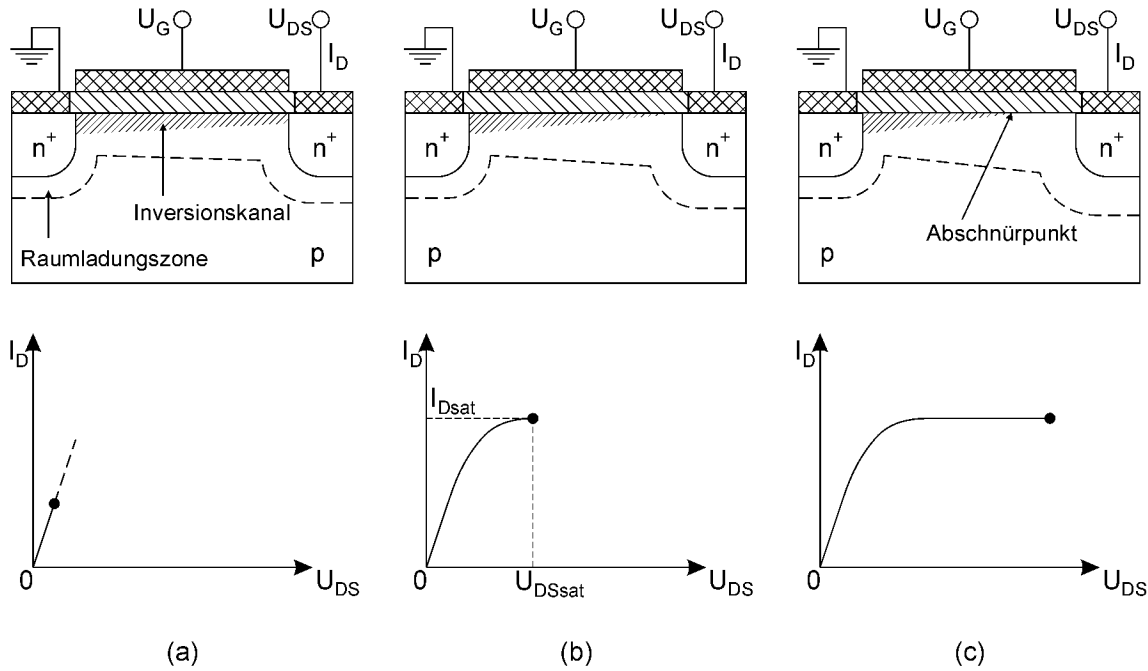


Abb. 3.4: Verschiedene Bereiche der Kennlinie. Es ist  $U_G > U_{th}$ . (a) kleine Drainspannung. Die Abhängigkeit ist annähernd linear. (b) Die Drainspannung hat den Sättigungswert  $U_{Dsat}$ , es fließt der Sättigungsstrom  $I_{Dsat}$ . Der Inversionskanal wird auf der Drainseite abgeschnürt. (c) große Drainspannung. Der Drainstrom befindet sich im Sättigungsbereich, zwischen dem Abschnürpunkt und der Drainelektrode fließt der Drainstrom aufgrund von injizierten Elektronen in einer Sperrschicht ohne Inversionskanal.

Eine wichtige Größe zur Charakterisierung der Gleichstromeigenschaften eines Transistors ist die Steilheit  $g$ . Sie ist definiert als ([9] S. 206)

$$g = \left. \frac{\partial I_D}{\partial U_G} \right|_{U_{DS}=const.} \sim \frac{W}{L \cdot t_{ox}}$$

mit Kanalweite  $W$  und Oxiddicke  $t_{ox}$ , und ist ein Maß für das Steuerverhalten des Transistors. Um die Werte von Transistoren unterschiedlicher Geometrie miteinander vergleichen zu können, wird die Steilheit auf die Kanalweite  $W$  normiert:

$$g_m = \frac{g}{W_G} \sim \frac{1}{L_G \cdot t_{ox}}$$

Hohe Steilheit bedeutet: Eine kleine Änderung der Gatespannung reicht aus, um eine große Änderung des Drainstroms zu bewirken. Da die Betriebsspannungen zukünftiger Schaltkreise immer kleiner werden, um die Leistungsaufnahme zu begrenzen, ist eine hohe Steilheit angestrebt. Diese kann durch Verkleinerung der Kanallänge  $L$  und der Oxiddicke  $t_{ox}$  erreicht werden.

Die Transferkennlinie eines Transistors beschreibt die Steuerwirkung. Der Drainstrom wird als Funktion der Gatespannung aufgetragen. Bei einem idealen MOSFET sind die Transferkennlinien unabhängig von der angelegten Source-Drain-Spannung.

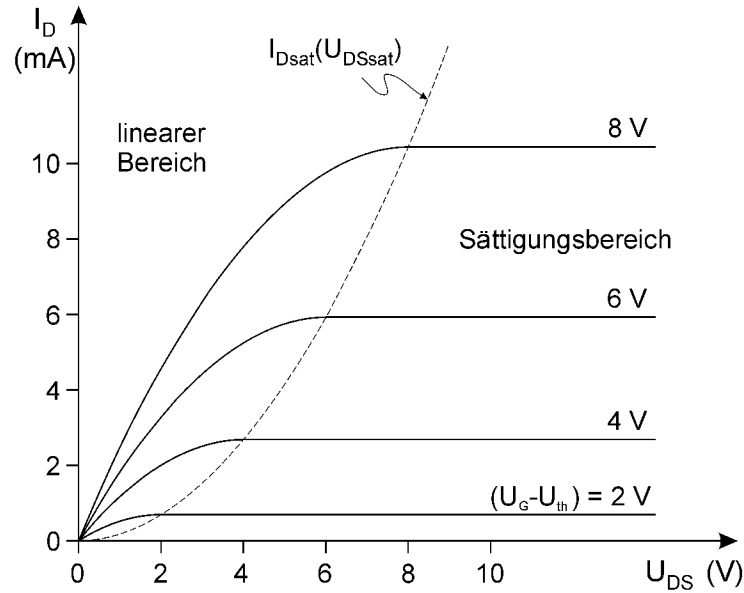


Abb. 3.5: Ausgangskennlinienfeld eines MOSFETs in schematischer Darstellung. Der Drainstrom  $I_D$  ist gegen die Drainspannung  $U_{DS}$  aufgetragen; die Differenz zwischen Gate-spannung und Schwellenspannung ( $U_G - U_{th}$ ) erzeugt bewegliche Ladungsträger und ist Parameter.

Wie in Abbildung 3.6 a ersichtlich, kann die Kennlinie in drei Bereiche unterteilt werden:

1. Ausgeschalteter Bereich. Der Transistor ist ausgeschaltet, wenn kein Kanal vorhanden ist. Dies ist bei Gatespannungen der Fall, bei denen das Silizium im Kanalgebiet in Verarmung ist. Die Leckströme, die zu  $I_{off}$  zusammengefaßt werden, sollten möglichst klein sein.
2. Unterschwellenspannungsbereich. Die MOS-Struktur ist in schwacher Inversion. Es fließt ein Drainstrom, der exponentiell mit der angelegten Gatespannung ansteigt. Für  $U_{DS} > 3kT/q$  ist  $I_D$  von  $U_{DS}$  unabhängig [12]. Die Steigung in diesem Bereich ist ein wichtiges Maß für die Schalteigenschaften des Transistors. Diese „Unterschwellenspannungssteigung“  $S$  ist die reziproke Steigung im logarithmischen Auftrag von  $I_D$  und hat die Einheit  $mV/dec$ .
3. Überschwellenspannungsbereich. Die MOS-Struktur ist in starker Inversion. Die Einsatzspannung der starken Inversion ist  $U_G = U_{th}$ . Wegen der quadratischen Abhängigkeit des Sättigungsdrainstroms von  $U_G - U_{th}$ , kann die Schwellenspannung grafisch bestimmt werden (Abb. 3.6 b).

Die wichtigsten technologischen Kenngrößen sind in Abbildung 3.6 (c) veranschaulicht: Die Gatelänge  $L$  ist der Abstand zwischen Source (S) und Drain (D), zwei dem Substrat gegendotierte Wannen,  $W$  die Gateweite (in der Abbildung die Ausdehnung in  $z$ -Richtung). Die Gateoxiddicke wird mit  $t_{ox}$  bezeichnet, die Dicke der Verarmungszone unter dem Gate mit  $t_{dep}$  (von *depletion*, engl. für Verarmung).  $\varepsilon_{ox}$  und  $\varepsilon_{si}$  sind die Dielektrizitätskonstanten für Oxid bzw. Silizium.

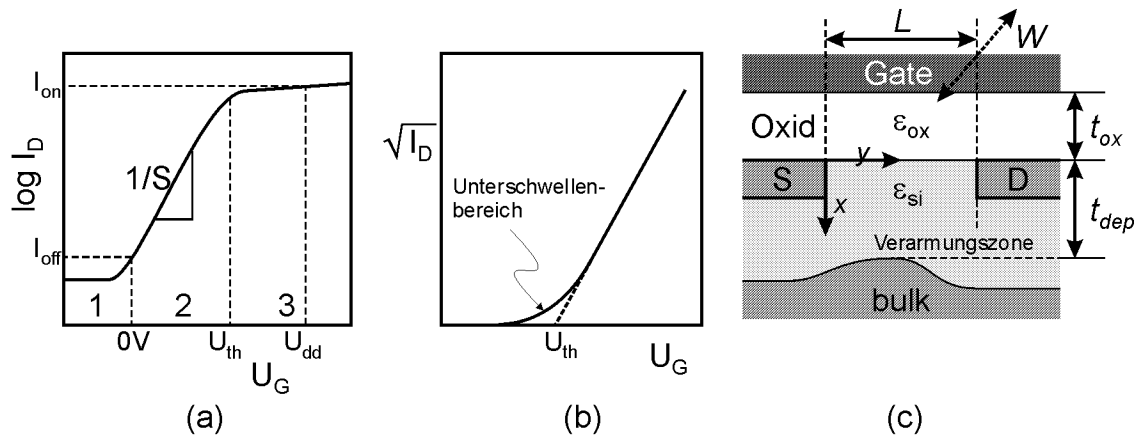


Abb. 3.6: (a) Schematische Darstellung der Transferkennlinie eines MOSFETs mit den drei Bereichen: Restströme, Unterschwellenspannungsbereich und Überschwellenspannungsbereich. Es sind die wichtigsten elektrischen Kenngrößen eingezeichnet.  $U_G = U_{DS} = U_{DD}$  ist der Arbeitspunkt im eingeschalteten Zustand,  $U_G = 0$  und  $U_{DS} = U_{DD}$  der bei ausgeschaltetem Zustand. (b) Grafische Bestimmung der Schwellenspannung  $U_{th}$ . (c) Technologische Kenngrößen eines MOSFETs.

Motor der Weiterentwicklung von MOSFETs ist die Ökonomie. Kleinere Transistoren auf größeren Wafern bedeuten billigere Bauelemente mit besseren Eigenschaften. Derzeit an der 100 nm-Grenze (DRAM half pitch), nähert sich die Technologie zwar noch nicht fundamentalen physikalischen Grenzen [2], während die Skalierung bislang jedoch einfachen Regeln folgte (Skalierungsmethode nach Dennard *et al.* [13]), die mit der Zeit nur leicht angepaßt werden mußten, können in wenigen Jahren wohl nur neuartige Bauelementkonzepte die geometrischen und elektrischen Anforderungen erfüllen. Im Folgenden werden die Skalierungsregeln vorgestellt und die Grenzen diskutiert.

### 3.4 MOSFET-Skalierung

Um die guten Schalteigenschaften der Transistoren zu bewahren, schlugen Dennard *et al.* [13] 1974 eine einfache Skalierungsregel vor (s. Abb. 3.7): Wenn sämtliche geometrische Größen eines planaren MOSFETs mit dem Faktor  $\alpha$  verkleinert werden, ebenso die angelegten Spannungen, dabei gleichzeitig die Substratdotierung um den Faktor  $\alpha$  vergrößert wird, so ändert sich das elektrische Feld im Bauelement nicht. Die Drainströme werden dabei ebenfalls um den Faktor  $\alpha$  verkleinert, behalten aber ihre Charakteristik bei. Die Schaltzeiten verkleinern sich ebenfalls um den Faktor  $\alpha$  und ermöglichen so höhere Schaltfrequenzen.

Grenzen dieser Skalierung treten für Submikrometer-MOSFETs auf, zum einen, wenn nicht alle Parameter gleichmäßig skaliert werden (können), zum anderen, wenn zusätzliche Effekte auftreten, die bei größeren Dimensionen vernachlässigt werden konnten. Nicht im selben Maß skaliert wie die anderen Größen wurde in der Vergangenheit z. B. die Betriebsspannung  $U_{dd}$ . Ein Grund dafür ist die Kompatibilität der

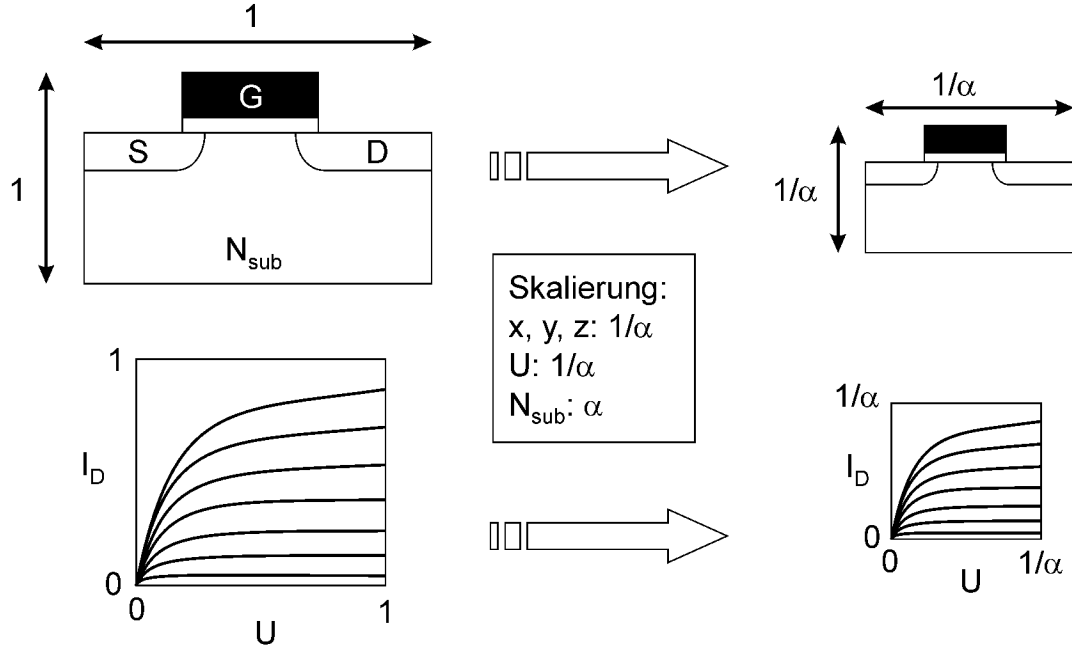


Abb. 3.7: Skalierungsregeln nach Dennard et al.: Werden die Dimensionen eines MOSFETs sowie die angelegte Spannung um den gleichen Faktor  $1/\alpha$  skaliert, dabei die Substratdotierung um den Faktor  $\alpha$  erhöht, so bleiben die Felder im Transistor gleich und die Charakteristik des Transistors ändert sich nicht.

verschiedenen Bauelementgenerationen, ein anderer, daß die Schwellenspannung  $U_{th}$  nicht skaliert werden kann. Diese wiederum ist hauptsächlich an die Unterschwellenspannungssteigung  $S$  gekoppelt: Bleibt sie unverändert, so sind geringere Werte von  $U_{th}$  mit höheren Unterschwellenströmen und größerem  $I_{off}$  verbunden. Die Verringerung von  $U_{th}$  um 0,1 V bedeutet dann einen Anstieg des Leckstroms bei  $I_{off}$  um den Faktor zehn [14]. Daß  $S$  nicht oder nur wenig verbessert werden kann, ist in ihrer starken Abhängigkeit von der thermischen Diffusion der Elektronen begründet [15]:

$$\begin{aligned}
 S &= \left( \frac{kT}{q} \ln 10 \right) \left( 1 + \frac{\varepsilon_{si} t_{ox}}{\varepsilon_{ox} t_{dep}} \right) \\
 &= 60 \text{ mV} \left( 1 + \frac{\varepsilon_{si} t_{ox}}{\varepsilon_{ox} t_{dep}} \right) \quad (\text{für RT}) \\
 &= 60 \text{ mV} \cdot m
 \end{aligned}$$

$S$  ist also von der Betriebsspannung und der Kanallänge unabhängig und hat bei Raumtemperatur eine Untergrenze bei 60 mV. Der Koeffizient  $m$  ist von Materialkonstanten und der Ausdehnung der Raumladungszone  $t_{dep}$  abhängig und liegt üblicherweise bei 1,2–1,5 [14]. Für  $t_{dep}$  wiederum gilt [9]:

$$t_{dep} \approx \sqrt{\frac{4 \varepsilon_{si} kT \ln(N_{sub}/n_i)}{q^2 N_{sub}}}$$

mit  $n_i$  der intrinsischen Ladungsträgerkonzentration von Silizium und  $q$  der Elementarladung. Die Unterschwellenspannungssteigung  $S$  kann somit nicht skaliert werden

und könnte nur über das Absenken der Betriebstemperatur unter den Grenzwert von 60 mV/dec gedrückt werden. Der Skalierungsparameter  $t_{dep}$  kann für einen „normalen“ MOSFET nur über die Erhöhung der Kanaldotierung verkleinert werden, was jedoch zu einer Erhöhung von  $U_{th}$ , zu größeren Feldern am Oxid und zu einer Verringerung der Ladungsträgerbeweglichkeit führt [16, 17].

Eine physikalische Grenze wird bald bei der Verwendung von  $\text{SiO}_2$  als Gateoxid erreicht: In einer MOS-Struktur sind mindestens vier Atomlagen  $\text{SiO}_2$  ( $= 0,7 \text{ nm}$ ) notwendig, damit sich eine Bandlücke ausbilden kann; auf Grund der Oberflächenrauigkeit heutiger Herstellungstechniken sind  $1,2 \text{ nm}$  die Grenze [18]. Diese Gateoxiddicken werden bereits für 2006 oder 2010 (je nach Anwendung) benötigt [19]. Ein weiteres Problem dabei sind Tunnelströme, die bei Schichtdicken unterhalb von  $3 \text{ nm}$  signifikant zunehmen. Ein Ausweg sind alternative Gateoxide, sogenannte „high- $k$ -Dielektrika“<sup>3</sup>. Ziel der dünneren Gateoxide ist eine Vergrößerung der Gatekapazität  $C_{ox}$ . Dies kann jedoch auch durch andere Dielektrika erzielt werden, sofern die Dielektrizitätskonstante  $k > 3,9$  ist – dies ist der Wert für  $\text{SiO}_2$ . Es gilt dann:

$$C_{\text{high-}k} = \frac{k_{\text{high-}k} \varepsilon_0}{t_{\text{high-}k}}$$

Damit ergibt sich für die äquivalente Oxiddicke  $EOT$ <sup>4</sup>:

$$EOT = \frac{k_{ox}}{k_{\text{high-}k}} t_{\text{high-}k}$$

Das heißt, bei der Verwendung von alternativen Gateoxiden kann das Oxid um den Faktor  $k_{\text{high-}k}/k_{ox}$  dicker sein als  $\text{SiO}_2$ , so daß Tunneln unterdrückt wird. Damit hofft man die Skalierungsregeln der kommenden Generationen erfüllen zu können, auch wenn sich mit der Auswahl und den Eigenschaften der alternativen Gateoxide weitere Problemfelder eröffnen.

### 3.5 Kurzkanaleffekte

Die Herleitung der Transistoreigenschaften galt für bestimmte Bedingungen [12]. Diese sind eine ideale MOS-Struktur, eine konstante Ladungsträgerbeweglichkeit  $\mu$  im Inversionskanal, homogene Kanaldotierung, vernachlässigbare Leckströme und ein im Vergleich zur lateralen Feldstärke großes transversales Feld im Kanalgebiet. Diese Voraussetzungen sind bei relativ großen Abmessungen (Kanallängen im  $\mu\text{m}$ -Bereich und Oxiddicken  $> 10 \text{ nm}$ ) in der Regel erfüllt, so daß man von „Langkanalverhalten“ spricht. Langkanalverhalten entspricht dem bisher beschriebenen Transistorverhalten.

Abweichungen vom Langkanalverhalten bezeichnet man allgemein als „Kurzkanalverhalten“, da sie insbesondere bei kleindimensionierten Bauelementen zu finden sind. Im Folgenden werden die wichtigsten Kurzkanaleffekte beschrieben. Abbildung 3.12 verdeutlicht abschließend ihre Auswirkungen auf die Kennlinienfelder.

<sup>3</sup>Teilweise wird statt  $k$  auch  $\kappa$  verwendet, im Deutschen meist  $\varepsilon_r$ .

<sup>4</sup>*Equivalent oxide thickness*, die entsprechende Dicke  $\text{SiO}_2$ , die zu einer gleichen Gatekapazität führt.

### 3.5.1 $U_{th}$ -Roll-Off

Mit dem Begriff „ $U_{th}$ -Roll-Off“ wird die Abhängigkeit der Schwellenspannung  $U_{th}$  von der Kanallänge  $L$  bezeichnet. Wie in 3.1 beschrieben, ist bei  $U_G = U_{th}$  die Raumladungszone unter dem Gate maximal, und bei größeren Gatespannungen bildet sich ein Inversionskanal aus. Die Größe dieser Zone, die durch die Gatespannung verarmt werden muß, bestimmt also den Wert von  $U_{th}$ . In der Abbildung 3.8 a wäre die vollständige Raumladungszone das Rechteck ABEF. An den Rändern jedoch wird ein Teil dieses Kanalgebietes durch die p-n-Übergänge an Source und Drain, welche ebenfalls Raumladungszone ausbilden, verarmt. Es wird also effektiv nur noch das Trapez ACDF durch das Gate kontrolliert.<sup>5</sup> Aus diesem Grund setzt die Inversion bei kleineren Schwellenspannungen ein. Ist die Gatelänge in der Größenordnung der Verarmungszonen, so ist der relative Anteil der nicht vom Gate kontrollierten Gebiete signifikant größer als bei langen Kanälen (Abb. 3.8 b). Die Verminderung der Schwellenspannung mit abnehmender Gatelänge ist in Abbildung 3.8 c dargestellt.

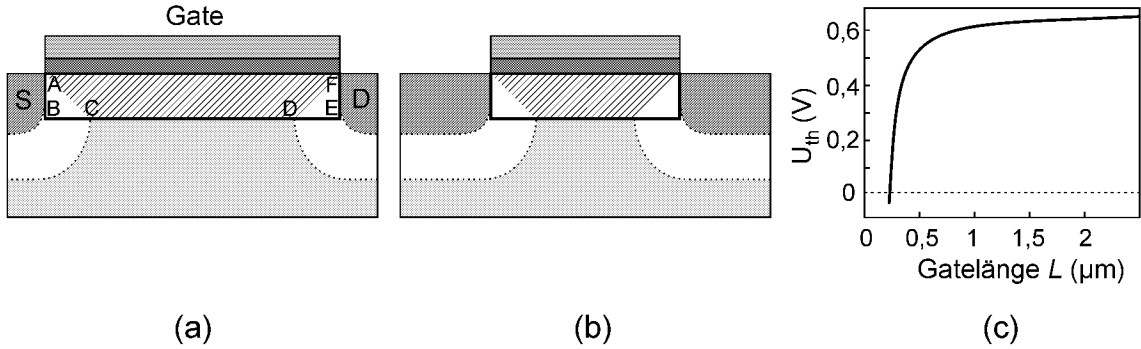


Abb. 3.8:  $U_{th}$ -Roll-Off: An den Rändern des Kanalgebietes sind Bereiche der Raumladungszone nicht durch die Gatespannung kontrolliert (Dreiecke ABC und DEF in (a)). Dieser Anteil wächst relativ zum Gesamtgebiet, wenn die Gatelänge schrumpft (b). Kommt  $L$  in die Größenordnung der Raumladungszone, so ist der Effekt dominierend. In (c) ist die Abhängigkeit beispielhaft dargestellt.

### 3.5.2 Drain-Induced Barrier Lowering

Dieser Effekt beschreibt die Abhängigkeit der Schwellenspannung  $U_{th}$  von der Source-Drain-Spannung  $U_{DS}$ . Für die Dicke der Verarmungsgebiete an Source und Drain gilt [9]:

$$t_S = \sqrt{\frac{2\varepsilon_{si}}{qN_{sub}}(2\Psi_B + U_{SB})}$$

$$t_D = \sqrt{\frac{2\varepsilon_{si}}{qN_{sub}}(2\Psi_B + U_{SB} + U_{DS})}$$

<sup>5</sup>Dieses Modell, bei dem die Größe der Verarmungszone betrachtet wird, die tatsächlich vom Gate kontrolliert wird, heißt „Trapez-Ansatz“ oder „charge-sharing“.

mit  $U_{SB}$  einer möglichen Substratvorspannung. Die Raumladungszone dehnt sich also am Drain mit ansteigender Source-Drain-Spannung aus, da dieser p-n-Übergang in Sperrichtung geschaltet ist. Dies hat zur Folge, daß ein weiterer, von  $U_{DS}$  abhängiger Anteil der Raumladungszone unter dem Gate nicht von der Gatespannung kontrolliert wird (Abb. 3.9 a). Dieser Effekt kann auch mit Hilfe der Potentialbarriere für die Ladungsträger zwischen Source und Drain am Kanal beschrieben werden. Bei kurzen Kanälen sinkt diese Barriere durch die angelegte Source-Drain-Spannung und die Inversion setzt bei geringeren Spannungen  $U_{th}$  ein (Abb. 3.9 b). Dieser Ansatz ist der Grund für die Bezeichnung „Drain-Induced Barrier Lowering“ (DIBL). Eine Folge des DIBL sind höhere Leckströme  $I_{off}$ .

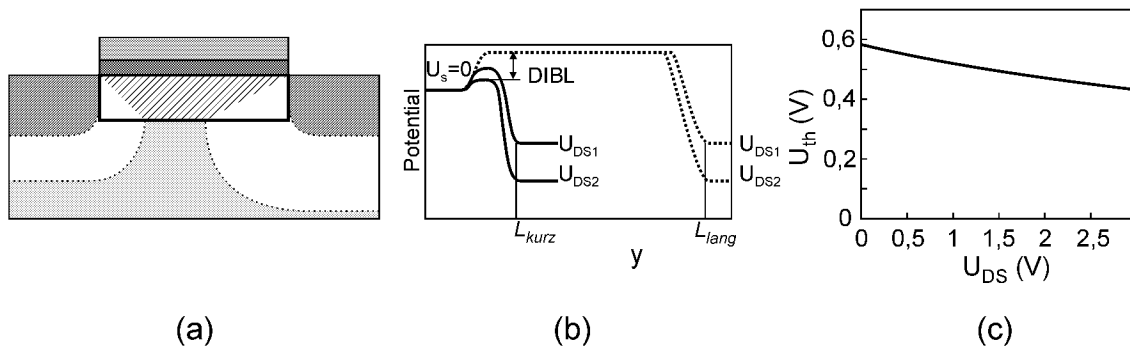


Abb. 3.9: *Drain-Induced Barrier Lowering: (a) Bei angelegter Source-Drain-Spannung dehnt sich die Raumladungszone am Draingebiet aus. (b) Eine alternative Beschreibung betrachtet die Potentialbarriere zwischen Source und Drain, die bei kurzen Kanälen durch  $U_{DS}$  abgesenkt wird. (c) Beispielhafter Verlauf der Abhängigkeit.*

### 3.5.3 Punch-Through

Wird die Raumladungszone am Drain so groß, daß  $t_S + t_D = L$  ist, berühren sich also die Raumladungszone von Source und Drain (Abb. 3.10), so fließt ein zusätzlicher, nicht vom Gate kontrollierter Strom auch unterhalb der Schwellenspannung [12]. Der Strom ist proportional zu  $U_{DS}^2$  und kann durch Erhöhung der Substratdopierung verhindert werden, da sich dann die Raumladungszone verkleinern. Bei Kurzkanal-MOSFETs bestimmt der *Punch-Through* meist die maximale Source-Drain-Spannung.

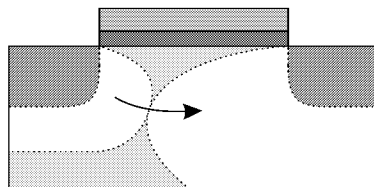


Abb. 3.10: *Punch-Through: Berühren sich die Raumladungszone von Source und Drain, so fließt ein nicht vom Gate kontrollierter Strom.*

### 3.5.4 Kanallängenmodulation

Als Kanallängenmodulation wird die Kanallängenverkürzung im Sättigungsbereich bezeichnet. Wenn für  $U_{DS} > U_G - U_{th} = U_{DSsat}$  der Inversionskanal am Drain abschnürt (siehe Abb. 3.4), verkleinert sich die effektive Kanallänge  $L_{eff}$ , da die gesamte Spannung nur über dem verkürzten Inversionskanal abfällt (Abb. 3.11 a). Für die Verkürzung gilt [20]:

$$\Delta L = \sqrt{\frac{2 \varepsilon_{si}}{q N_{sub}} (U_{DS} - U_{DSsat})}$$

Diese, von der Kanallänge unabhängige Änderung ist ebenfalls für kleine Kanallängen signifikanter als für große. Die Folge ist eine größere laterale Feldstärke, so daß mehr Ladungsträger aus dem Source gezogen werden. In den Kennlinien äußert sich die Kanallängenmodulation also durch fehlende Sättigung: Der Sättigungsstrom steigt linear mit der Drainspannung. Die Sättigungsstromgleichung  $I_{Dsat} = k U_{DSsat}^2$  von Langkanal-MOSFETs (siehe Abb. 3.3) verändert sich dann zu

$$I_{Dsat} = k U_{DSsat}^2 \cdot (1 + \lambda U_{DS})$$

Der Kanallängenmodulationsfaktor  $\lambda$  der Parametrisierung kann grafisch aus dem Ausgangskennlinienfeld bestimmt werden (Abb. 3.11 b).

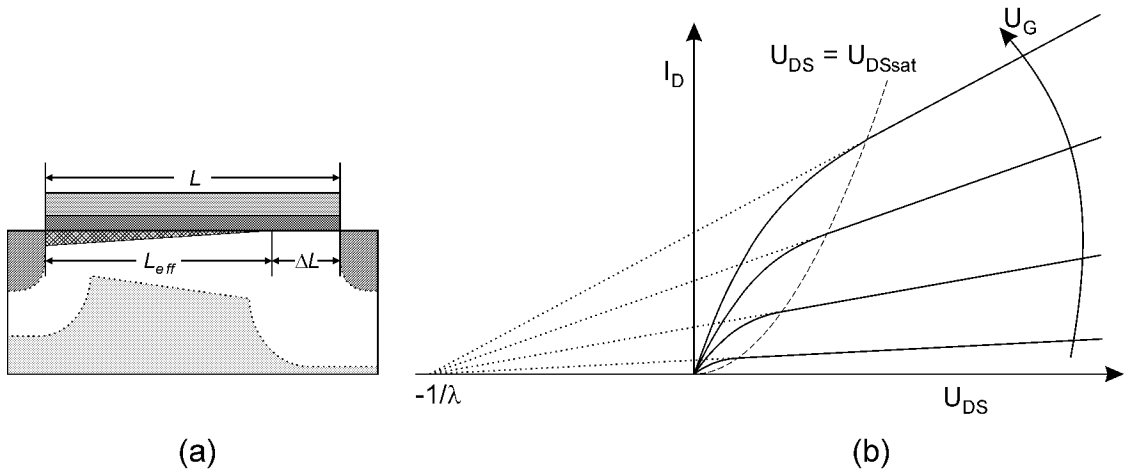


Abb. 3.11: Kanallängenmodulation: (a) Die effektive Kanallänge bei Sättigung ist nicht mehr der geometrische Abstand zwischen Source und Drain, da der Inversionskanal abschnürt. (b) Der lineare,  $U_{DS}$ -abhängige Anstieg im Sättigungsbereich läßt sich durch  $\lambda$  beschreiben und geometrisch aus den Ausgangskennlinien bestimmen.

### 3.5.5 Unterschwellströme

Nach dem vereinfachten Modell fließt im Unterschwellbereich ( $U_G < U_{th}$ ) überhaupt kein Drainstrom. Tatsächlich aber klingt  $I_D$  exponentiell ab, wenn  $U_G$  unter  $U_{th}$  sinkt:

$$I_D \propto \frac{1}{L_{eff}} e^{q U_G / k T}$$

Bei schwacher Inversion fließt also ein Reststrom. Während die Unterschwellströme bei Langkanaltransistoren für  $U_{DS} > 3kT/q$  von der Source-Drain-Spannung unabhängig sind [12], steigen sie für kleine Gatelängen stark mit  $U_{DS}$ . Grund ist die Proportionalität zu  $1/L_{eff}$ .  $L_{eff}$  ist hier die um die Raumladungszonen an Source und Drain reduzierte Gatelänge und ist – wie beim DIBL beschrieben – für kleine  $L$  stark von  $U_{DS}$  abhängig.

Restströme unterhalb der Einsatzspannung können die statische Stromaufnahme von integrierten Schaltungen wesentlich beeinflussen, denn diese Ströme fließen bei jedem inaktiv geschalteten Transistor. Auch werden Haltezeiten dynamischer Schaltungen, bei denen digitale Informationen auf Kapazitäten gespeichert werden, verkürzt.

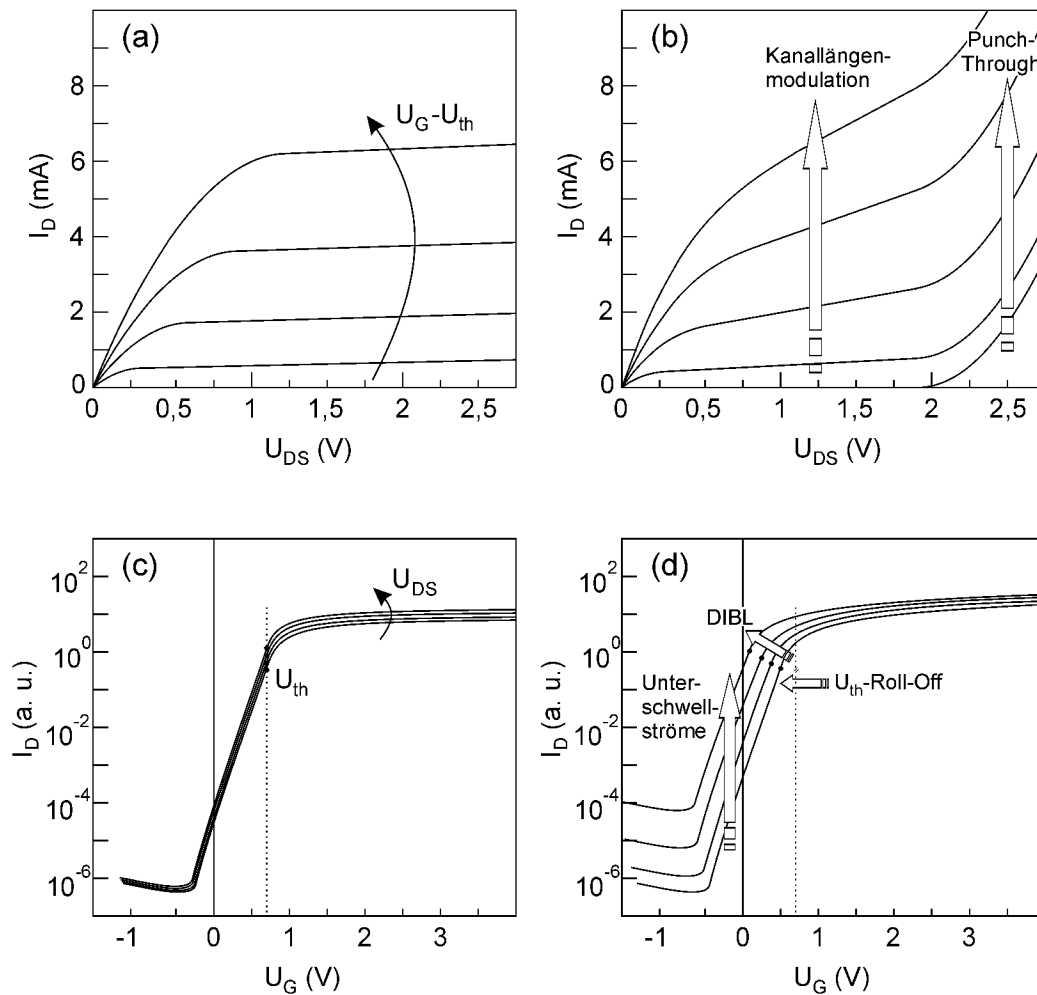


Abb. 3.12: Auswirkungen der Kurzkanaleffekte auf die Kennlinien eines MOSFETs (schematisch). Ausgangskennlinienfelder beim Langkanal- und Kurzkanal-Transistor (a)+(b). Transferkennlinienfelder beim Langkanal- und Kurzkanal-Transistor (c)+(d).

### 3.6 Neue MOSFET-Konzepte

Für die Chipgenerationen der vergangenen Jahre mußten die Designs der MOSFETs abweichend von der Skalierungsregel verändert werden, damit das beschriebene Kurzkanalverhalten unterdrückt werden konnte. Zunächst wurde das elektrische Feld mit einem anderen Skalierfaktor entworfen, dann Transistordimensionen und deren Zuleitungen entkoppelt [21]. Derzeitige MOSFETs unterdrücken Kurzkanaleffekte mit Dotierprofilen im Kanalbereich („channel design“, Abb. 3.13). Dadurch werden die Verarmungszonen an Source und Drain verkleinert, wodurch das vom Gate kontrollierte Gebiet maximiert wird. Für zukünftige MOSFET-Generationen jedoch ist das größte Problem die Tiefe der Verarmungszone  $t_{dep}$ , die nicht wie die anderen Parameter skaliert werden kann und dadurch Hauptursache für den  $U_{th}$ -Roll-Off ist. Neue MOSFET-Konzepte setzen deshalb auf die Begrenzung des Kanalgebietes durch die Verwendung dünner Siliziumschichten, die dann vollständig verarmen. Hier wiederum kann man unterscheiden zwischen planaren MOSFETs, die aus einer Siliziumschicht auf Oxid hergestellt werden – sogenannten SOI-MOSFETs<sup>6</sup> – und vertikalen MOSFETs, bei denen die Siliziumschicht als Steg senkrecht zur Waferoberfläche angeordnet ist.

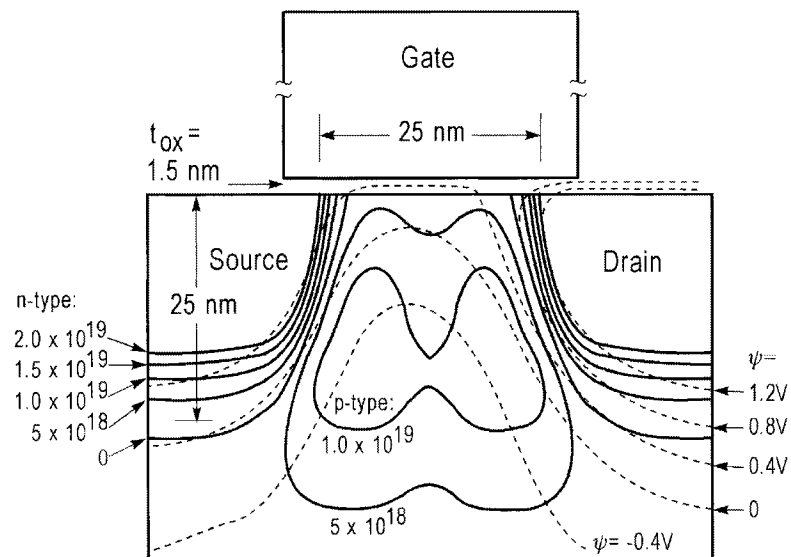


Abb. 3.13: Der  $U_{th}$ -Roll-Off kann durch geeignetes Doping im Kanalbereich vermindert werden. Diese hier gezeigte Form wird „Halo-Doping“ genannt (aus [16]).

#### 3.6.1 SOI-MOSFETs

Bei einem SOI-MOSFET ist die Tiefe der Verarmungszone  $t_{dep}$  durch die Dicke der Siliziumschicht  $t_{si}$  begrenzt. (Abb. 3.14 a). Ist die Schicht so dünn, daß das gesamte Kanalgebiet verarmt, vermindern sich die Kurzkanaleffekte (Abb. 3.14 b). Kleine-

<sup>6</sup>SOI, *silicon on insulator*, Silizium auf Isolator.

re Gatelängen können somit erreicht werden. Nachteilig sind die extrem dünnen Siliziumschichten (5–10 nm) sowie die relativ hohen Kosten für SOI-Wafer.

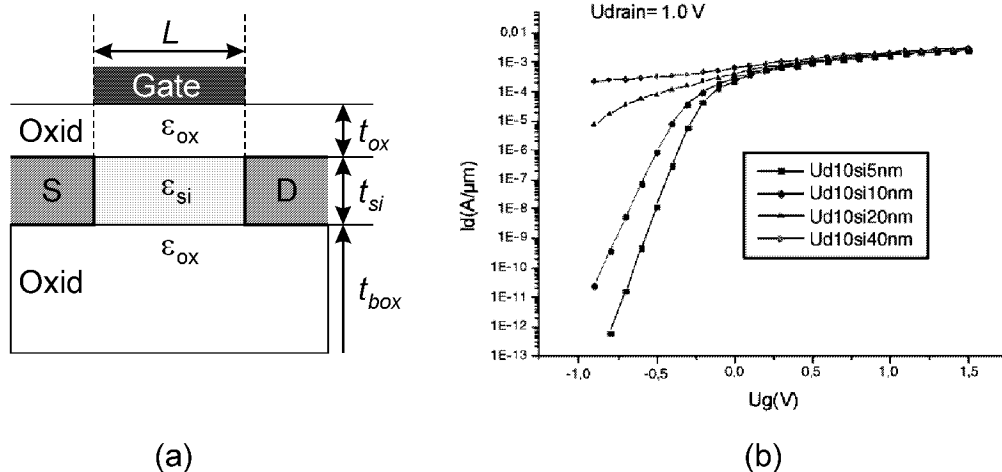


Abb. 3.14: (a) Bei einem SOI-MOSFET ist das Kanalgebiet vom Substrat durch eine Isolierschicht getrennt. (b) Die Folge sind verminderte Kurzkanaleffekte: Je dünner die Oxidschicht ist, desto größer sind Schwellenspannung und Off-Ströme und desto kleiner die Unterschwellenspannungssteigung (Simulationen für  $L = 50 \text{ nm}$ ,  $t_{si} = 40 \text{ nm}$  und  $t_{ox} = 2 \text{ nm}$ , aus [2]).

### 3.6.2 Double-Gate-MOSFETs

Weitere Vorteile können erzielt werden, wenn das Kanalgebiet von beiden Seiten jeweils durch ein Gate gesteuert wird; man spricht dann von „Double-Gate“ (DG). Es gibt drei verschiedene Arten von DG-MOSFETs, die sich in der Orientierung ihrer Kanäle unterscheiden:

Planarer DG	Kanal und Strom parallel zur Waferoberfläche	Abb. 3.15 a
FinFET	Kanal senkrecht und Strom parallel zur Wafer-OF	Abb. 3.15 b
Vertikaler DG	Kanal und Strom senkrecht zur Waferoberfläche	Abb. 3.15 c

Der planare DG-MOSFET entspricht einem SOI-MOSFET, bei dem auf der Unterseite des Kanals ein zweites Gate angebracht ist. Beim FinFET wird der Steg aus einer SOI-Schicht herausgeätzt. Source, Kanal und Drain sind lateral angeordnet, so daß der Strom im Steg, aber parallel zur Waferoberfläche fließt. Beim vertikalen DG sind Source, Kanal und Drain übereinander angeordnet in einem Steg, der aus dem Substrat geätzt wurde. Das Kanalgebiet ist also über die Dicke der mittleren Schicht definiert, nicht durch Lithographie, wie bei den anderen Versionen. Der Strom fließt vertikal zur Waferoberfläche.

Auf Grund des symmetrischen Aufbaus können das elektrische Feld am Drain noch besser unterdrückt und bessere Transistoreigenschaften erreicht werden. Die Siliziumschicht muß nicht so dünn sein wie bei SOI-MOSFETs, da an jedem Gate nur die halbe Dicke der Siliziumschicht verarmt werden muß:  $t_{dep} = t_{si}/2$ . Dabei fließt

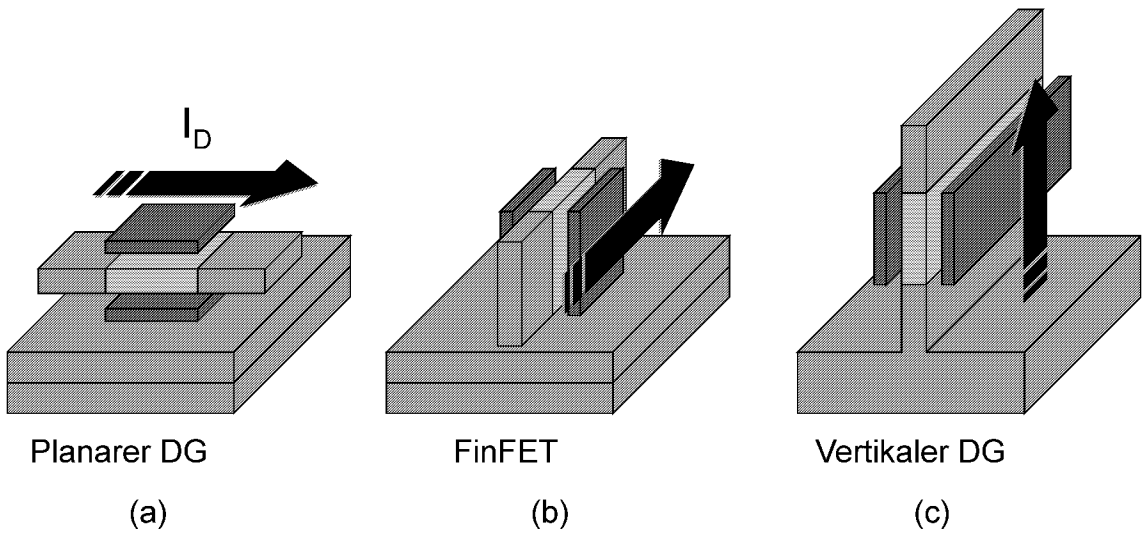


Abb. 3.15: Es gibt drei verschiedene Arten von Double-Gate-MOSFETs: (a) Planarer DG, (b) FinFET und (c) vertikaler DG. Die ersten beiden Variationen werden aus SOI-Wafern hergestellt, und der Stromfluß ist bei beiden parallel zur Waferoberfläche, beim planaren DG Flach in der Schicht, beim FinFET im Steg. Der vertikale DG ist aus Substrat geätzt, und der Strom fließt senkrecht im Steg.

durch die Parallelschaltung der beiden Kanäle der doppelte Strom [2]. Auch können Dotierprofile im Kanalgebiet vermieden werden, die einen zusätzlichen Beitrag zum elektrischen Feld an der Oberfläche leisten [21]. Undotierte Kanalgebiete verbessern sogar den Ladungstransport: Coulombstreuung wird reduziert bei weniger ionisierten Störstellen im undotierten/gering dotierten Kanal, und geringere elektrische Felder an der Oberfläche verringern Oberflächenstreuungen [21].

Eine Besonderheit von Double-Gate-MOSFETs ist die Möglichkeit, bei hinreichend dünnem Silizium eine Volumeninversion erreichen zu können. Simulationen von Schulz von vertikalen DG-MOSFETs [4] haben gezeigt, daß diese Volumeninversion mit den Standardmodellen „Drift-Diffusion“<sup>7</sup> und „Hydrodynamik“<sup>8</sup> nicht erfaßt werden können (Abb. 3.16, links). Quantendynamische Simulationen jedoch zeigen eine maximale Elektronenkonzentration in der Stegmitte für  $t_{si} \leq 10$  nm.

Vollständige Verarmung bei vertikalen DG-MOSFETs mit einer Kanallänge von 100 nm und einer Substratdotierung von  $N_{sub} = 10^{18} \text{ cm}^{-3}$  wurde bei Simulationen für  $t_{si} < 70$  nm erreicht, bei skalierten Transistoren mit  $L = 50$  nm und  $N_{sub} = 5 \cdot 10^{18} \text{ cm}^{-3}$  für  $t_{si} < 30$  nm. Mit diesen vollständig verarmten DG-MOSFETs können Kurzkanaleigenschaften jedoch bereits verbessert werden (siehe Abb. 3.17).

<sup>7</sup>Beim Drift-Diffusions-Modell setzt sich die Ladungsstromdichte aus dem Driftstrom und dem Diffusionsstrom zusammen.

<sup>8</sup>Beim Hydrodynamik-Modell werden zusätzlich sog. „heiße Elektronen“ berücksichtigt durch Einführung einer teilchenabhängigen Ladungsträgertemperatur.

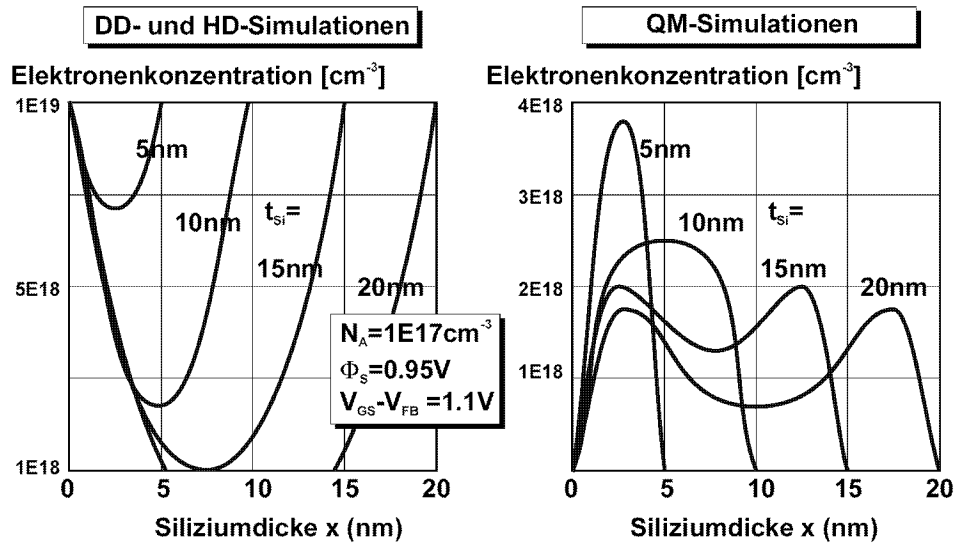


Abb. 3.16: Simulationen der Elektronenverteilung im DG-MOSFET mit dünnen Siliziumstegen. Volumeninversion ist nur bei quantenmechanischer Simulation bei Schichtdicken von höchstens 10 nm zu beobachten. (Abb. aus [4]).

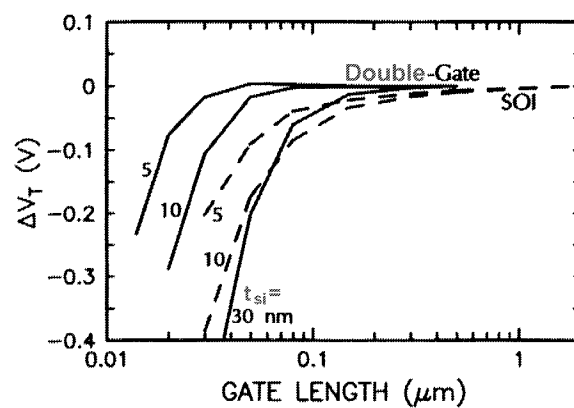


Abb. 3.17: Das Einbrechen der Schwellenspannung  $U_{\text{th}}$  bei Verringerung der Kanallänge tritt bei DG-MOSFETs deutlich später auf als bei SOI-MOSFETs. (Simulation, nach [14].)

# Kapitel 4

## Verwendete Technologie

In diesem Kapitel werden die wichtigsten Technologien beschrieben, die für die Herstellung der Transistoren verwendet wurden. Die genauen Prozeßparameter sind im Kapitel 5 angegeben, in dem die Prozeßfolge der drei Varianten des NovaFETs erläutert sind.

### 4.1 Technologien zur Strukturierung

Die Strukturierung gliedert sich in zwei Bereiche: den der Strukturdefinition, die hier mittels Elektronenstrahlschreiben oder optischer Lithographie erfolgt, und den der Strukturübertragung, die in der Regel durch Ätzung erreicht wird.

#### 4.1.1 Elektronenstrahl-Lithographie

Neben dem üblichen Verfahren der optischen Lithographie (siehe 4.1.2), bei dem die Energie durch Photonen übertragen wird, können auch Elektronen verwendet werden. Während das parallele Schreiben (Projektionsverfahren) mit Elektronen noch in den Anfängen steckt, da die Masken hierbei ein Problem darstellen, ist das serielle oder „direkte“ Schreiben etabliert. Möglich ist dies, da Elektronen geladene Teilchen sind, deren Strahlen fokussiert werden können; ein optisches System wird dadurch möglich. Das darauf basierende Lithographieverfahren heißt Elektronenstrahl-Lithographie oder „E-Beam-Schreiben“ und ist üblich zur Maskenherstellung oder Einzelfertigung von Wafern. Dabei wird ein dünner Strahl über die Probe geführt und so der für diese Energie empfindliche Resist seriell belichtet. Der Resist wird – wie bei der optischen Lithographie – vorher auf die Probe aufgeschleudert und getrocknet. Näheres zu Auswahl und Eigenschaften des hier verwendeten Resists ist in Kapitel 6 beschrieben. Vorteil von E-Beam-Schreiben ist das hohe Auflösungspotential (die Wellenlänge ist nicht begrenzend), Nachteil der geringe Durchsatz. Die wichtigsten Variablen beim E-Beam-Schreiben sind die Strahlgröße (*beam size*), die Schrittweite (*beam step size*), die Dosis und die Beschleunigungsspannung.

Der hier verwendeten Elektronenstrahlschreiber EBPG 5HR der Firma Leica verwendet das Gauß-Rundstrahl-System im Vektor-Scan-Verfahren. Dabei wird ein runder, gaußförmig gebündelter Elektronenstrahl nur auf die zu belichtenden Flächen

gelenkt. Andere mögliche Systeme von E-Beam-Schreibern sind z. B. das Rechteckstrahl- oder das Variabler-Strahl-System, ein anderes Verfahren das Raster-Scan-Verfahren, bei dem das gesamte Feld mit dem Elektronenstrahl überstrichen wird; die Blende wird dann nur bei den zu belichtenden Flächen geöffnet.

Während die Spannung beim Standardwert 50 kV unverändert gelassen wurde, waren die anderen Parameter Optimierungsgrößen. Bei der Strahlgröße wurde unterschieden zwischen Linienschreiben und Flächenschreiben: Die Flächen (Pads, Justiermarken etc.) wurden mit einem 300 nm-Strahl geschrieben, die Linien mit dem dünnsten Strahl, dessen Durchmesser nominell 15 nm beträgt.<sup>1</sup> Die Schrittweite ist abhängig von der Strahlgröße: Für glatte Kanten und gleichmäßige Flächenbelichtung muß die Schrittweite kleiner als der Strahldurchmesser sein, für eine akzeptable Schreibzeit darf sie nicht zu klein sein. Die Flächen wurden deshalb mit einer Schrittweite von 250 nm, die Linien mit einer Schrittweite von 12,5 nm geschrieben. Eine Ausnahme bildeten hier die *Single-Lines*, die mit einer Schrittweite von 5 nm geschrieben wurden (siehe Abb. 4.1).

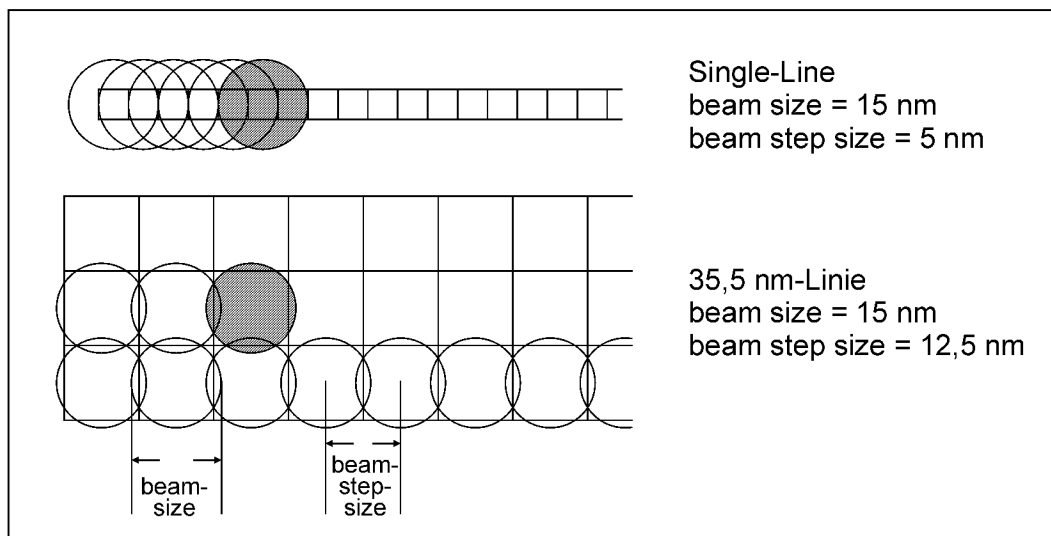


Abb. 4.1: Strahlführung beim Elektronenstrahlschreiben. Im gleichen Maßstab sind hier eine Single-Line und eine 37,5 nm-Linie schematisch dargestellt.

Die Abbildung 4.2 zeigt Auswirkung der Streuung in der Probe auf die Intensität im Resist: Mehrere Mikrometer abseits des Hauptstrahls wird der Resist durch tief im Substrat rückgestreute Elektronen belichtet. Dieser sog. Proximityeffekt ist insbesondere bei *lines and spaces*<sup>2</sup> sowie bei unbelichteten Flächen inmitten großer belichteter Flächen spürbar, da sich die geringe Intensität durch die Rückstreuung integriert, wenn viele Orte innerhalb der HWB<sub>2</sub> eines unbelichteten Ortes belichtet werden. Diese integrierte Proximitydosis kann in der Größe der Dosis durch den Hauptstrahl liegen.

<sup>1</sup>Auch bei gaußförmigem Strahl wird weiterhin von Strahlgröße oder -durchmesser statt von Halbwertsbreite gesprochen.

<sup>2</sup>*Lines and spaces* nennt man Linienfelder, bei denen der Abstand zwischen den Linien in der Größenordnung der Linienbreite ist.

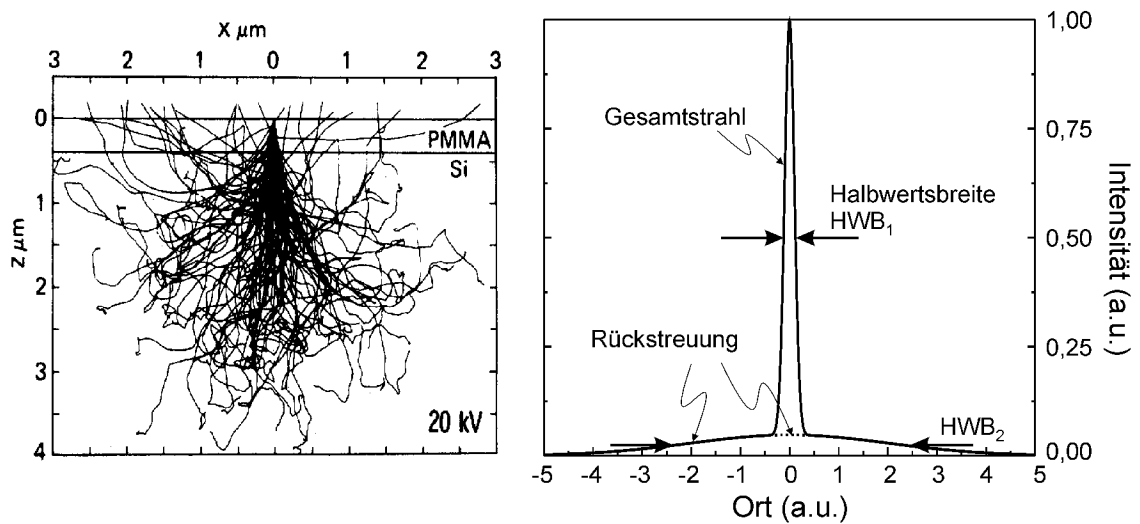


Abb. 4.2: Strahlform beim E-Beam-Schreiben. Streuung der Elektronen tief im Substrat bewirkt, daß der Resist auch mehrere Mikrometer abseits des Hauptstrahls belichtet wird. Links eine Monte-Carlo-Simulation von 20 kV-Elektronen in PMMA mit Silizium als Substrat (aus [22]). Rechts schematisch die resultierende Intensität im Resist. Die Halbwertsbreiten von Vorwärts- und Rückwärtsstreuung sind nicht maßstabsgetreu; der Hauptstrahl ist deutlich schärfer.

### 4.1.2 Optische Lithographie

Außer den dünnen Linien, die den aktiven Bereich des Transistors beinhalten, werden die übrigen Strukturen mit optischer Lithographie übertragen. Dazu wird ein lichtempfindlicher Film (Lack oder Resist genannt) auf die Probe aufgeschleudert. Die Strukturen liegen auf einer Maske vor und werden von dort auf die Probe abgebildet. Die Maske besteht aus Quarzglas mit einer strukturierten Chromschicht; die Bereiche der Probe, die nicht belichtet werden sollen, werden durch das Chrom maskiert. Bei der Belichtung verändert sich die chemische Struktur des Photolackes derart, daß er sich bei der anschließenden Entwicklung in der Entwicklerlauge löst (Positivprozeß).<sup>3</sup> Man unterscheidet zwischen Projektions-, Proximity- und Kontaktbelichtung. Bei der Projektionsbelichtung wird die Maske so in den Lichtweg eingebaut, daß ein (meist verkleinertes) Bild der Strukturen auf die Probe projiziert wird. Bei der Proximitybelichtung wird ein direkter Schatten auf die Probe geworfen, während die Maske einige Mikrometer von der Probe entfernt ist. Dies schont die Maske und die Probe, verringert aber die Auflösung. Beim hier angewendeten Verfahren der Kontaktbelichtung wird die Maske auf den Wafer gepreßt bzw. im Vakuummodus der Raum zwischen Maske und Wafer leergepumpt und so der Wafer an die Maske gesaugt.

Das Prinzip der Kontaktbelichtung hat verschiedene begrenzende Auflösungskriterien: zum einen die theoretische „minimal feature size“ (*MFS*, kleinste darstellbare Strukturgröße), die sich aus verwendeter Lackdicke  $d$  (Schirmabstand) und verwen-

<sup>3</sup>Beim Negativprozeß löst sich der unbelichtete Resist im Entwickler, während der belichtete Teil stehen bleibt.

deter Wellenlänge  $\lambda$  ergibt. Für Nahbeugung gilt

$$MFS \approx \sqrt{\lambda \cdot d}$$

Hieraus ergibt sich für  $\lambda = 250 \text{ nm}$  und Lackdicke  $d \approx 500\text{--}600 \text{ nm}$  eine theoretische minimale Strukturgröße von  $350\text{--}400 \text{ nm}$ .

Dieser theoretische Wert kann allerdings nur erreicht werden, wenn die Strukturgrößen auf der Maske mit mindestens dieser Auflösung realisiert werden können, im Gegensatz zur Projektionsbelichtung, bei der die Maske auf den Wafer verkleinert projiziert wird und sich so die Auflösung gegenüber der Maske verbessern läßt. Aus diesem Kriterium folgt die hier mögliche minimale Strukturgröße von ca.  $700 \text{ nm}$  durch optische Lithographie. Dies entspricht ungefähr dem Industriestandard von 1989.

Die Belichtung erfolgte hier durch den Belichter *MA 6* der Firma *Süss* (Abb. 4.3). Als Lichtquelle dient eine Xenon-Quecksilberdampfampe mit einer hohen Intensität bei  $\lambda \approx 250 \text{ nm}$ .

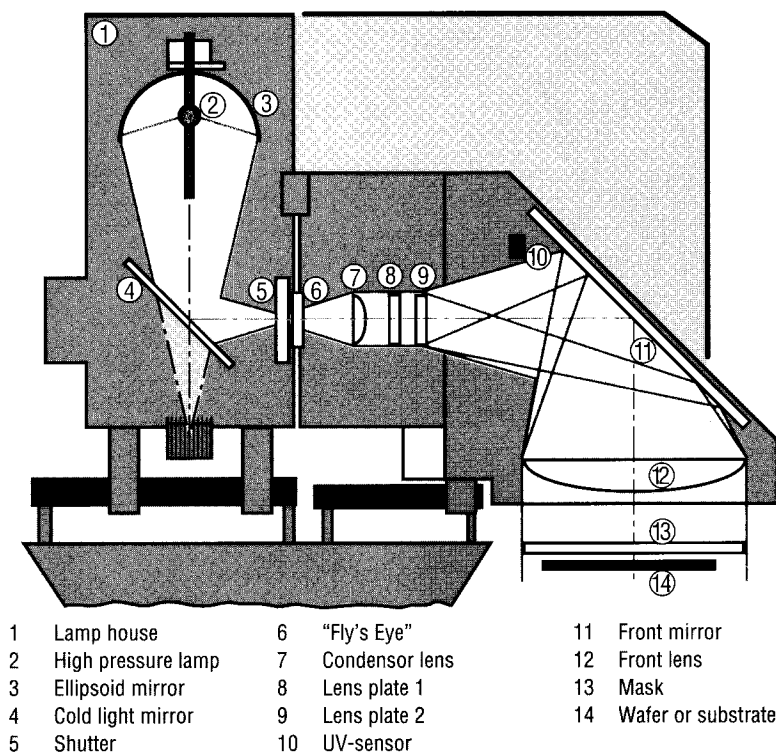


Abb. 4.3: Optisches System des Belichters *MA6* der Firma *Süss*.

### 4.1.3 Strukturübertragung durch RIE-Ätzen

Man kann zwei verschiedene Ätzprinzipien unterscheiden: die naßchemische Ätzung und die Trockenätzung. Bei der naßchemischen Ätzung wird die Probe in eine Ätzlösung eingebracht, die chemisch mit dem zu entfernden Material reagiert und

dieses dadurch auflöst. Nachteile sind die Isotropie der Ätzung und Benetzungsprobleme beim Ätzen kleinster Löcher. Die Trockenätzung stellt heute den wichtigsten Strukturübertragungsprozeß in der Halbleiterfertigung dar, weil durch ihre Anwendung eine sehr genaue Übertragung der immer kleiner werdenden Maskenstrukturen erzielt werden kann. Das „Reaktive Ionen Ätzen“<sup>4</sup> ist ein Trockenätzverfahren, das sich als Standardverfahren in der Halbleitertechnik etabliert hat.

In einer RIE (Abb. 4.4 (a)) liegt die zu ätzende Probe auf einer Elektrode, an die gegenüber einer anderen, geerdeten Elektrode im Hochvakuum ( $\sim 10^{-5}$  bar) eine hochfrequente Wechselspannung angelegt wird. Durch diese Wechselspannung werden die in den Reaktor geleiteten Prozeßgase teilweise dissoziiert ( $\sim 0,1\text{--}10\%$ ) bzw. ionisiert ( $\sim 0,001\text{--}0,01\%$ ). Durch die einsetzende Plasmaentladung bildet sich zwischen den Elektroden ein Potentialverlauf aus (Abb. 4.4 (b)).

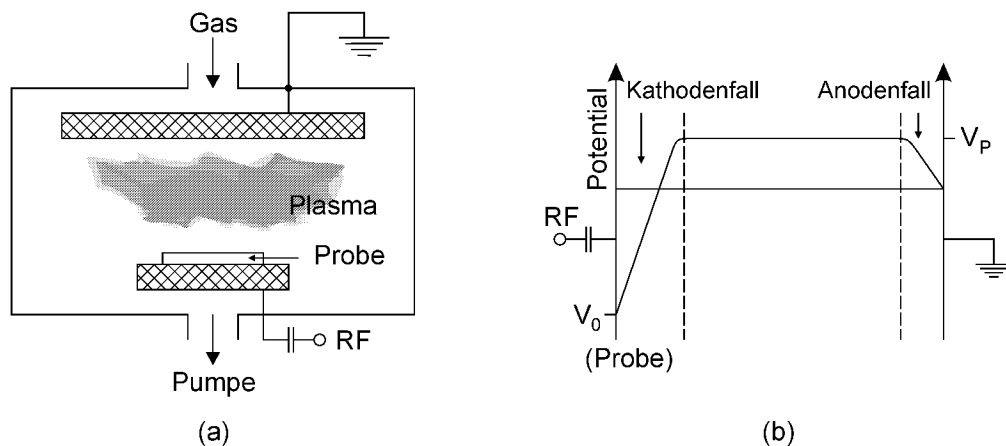


Abb. 4.4: Schematische Darstellung der RIE. (a) Reaktorgeometrie eines RIE-Parallelplattenreaktors. (b) Potentialverlauf während des Ätzprozesses.

Das Plasmapotential  $V_P$  ist durch die Prozeßparameter Gasfluß, Leistung, Wechselspannung und Prozeßdruck bestimmt. Im elektrischen Feld werden die reaktiven Ionen zur Probe hin beschleunigt; ist die Potentialdifferenz  $V_P - V_0$  zwischen Plasmapotential und Probenort groß, können die Ionen fast senkrecht auf die Probe auftreffen und anisotropes Ätzen ist möglich.

Der Ätzmechanismus ist eine Kombination aus physikalischem Ätzen (Sputtern), bei dem das Material durch den Impuls der auftreffenden Ionen abgetragen wird, und chemischem Ätzen, bei dem das Material mit den Ionen chemisch reagiert und flüchtige Reaktionsprodukte bildet. Die Ätzkomponenten beeinflussen sich gegenseitig, da die Bindungen im Material chemisch geschwächt werden und dadurch das Material leichter gesputtert werden kann. Andererseits werden durch den Ionenbeschuss Oberflächenschäden induziert und Energien bereitgestellt, die das chemische Ätzen erleichtert. Die physikalische Komponente des Ätzprozesses führt zur Anisotropie, während die chemische Komponente selektives Ätzen ermöglicht. Ergebnisse isotropen und anisotropen Ätzens werden in Abbildung 4.5 gezeigt.

<sup>4</sup>Man spricht auch von „RIE-Ätzen“ von engl. *reactive ion etcher*

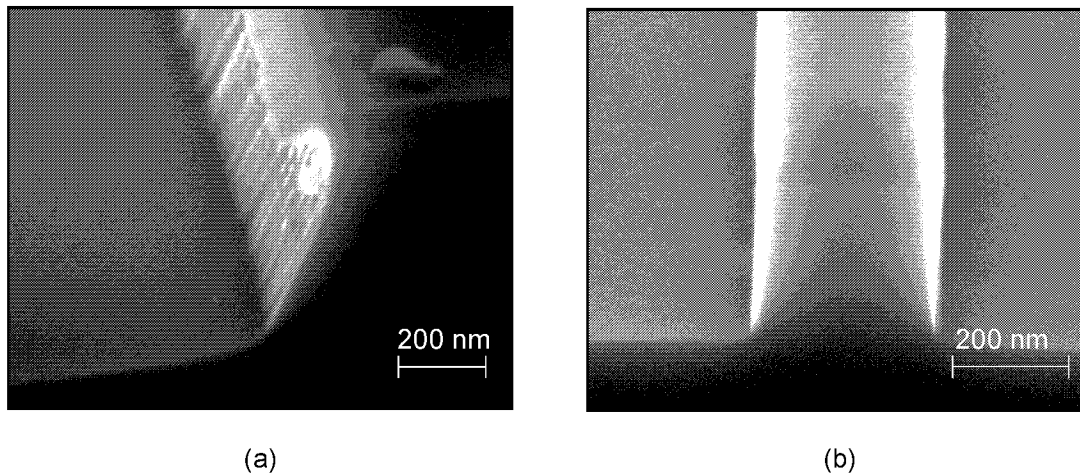


Abb. 4.5: REM-Aufnahmen: Beispiele für isotropes und anisotropes RIE-Ätzen. (a) Isotropes Ätzen von Silizium mit TEOS-SiO<sub>2</sub> als Maske. 7,5 min mit Plasma: Ar 40 ml, SF<sub>6</sub> 1 ml, 150 W, 20 mbar, auf Quarzteller. (b) Anisotropes Ätzen von Silizium mit Resist UV 6 als Maske. 10 min mit Plasma: Ar 40 ml, SF<sub>6</sub> 1 ml, 150 W, 20 mbar, auf Graphitteller.

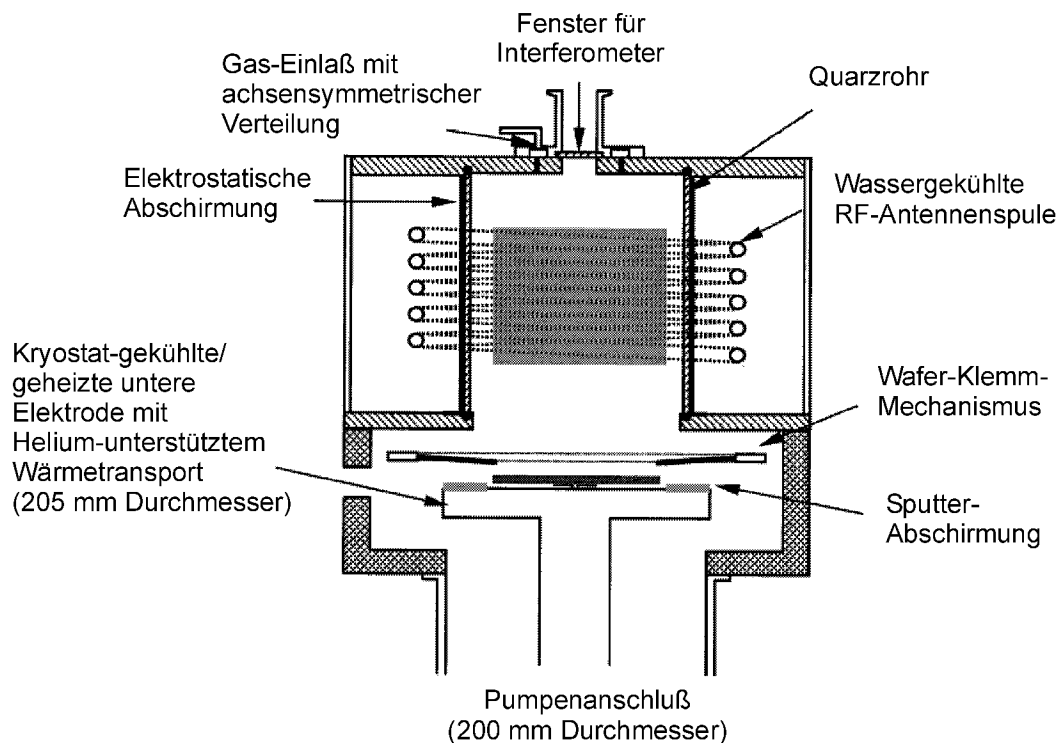


Abb. 4.6: Prinzip der ICP-RIE. Die Ionenstromdichte wird durch die induktiv gekoppelte RF-Leistung unabhängig von der Ionenenergie eingestellt, welche an der Substratelektrode eingestellt wird.

Bei dieser Standardtechnologie werden Ionenenergie und -stromdichte durch eine einzige angelegte Leistung erreicht. Es wurden Techniken entwickelt, um diese

Werte unabhängig voneinander zu kontrollieren: Mit einer „ICP-RIE“<sup>5</sup> wird das Plasma über eine Spule durch eine induktiv gekoppelte RF-Leistung erzeugt. Eine zweite, kapazitiv gekoppelte RF-Leistung an der Substratelektrode kontrolliert die Ionenenergie über die eingestellte Bias-Leistung. Die Komponenten einer solchen ICP-Kammer sind in Abbildung 4.6 dargestellt. Dieser weitere Freiheitsgrad ermöglicht besonders selektive Ätzprozesse (siehe auch 6.4).

## 4.2 Ionenimplantation

Die Ionenimplantation ist ein Dotierverfahren, bei der geladene Teilchen auf hohe Geschwindigkeit gebracht und auf einen Körper geschossen werden. Die Ionen werden im Innern des Targets durch Stöße abgebremst und lagern sich in den oberen Schichten ein. Ein Ionenimplanter besteht im Wesentlichen aus einer Ionenquelle, einer Beschleunigungseinheit, einer Separationseinheit, die die Ionen nach ihrer Masse und Ionisierung trennt, und einer Strahlablenkung (Scanner), die die Ionen gleichförmig über das Ziel verteilt (Abb. 4.7).

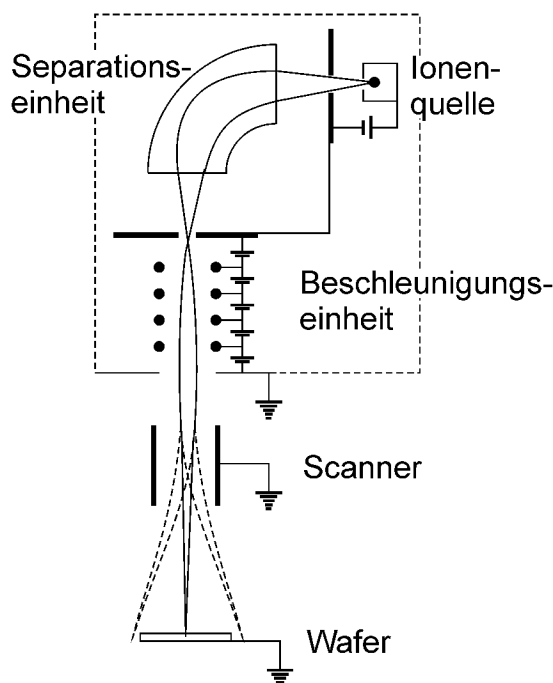


Abb. 4.7: Prinzip eines Ionenimplanters (nach [23]).

Der vorbeschleunigte Ionenstrahl besteht aus einem Gemisch von Ionen, aus denen die gewünschte Sorte herausgefiltert werden muß. In der einfachsten Form geschieht das in einem homogenen Magnetfeld senkrecht zum Ionenstrahl. Die Ionen werden auf Kreisbahnen abgelenkt, deren Radien von der Masse und der Ladung der Ionen abhängen. Der Austrittswinkel hinter der Separationseinheit ist also typisch

<sup>5</sup>Inductive Coupled Plasma, induktiv gekoppeltes Plasma

für die gewünschte Ionenart, so daß mit einer Blende die nicht benötigten Ionen herausgefiltert werden können. Am Ende seiner Bahn hat der Ionenstrahl nur einen geringen Durchmesser. Eine ionenoptische Aufweitung des Strahls wäre nicht homogen genug über einen Wafer, so daß durch waagerechten und senkrechten Feldplatten mit angelegter Wechselspannung der Strahl über eine größere Fläche geführt wird. Bei der Implantation in einen amorphen Festkörper mit zufälliger Stoßverteilung ergibt sich eine Gaußverteilung für das Tiefenprofil der implantierten Ionen, wie es auch in Abbildung 4.8 in Bereich A dargestellt ist. Ist das Target ein Einkristall, so ist jedoch sogenanntes *Channelling* möglich: Ionen, die in bestimmte kristallographische Richtungen einfallen, „sehen“ die Gitteratome hintereinander und können zwischen den Gitterebenen tief in den Kristall eindringen, ohne durch zufällige Streuung abgebremst zu werden (Abb. 4.8, Bereich C). Ein  $\langle 100 \rangle$ -Siliziumwafer z. B. darf deshalb nicht senkrecht zum Strahl eingebaut werden, sondern wird um  $7^\circ$  und  $21^\circ$  um die beiden anderen Raumachsen gekippt, um *Channelling* zu vermeiden.

Die implantierten Ionen bleiben allerdings meist auf Zwischengitterplätzen stehen – sie werden also nicht richtig in den Kristall eingebunden. Da die Ionen dann nicht zur elektrischen Leitung beitragen, müssen diese Störstellen durch eine thermische Nachbehandlung „aktiviert“ werden. Auch wird der Kristall während der Implantation teilweise zerstört, weil Atome aus ihren Gitterplätzen herausgeschlagen werden. Auch für die Beseitigung dieser Gitterschäden ist die nachfolgende Temperaturbehandlung notwendig. Um unerwünschte Diffusion der eingebrachten Ionen zu vermeiden, wird diese Temperung mit RTP durchgeführt.

In der Mikroelektronik ist die Ionenimplantation ein Standardverfahren zur Erzeugung von lokalen Dotierungen. Eine Vielzahl von Elementen kann dafür verwendet werden; die Beschleunigungsenergien reichen dabei von einigen hundert eV bis zu einigen MeV, wodurch die Eindringtiefe gesteuert werden kann.

Beim I-NovaFET-Prozeß wurde das Kanalgebiet durch Implantation und Diffusion definiert (siehe 5.2.2), beim EpiNovaFET die Schicht am Topkontakt (siehe 5.3.2). Für diese Prozesse wurde ein *Axcelis 8250 Implanter* verwendet, der mit Energien von 1 keV bis 250 keV arbeitet.

### 4.3 Secondary Ion Mass Spectrometry

*Secondary ion mass spectrometry* (SIMS) ist eine Analysetechnik zur Charakterisierung der Oberfläche und oberflächennahen Region von Festkörpern (und der Oberfläche einiger Flüssigkeiten). Der Beschuß mit primären Ionen bewirkt ein Sputtern an der Oberfläche, also den Abtrag der Oberflächenschichten in Form von ionisierten und neutralen Atomen oder Molekülen, die auch angeregt sein können. Die ionisierten Partikel, die Sekundär-Ionen, werden nach ihrer Masse analysiert und liefern somit bei zeitaufgelöster Messung ein Tiefenprofil der Zusammensetzung (Abb. 4.9) [25].

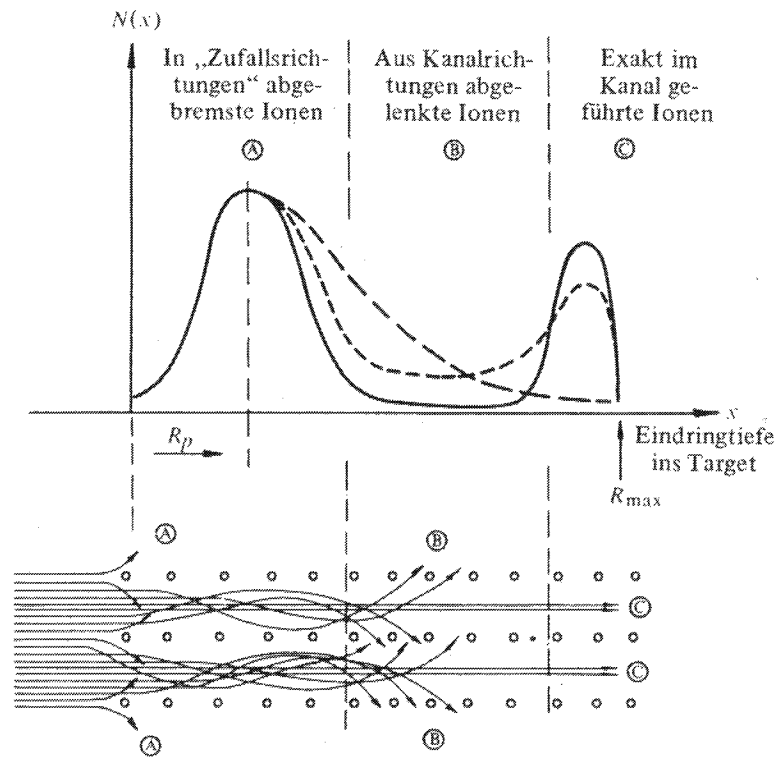


Abb. 4.8: Schematische Darstellung von möglichen Reichweitenprofilen für Ionen, die in einen Einkristall implantiert wurden.  $R_p$  ist die mittlere Reichweite für den amorphen Festkörper. Das Phänomen, daß Ionen tief eindringen können, wenn sie ungestreut tief in das Gitter eines Einkristalls geführt werden, nennt man „Channelling“. (Bild aus [24].)

## 4.4 Chemisch-Mechanisches Polieren (CMP)

Aus verschiedenen Gründen können Planarisierungen in einem Prozeßablauf notwendig sein. Am wichtigsten ist wohl eine planarisierte Fläche für die Verdrahtung der Bauelemente integrierter Schaltungen. Ein topographisches Dielektrikum verhindert die Herstellung gleichmäßiger Leiterbahnen, da die Schärfentiefe der optischen Lithographie von nur wenigen hundert Nanometern zu Unschärfe und Linienverbreiterungen führt, wenn Niveauunterschiede zu groß sind. Mehrere Planarisierungsmethoden sind etabliert, z. B. SOG,<sup>6</sup> Resist/Etchback<sup>7</sup> oder CMP. Während die anderen Verfahren lediglich lokal planarisieren (1–100  $\mu\text{m}$ ), ist mit CMP eine Nivellierung über den gesamten Wafer möglich.

Beim CMP wird der Wafer außermittig auf ein rotierendes Polierruch (*Pad*) gedrückt, während eine Polierflüssigkeit, die *Slurry*, dazugegeben wird (siehe Abb. 4.10). Die *Slurry* für Oxid-CMP besteht aus einer basischen Lösung, in der kleine

<sup>6</sup>*Spin On Glass*, eine Planarisierungsmethode, bei der ein gelöster Isolator aufgeschleudert und anschließend thermisch vernetzt wird. Durch das flüssige Auftragen wirkt es planarisierend. HSQ, der für die Nanolithographie verwendete Elektronenstrahlresist, ist ein solches SOG.

<sup>7</sup>Beim Resist/Etchback-Verfahren wird zunächst  $\text{SiO}_2$  konform deponiert, anschließend ein Resist aufgeschleudert. Beides wird dann mit gleicher Ätzrate zurückgeätzt.

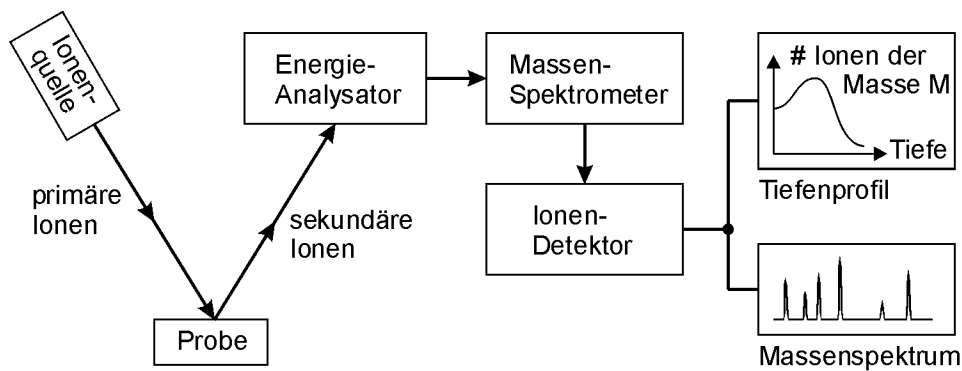


Abb. 4.9: Prinzip der SIMS: Durch Sputtern ausgelöste Ionen werden analysiert und liefern ein Tiefenprofil der Atomkonzentration.

Schleifkörper (Abrasive) aus  $\text{SiO}_2$  schweben.<sup>8</sup> Diese sind durch direkten Abrieb für die mechanische Komponente der Abtragung verantwortlich. Die chemische Komponente wird wesentlich durch das Wasser in der *Slurry* beeinflusst:  $\text{H}_2\text{O}$ -Moleküle passivieren die Oxidoberflächen sowohl an den Abrasiven als auch auf der Wafer. Zwischen diesen Passivierungen kann es nun zu einer Reaktion kommen, bei der molekulare Bindungen entstehen, die dann Abrasive und Oxidschicht verbinden. Ohne mechanische Komponente fände hier also eine Passivierung oder sogar eine Deposition statt, so allerdings werden die Partikel und damit auch Teile aus der Oxidschicht vom Poliertuch mitgerissen. Der chemische Anteil bewirkt eine deutlich schnellere Abtragung sowie eine bessere Oberflächenstruktur der polierten Schicht. (Erläuterungen nach [26].)

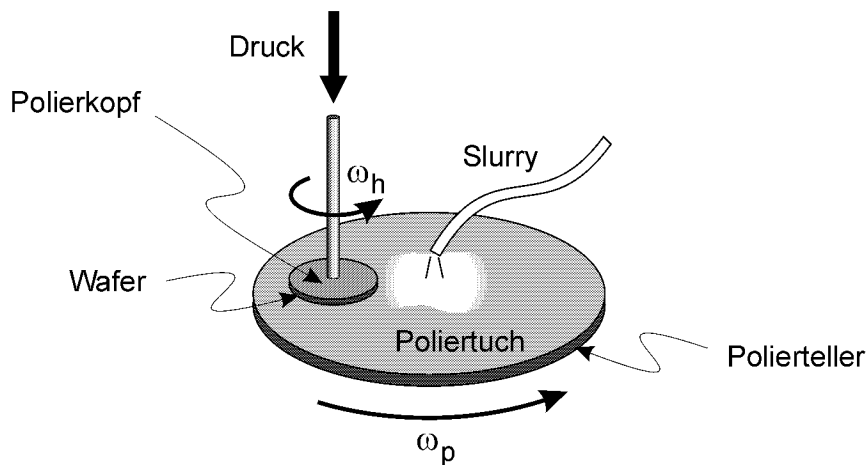


Abb. 4.10: Schematische Darstellung des Chemisch-Mechanischen Polierens. Auf ein rotierendes Poliertuch (Pad), das mit der sog. „Slurry“ bedeckt ist, wird der ebenfalls rotierende Wafer gepreßt.  $\text{SiO}_2$ -Partikel in der Slurry bewirken mechanisches Schleifen, die chemische Komponente eine Ätzung sowie das Auflösen der Schleifprodukte.

<sup>8</sup>Die Base (meist  $\text{NH}_4\text{OH}$  oder  $\text{KOH}$ ) stabilisiert die Abrasive in der Flüssigkeit. Sie trägt nicht zur chemischen Komponente bei.

Viele Parameter gehen in das Ergebnis der Planarisierung ein: Die zu polierende Oberfläche und deren Material, die Eigenschaften des Polierpads, die Zusammensetzung der Slurry sowie diverse geometrische und kinetische Größen. Das Know-how für das äußerst genaue Polieren ist in weiten Teilen aus dem Schleifen von Linsen transferiert worden. Einige Variablen können vom Anwender optimiert werden, um ein für die Anforderungen des jeweiligen Prozesses optimales Ergebnis zu erzielen: der Druck des Polierkopfes auf das Polierruch, die Drehzahlen von Polierkopf und -teller und natürlich die Dauer des Polierschrittes. Maße für das Polierergebnis sind die Abtragsrate und die Inhomogenität.



# Kapitel 5

## Der NovaFET-Prozeß

Der NovaFET ist ein vertikaler Double-Gate-MOSFET. Wie in Abbildung 5.1 ersichtlich ist, ist das Zentrum ein Siliziumsteg, der dünn genug sein muß, damit das Kanalgebiet vollständig verarmen oder – im besten Fall – sogar vollständig invertieren kann. Vertikal angeordnet in diesem Steg ist die p-n-p-Schichtabfolge für p-Kanal-Transistoren bzw. n-p-n-Folge für n-Kanal-Transistoren. Die Gatelänge ist somit die Dicke dieser mittleren Schicht, die im Gegensatz zur Stegbreite nicht lithographisch definiert werden muß, sondern durch Epitaxie oder Implantation und Diffusion eingestellt wird. Wichtige Prozesse der Herstellung sind die Definition der Stege, die Erzeugung der Schichtfolge und der Topkontakt.

Im Folgenden werden zunächst die Prozeßfolge des ursprünglichen Layouts vorgestellt und die Grenzen dieses Konzepts diskutiert. Zwei weiterentwickelte Layouts – die Implantationsvariante und der EpiNovaFET, die sich in der Herstellung der Schichtfolge unterscheiden – werden im Anschluß beschrieben und die für ihre Realisierung benötigten Schlüsselprozesse erläutert. Die Herstellung der Stege wird separat in Kapitel 6 behandelt.

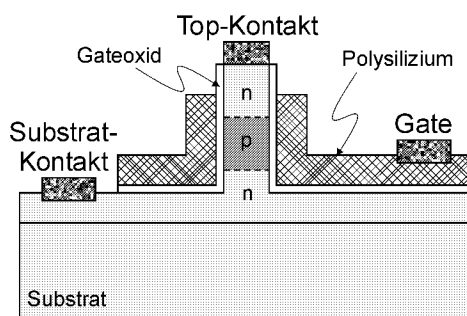


Abb. 5.1: Schematischer Aufbau des NovaFETs.

### 5.1 NovaFET – ursprüngliches Layout

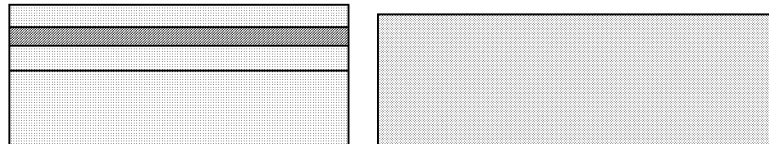
Bei dieser ursprünglichen Variante wird der aktive Bereich als p-n-p-Schichtfolge epitaktisch gewachsen. Der obere Kontakt wird entlang des Steges herausgeführt

und über ein Pad kontaktiert.

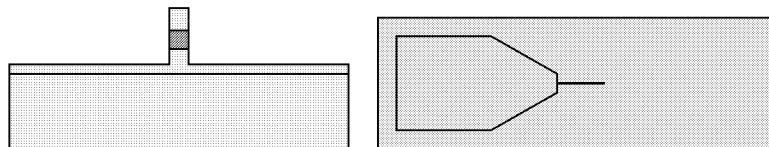
### 5.1.1 Prozeßfolge

Die Prozessierung des Transistors wird hier qualitativ beschrieben. Das genaue Prozeßprotokoll ist im Anhang A.1 dokumentiert.

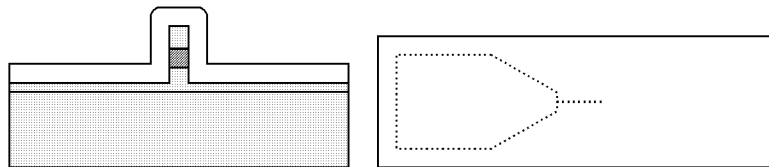
Epitaxie von p-, n- und p-Silizium auf p-Siliziumsubstrat  $\langle 100 \rangle$ .



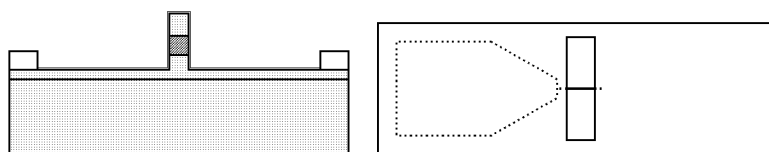
E-Beam-Lithographie und RIE-Ätzen von Pads und Stegen (s. Kap. 6).



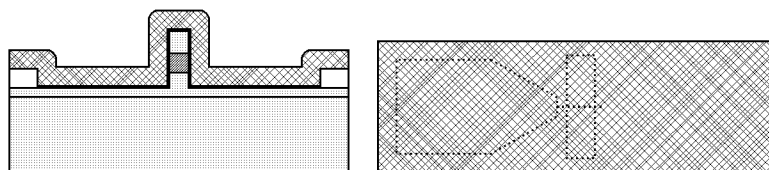
Deposition des ersten Deckoxids: 300 nm  $\text{SiO}_2$



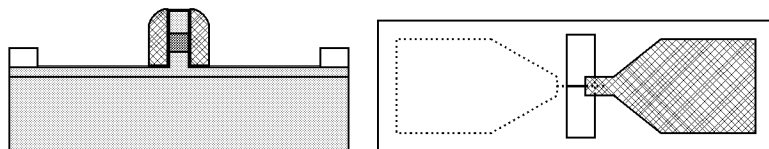
Öffnen des aktiven Bereichs durch opt. Lithographie und naßchem. Ätzen des Deckoxids, anssl. Gateoxidation.



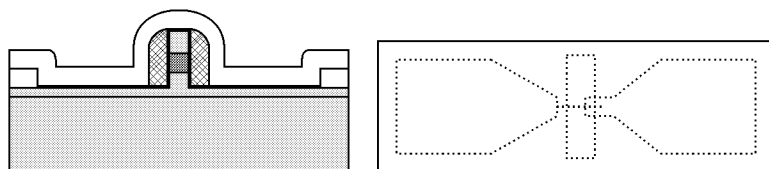
Deposition von 200 nm p-Polysilizium.

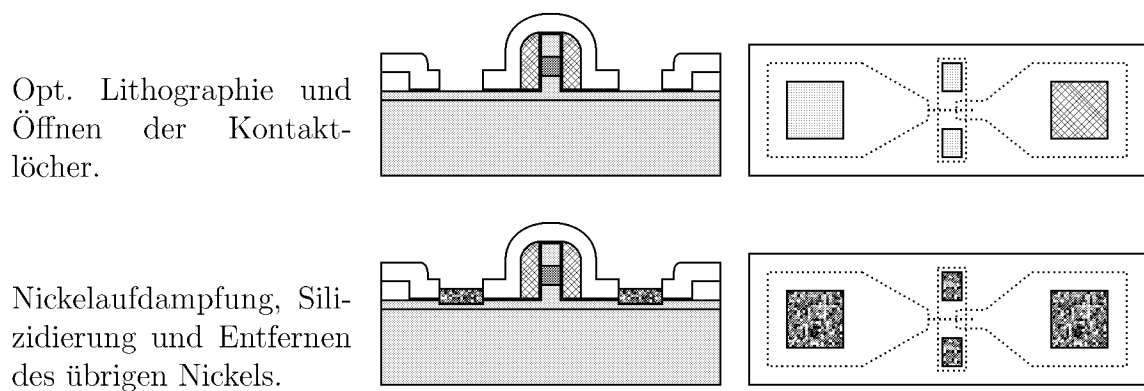


Opt. Lithographie für die Gateelektrode und anisotropes Ätzen des Poly-Si, um Gate-Spacer an den Stegen zu erhalten.



Deposition des zweiten Deckoxids: 300 nm  $\text{SiO}_2$





Zuleitungen zu den silizidierten Kontakten durch opt. Lithographie, Aufdampfung von Aluminium und Lift-Off.

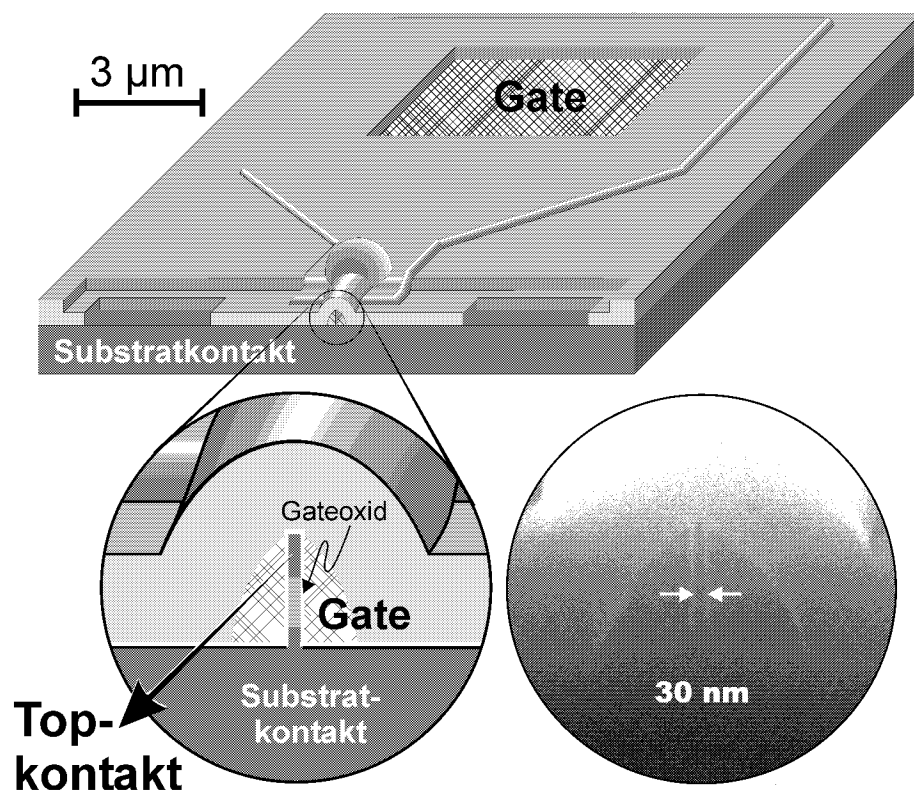


Abb. 5.2: Querschnitt durch den NovaFET im ursprünglichen Layout in realistischem Maßstab. Der Ausschnittsvergrößerung des aktiven Bereichs ist die REM-Aufnahme einer solchen Struktur gegenübergestellt.

### 5.1.2 Schlüsselprozesse

#### Dotierfolge

Die Schichtfolge wurde in einer LPCVD epitaktisch gewachsen. Abbildung 5.3 zeigt das SIMS-Spektrum<sup>1</sup> einer gleichzeitig gewachsenen Teststruktur, allerdings auf undotiertem Substrat. Die Dicke der mittleren, n-dotierten Schicht und somit die Kanallänge beträgt 140 nm. Die obere Schicht ist 130 nm dick. Dies bedeutet, daß die Stege mindestens 270 nm hoch sein müssen. Problematisch ist der flache untere p-n-Übergang. Er ist bedingt durch den sogenannten Memory-Effekt: Beim Wachsen einer Phosphor-dotierten Schicht deponiert ein Film an den Kammerwänden und auf dem Suszeptor, aus dem beim Wachsen der folgenden Bor-dotierten Schicht Phosphor wieder in die Kammer diffundiert. Dadurch fiel die Phosphor-Konzentration in der p-Schicht kaum ab, wenn keine Gegenmaßnahmen getroffen würden. Ein Tempersschritt von 1000°C für eine Stunde nach der Deposition der n-Schicht reinigt die Kammer, so daß die p-Schicht wieder mit großem Phosphor-Dotiergefälle gewachsen werden kann. Allerdings verursacht ein solcher Tempersschritt beim bereits bestehenden p-n-Übergang starke Diffusion, die sich auf die elektrischen Eigenschaften des Übergangs auswirkt: Obwohl der Transistor eigentlich symmetrisch funktionieren sollte, sind deutliche Unterschiede zwischen *Source on Top* und *Source on Bottom* sichtbar (siehe 7.1).

Eine weitere Einschränkung ist die Konzentration der Dotierung:  $3 \cdot 10^{17}$ – $2 \cdot 10^{18}$  at cm<sup>-3</sup> (untere Schicht) bzw.  $4$ – $5 \cdot 10^{18}$  at cm<sup>-3</sup> (obere Schicht) erzeugen zu geringe Ladungsträgerkonzentrationen. Höhere Konzentrationen und schärfere Übergänge können durch Implantationen hergestellt werden. Dies wurde beim weiterentwickelten Layout realisiert.

#### Gateoxid

Ein gebräuchliches Verfahren zur Herstellung von Gateoxiden ist die *Rapid thermal processing*-Oxidation (RTP). Beim RTP werden einzelne Wafer kurzzeitig durch Lampen aufgeheizt; die Kammer besteht aus Quarzglas, das für die Strahlung weitgehend transparent ist und sich somit nicht im gleichen Maße erwärmt. Da die Kammer nicht auf die Wafertemperatur gebracht werden muß, sind schnellere Temperaturrampen möglich als in einem Konvektionsofen. Das Gateoxid für diesen Transistor wurde mit einem Tempersschritt von 950°C für 3 min hergestellt und war 6 nm dick (Ellipsometermessung auf dem Substrat).

#### Gatespacer-Ätzung

Die Ätzung des Polysiliziums, um die Gatespacer zu erhalten, sollte im optimalen Fall eine Überätzung beinhalten, damit das Gate nicht mit der oberen Schicht parasitäre Kapazitäten bildet. Dafür ist jedoch ein Ätzprozeß mit hoher Selektivität zwischen Silizium und SiO<sub>2</sub> notwendig, da ansonsten sowohl die obere Zuleitung als auch die Substratzuleitung zerstört werden (Abb. 5.4). Bei der Herstellung des No-

---

<sup>1</sup>„Secondary ion mass spectrometry“, siehe 4.3.

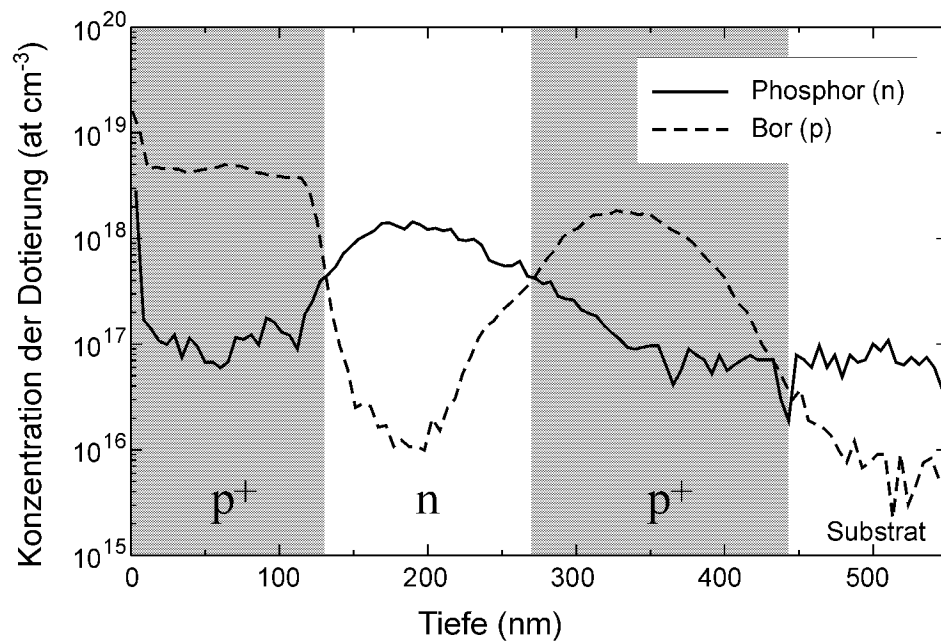


Abb. 5.3: SIMS-Messung der verwendeten Dotierfolge. Der untere p-n-Übergang ist auf Grund eines Temperschriffs deutlich flacher als der obere.

vaFETs im ursprünglichen Layout stand für das Silizium-Ätzen nur das anisotrope, aber kaum selektive Ar/SF<sub>6</sub>-Plasma zur Verfügung, so daß auf eine Rückätzung verzichtet werden mußte. Dies hatte auch Folgen für den Topkontakt, wie im Folgenden gezeigt wird. Im weiterentwickelten Layout konnte diese Rückätzung berücksichtigt werden.

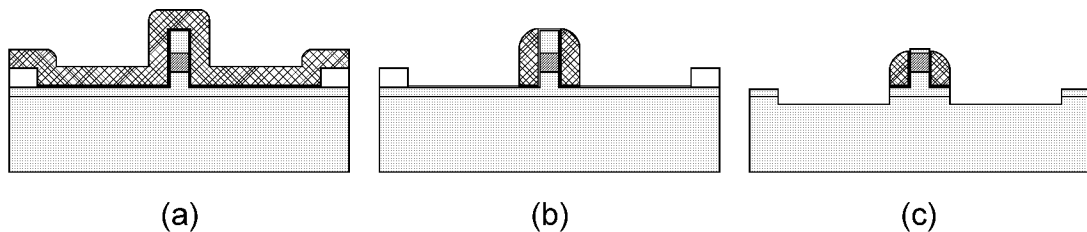


Abb. 5.4: (a) Querschnitt des NovaFETs vor der Polyspacer-Ätzung. (b) Bei dem nicht selektiven Plasma muß ohne Überätzung gestoppt werden, so daß parasitäre Kapazitäten mit der oberen Zuleitung entstehen. (c) Rückätzung mit dem nicht selektiven Ar/SF<sub>6</sub>-Plasma würde die obere Schicht und die Substratzuleitung zerstören.

### Topkontakt

Für den Topkontakt gibt es grundsätzlich zwei Möglichkeiten: Entweder wird der Steg in seinem aktiven Bereich von oben kontaktiert, oder der Strom wird durch die obere Schicht nach außen geführt und über ein Pad kontaktiert. Die zweite Variante verursacht hohe Zuleitungswiderstände, da die obere Schicht praktisch ein Halbleiterdraht mit  $140 \times 50 \text{ nm}^2$  Querschnitt ist. Allerdings ist für den direkten

Topkontakt die Rückätzung des Polyspacers nötig, um Kurzschlüsse zwischen Metallisierung und Gate zu verhindern (Abb. 5.5). Diese Rückätzung jedoch war, wie schon zuvor beschrieben, zu diesem Zeitpunkt nicht möglich. Für das neue Layout jedoch war diese Beschränkung nicht mehr gegeben, so daß dort direkte Topkontakte angebracht werden konnten.

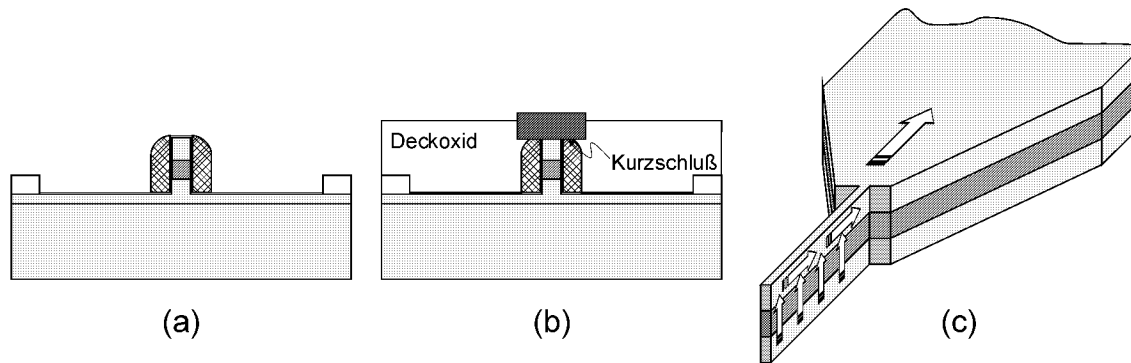


Abb. 5.5: Topkontakt beim NovaFET im ursprünglichen Design: (a) Polyspacer ohne Rückätzung. (b) Direkte Topkontakte ohne Rückätzung verursachen Kurzschlüsse. (c) Der Topkontakt muß deshalb durch den Steg nach außen geführt und an einem Pad abgenommen werden.

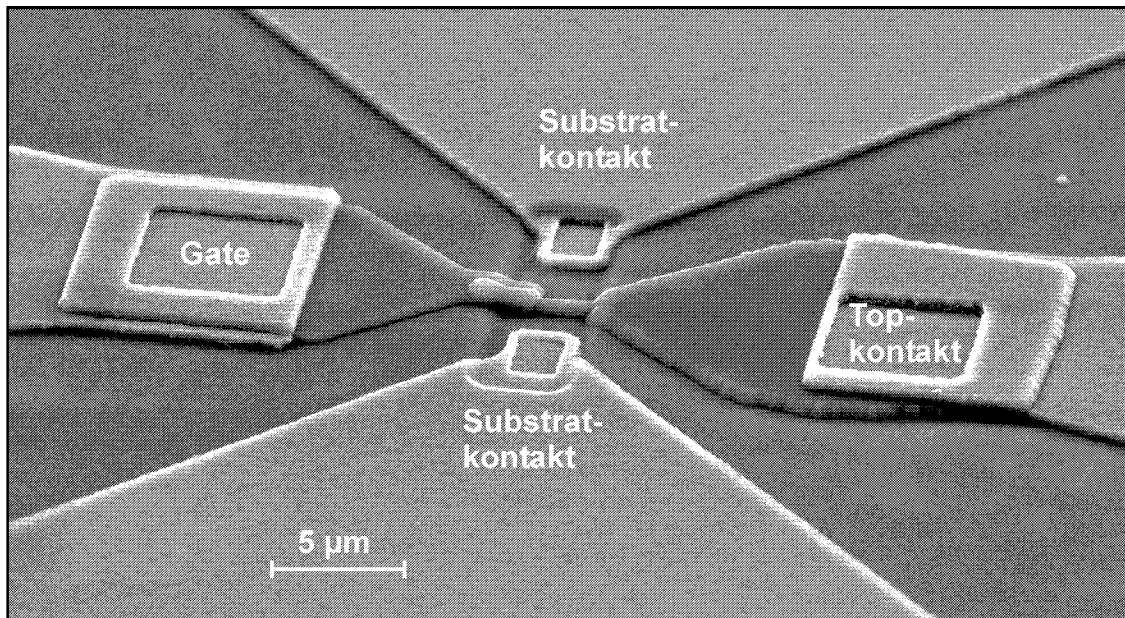


Abb. 5.6: REM-Bild des NovaFETs mit 3 µm aktiver Steglänge.

## 5.2 I-NovaFET – die Implantationsvariante

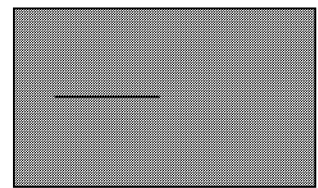
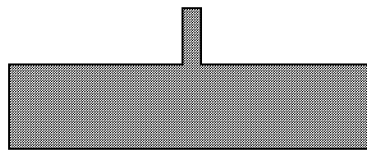
Das hier verwendete Layout unterscheidet sich in zwei wesentlichen Punkten vom vorherigen: Zum einen wird die Dotier-Abfolge hier alleine mit Implantation reali-

siert, und zum anderen wird die obere Schicht direkt oben an der Mesa kontaktiert. Dies erfordert eine Planarisierung, die hier mit CMP durchgeführt wird.

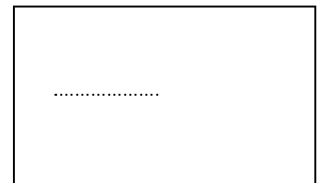
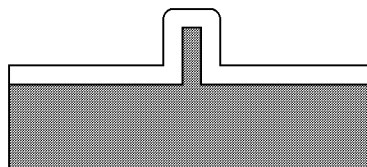
### 5.2.1 Prozeßfolge

Die Prozessierung des Transistors wird hier qualitativ beschrieben. Das genaue Prozeßprotokoll ist im Anhang A.2 dokumentiert.

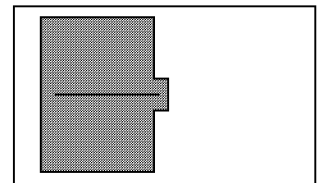
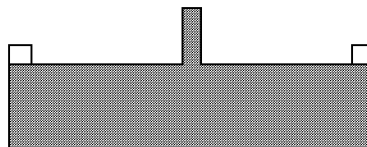
E-Beam-Lithographie und Ätzen der Stege in n-Silizium-Substrat mit Kanaldotierung (p-Silizium-Substrat für n-Kanal-MOSFET).



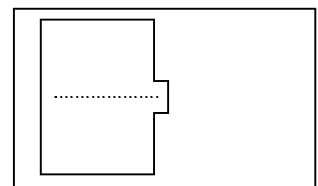
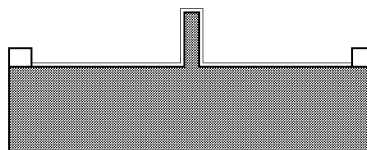
Deposition der Implantationsmaske: 100 nm TEOS-Oxid



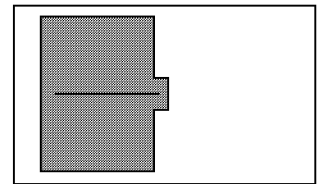
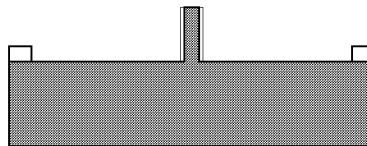
Öffnen des Implantationsbereichs durch opt. Lithographie und naßchem. Ätzen des Oxids.



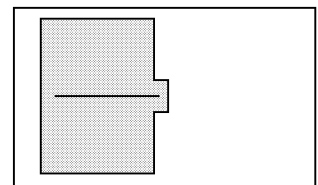
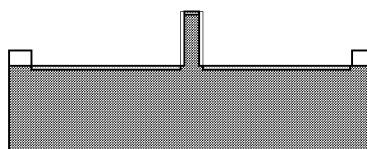
Therm. Oxidation von 15 nm SiO<sub>2</sub>.



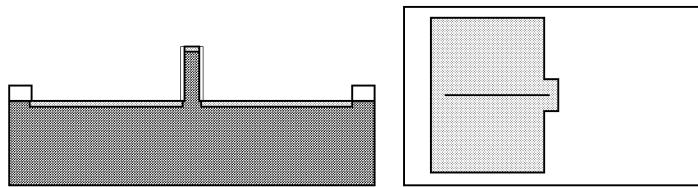
Anisotropes Ätzen des therm. Oxids, um Schutz-Spacer an den Stegen zu erhalten.



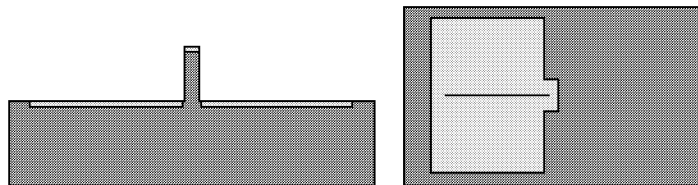
Implantation mit Bor (Arsen für n-Kanal-MOSFET).



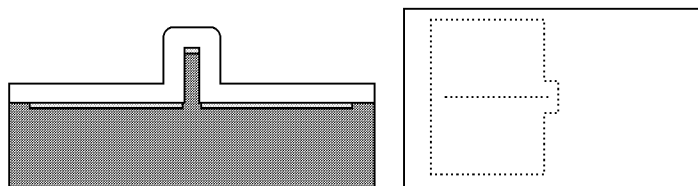
Annealing und Diffusion 1



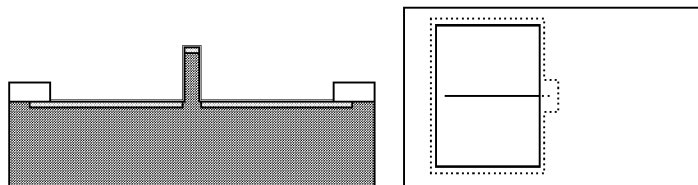
Entfernen der Spacer und der TEOS-Maske in 0,5%iger Flußsäure.



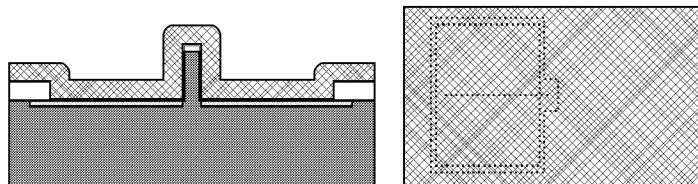
Deposition des ersten Deckoxids: 100 nm  $\text{SiO}_2$



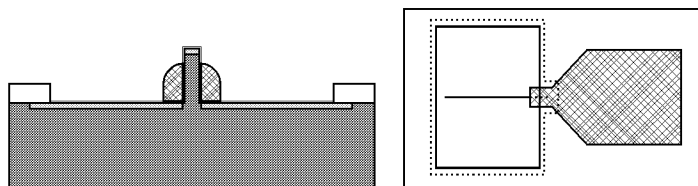
Öffnen des aktiven Bereichs durch opt. Lithographie und naßchem. Ätzen des Deckoxids, anschl. Gateoxidation.



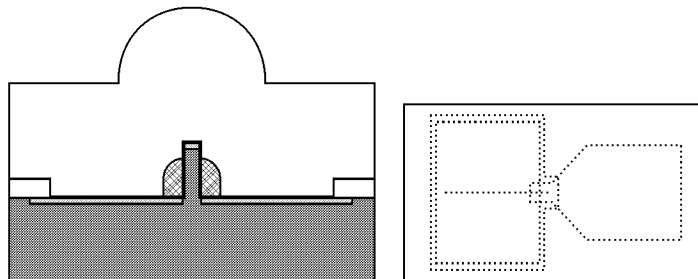
Deposition von 150 nm p-Polysilizium (bzw. n-Polysilizium).



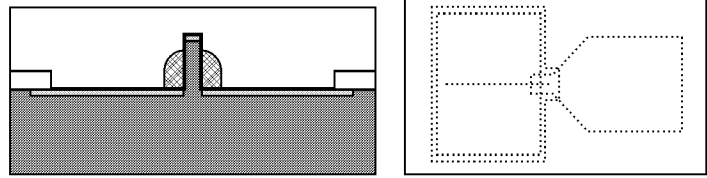
Opt. Lithographie für die Gateelektrode und anisotropes Ätzen des Poly-Si, um Gate-Spacer an den Stegen zu erhalten.



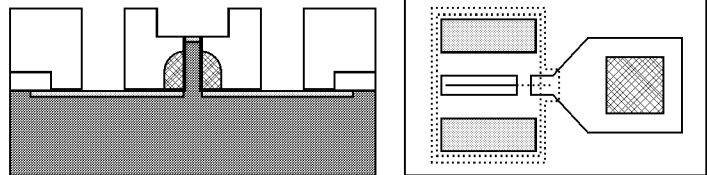
Deposition des zweiten Deckoxids: 700 nm  $\text{SiO}_2$



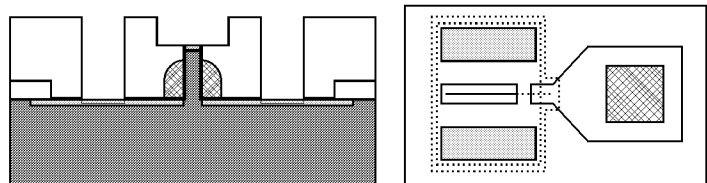
Chemisch-mechanisches Polieren (CMP) zum Planarisieren.



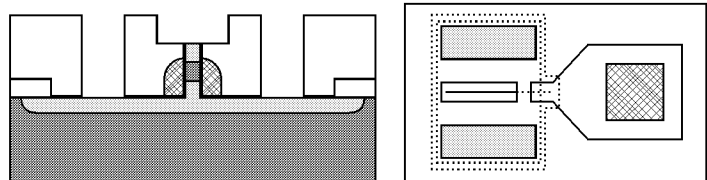
Opt. Lithographie und Öffnen der Kontaktlöcher.



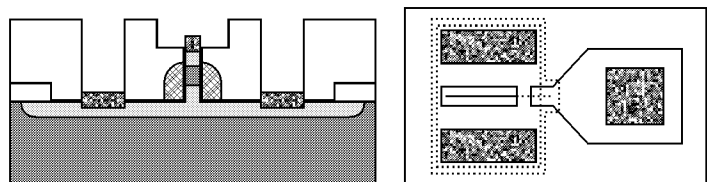
Kontaktimplantation mit Bor (Arsen für n-Kanal-MOSFET).



*Annealing* und Diffusion 2: Bildung des Kanalbereichs.



Nickelaufdampfung, Silizidierung und Entfernen des übrigen Nickels.



Zuleitungen zu den silizidierten Kontakten durch Aufdampfung von Aluminium, opt. Lithographie und Ätzen des Aluminiums.

### 5.2.2 Schlüsselprozesse

#### Dotierfolge

Höhere Dotierungen und schärfere Übergänge als mit LPCVD-Epitaxie können mit Implantation hergestellt werden. Um einen Kanal zu erzeugen, ist es jedoch notwendig, in die bestehenden Nanostrukturen zu implantieren. Durch die Diffusions- und *Annealing*-Tempereschritte bildet sich auch im unteren Bereich des Steges ein p-n-Übergang, so daß ein Kanal ausgebildet wird, dessen Länge nun von der Implantation, der Temperung und der Höhe des Steges abhängig ist. Dieses Verhalten konnte in

Simulationen bestätigt werden,<sup>2</sup> jedoch ist die experimentelle Bestimmung mittels SIMS für Nanostrukturen schwierig, da hierbei keine Ortsauflösung gegeben ist. Für die Implantation müssen Schutzspacer an den Seitenwänden des Steges angebracht werden, damit die Implantation nur in die Oberseite und ins Substrat, nicht aber in den gesamten Steg erfolgt. Diese Schutzspacer werden durch thermische Oxidation von 15 nm SiO<sub>2</sub> und anschließendem anisotropen RIE-Ätzen erzeugt. Dadurch werden die Stege noch einmal um insgesamt ca. 10 nm ausgedünnt und gleichzeitig die Oberfläche geglättet. Bild 5.7 zeigt Simulationen von Bor-Implantation in dünne Siliziumstege (nur Topkontakt). Implantiert wurde nach Optimierung der Simulationen mit folgenden Werten:

---

p-Kanal-MOSFET:	Bor	1 keV	$2 \cdot 10^{15} \text{ cm}^{-2}$	7°/21°
n-Kanal-MOSFET:	Arsen	10 keV	$2 \cdot 10^{15} \text{ cm}^{-2}$	7°/21°

---

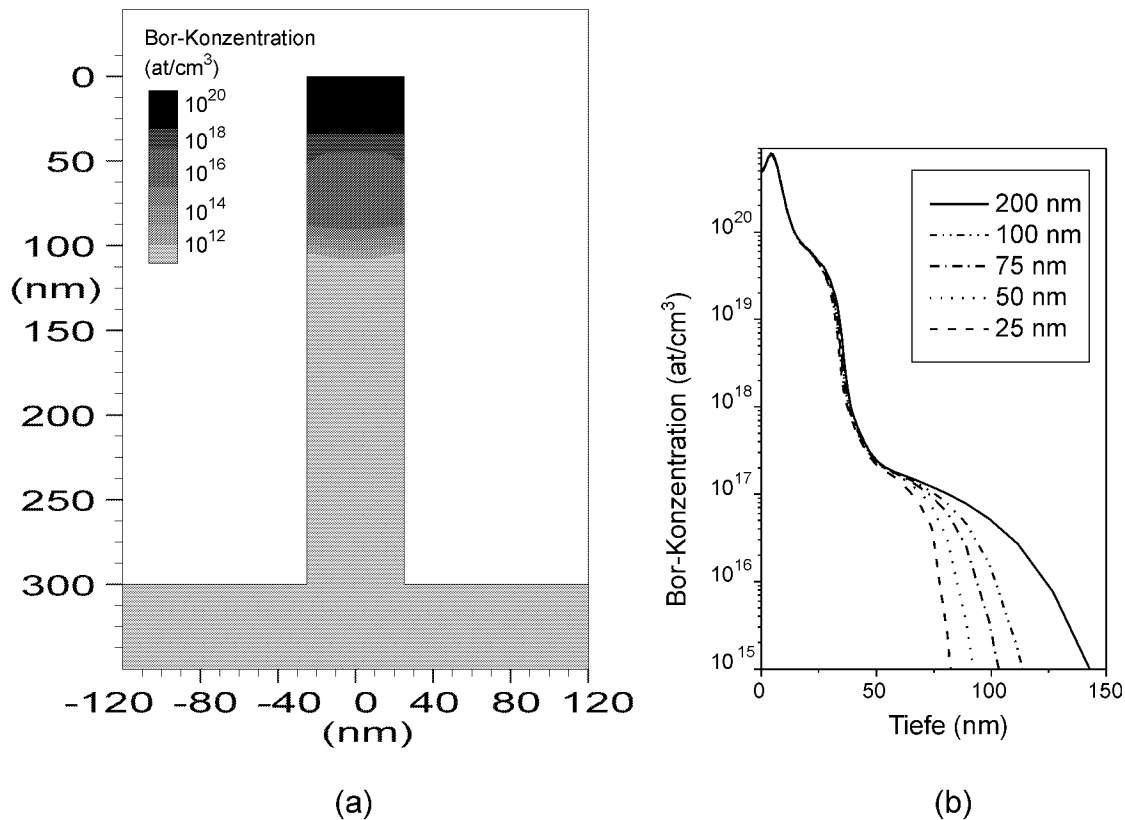


Abb. 5.7: Implantation in dünne Siliziumstege: (a) Simuliertes Dotierprofil eines 50 nm weiten Steges. Implantation mit  $1 \times 10^{15} \text{ at/cm}^2$ , 1 keV Bor, Temperung 1050° C Spike-Annealing. (b) Querschnitt des Dotierprofils im Zentrum des Steges für verschiedene Stegbreiten. (Abbildungen nach [27].)

---

<sup>2</sup>Verwendet wurde SSUPREM IV der Firma *Silvaco*.

### Gateoxid

RTP-Trockenoxidation ist die übliche Methode zur Herstellung von Gateoxiden und fand auch bei der ursprünglichen Variante des NovaFETs Anwendung (siehe 5.1.2). Da bei einer RTP-Oxidation der Wafer jedoch – wenn auch nur für wenige Sekunden – hohen Temperaturen ausgesetzt wird (Größenordnung: 3 min 950°C), kommt es zu einer verstärkten Diffusion der Dotierungen aus den Source- und Draingebieten in den Kanal, besonders bei Bor-Dotierungen (p-Kanal-MOSFETs). Deshalb ist ein Niedrigtemperatur-Feuchtoxidationsprozeß, der von Appenzeller *et al.* [28] entwickelt wurde, übertragen worden. Die Feuchtoxidation ist ein Standardverfahren zur thermischen Oxidation von Silizium. Dabei wird das Substrat in einem Konvektions-Oxidationsofen bei Normaldruck und Temperaturen üblicherweise zwischen 800°C und 1200°C unter Wasserdampf oxidiert. Die vereinfachte Oxidationsreaktion ist  $\text{Si} + \text{H}_2\text{O} \rightarrow \text{SiO}_2 + \text{H}_2$ . Bei relativ dünnen Oxiden ist die Oxidationsrate proportional zur Oxidationszeit, die Temperatur geht dagegen exponentiell ein. Die etablierten Prozesse liefern jedoch zu große Oxidationsraten, als daß diese Schichten als Gateoxide in Frage kämen.

Der modifizierte Prozeß von Appenzeller jedoch arbeitet bei deutlich niedrigeren Temperaturen. So wurde für den NovaFET ein Prozeß etabliert, der z. B. bei 600°C in vier Stunden eine Oxidschicht der Dicke 3,5–4 nm erzeugt. Trotz einer etwas höheren Grenzflächenzustandsdichte gegenüber den RTP-Oxiden (ca.  $3 \cdot 10^{11} \text{ cm}^{-2} \text{ eV}^{-1}$  gegenüber  $8 \cdot 10^{10} \text{ cm}^{-2} \text{ eV}^{-1}$ , [29]) überwiegen die Vorteile, so daß dieser Prozeß bei der Implantationsvariante erstmals angewandt wurde.

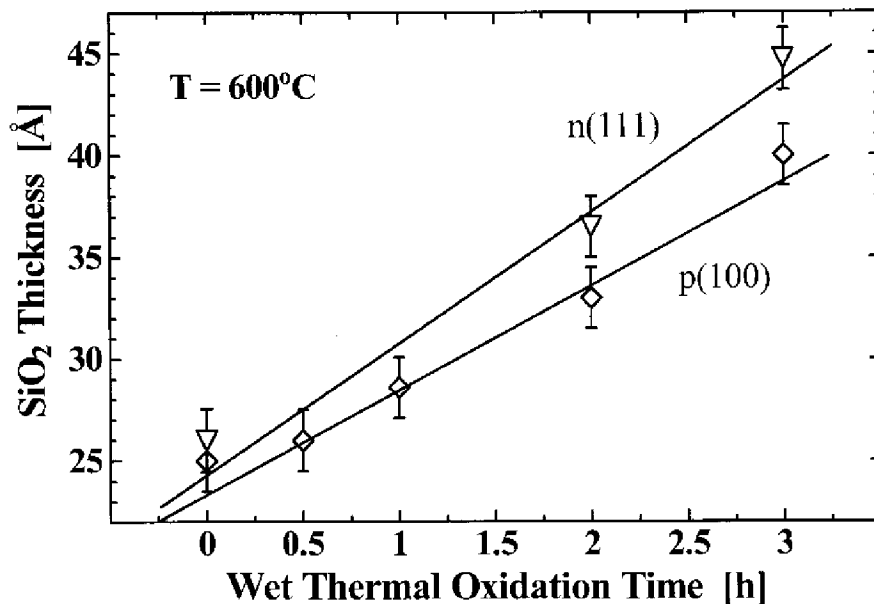


Abb. 5.8: Mit Niedrigtemperatur-Feuchtoxidation können bei 600°C Gateoxide definiert hergestellt werden. (Bild aus [28].)

### Gatespacer-Ätzung

Bei der Ätzung der Gatespacer ist eine kontrollierte Rückätzung wünschenswert für minimale parasitäre Kapazitäten zwischen Gate und Topkontakt, insbesondere aber für die direkten Topkontakte, bei denen ansonsten Kurzschlüsse aufträten (vergl. Abb. 5.5 b). Allerdings muß diese Polysilizium-Ätzung hier mit sehr hoher Selektivität zu  $\text{SiO}_2$  erfolgen, da ansonsten der Steg und die Substratzuleitung zerstört würden (vergl. Abb. 5.4 c). Das Gateoxid von nur wenigen nm Dicke muß hier als Ätzstopp ausreichen. In 6.4 wird im Zusammenhang mit der Herstellung der Siliziumstege ein solcher Ätzprozeß vorgestellt, der mit einer ICP-RIE und HBr-basierenden Plasmen arbeitet. Da der zweite Ätzschritt mit einem  $\text{HBr}/\text{O}_2$ -Plasma eine Selektivität von 150 hat, kann der Polyspacer definiert zurückgeätzt werden. Eine Überätzzeit von ca. 30 s – sowohl für p- als auch für n-Polysilizium – erzeugt die gewünschten Polysilizium-Spacer (Abb. 5.9).

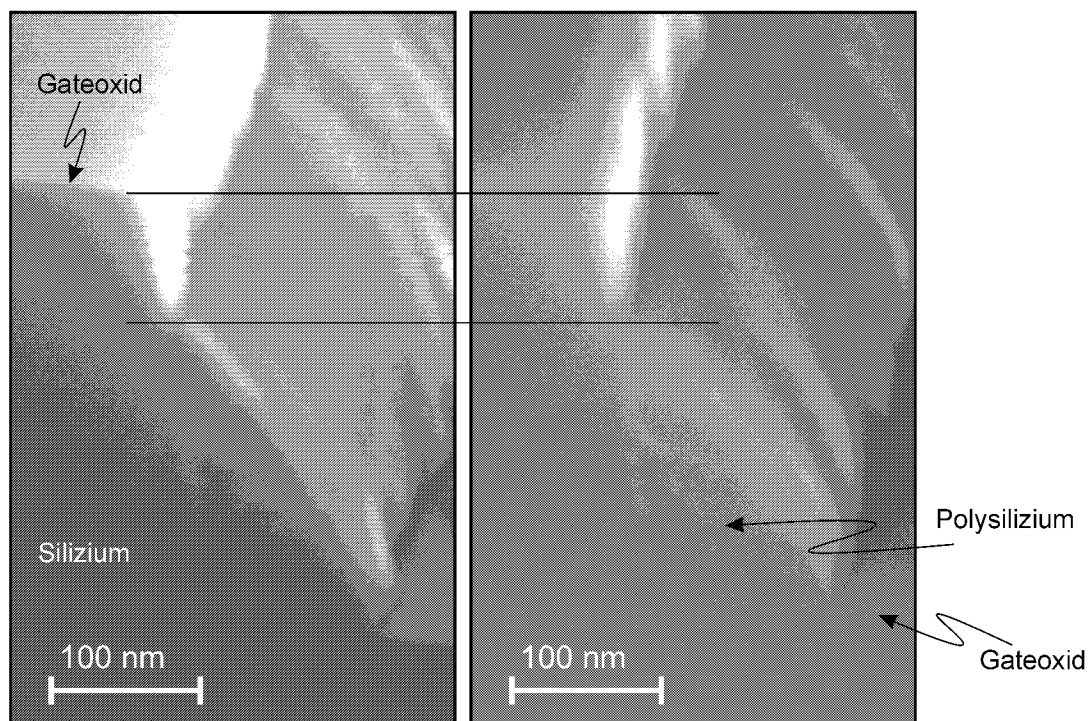


Abb. 5.9: REM-Aufnahmen von n-Polysilizium- (a) und p-Polysilizium-Spacern (b) an einer Siliziumkante. Die Rückätzung ist in beiden Fällen gleich. Das Gateoxid reicht als Ätzstopp aus.

### Topkontakt

Durch die Rückätzung des Gatespacers wird es möglich, den Topkontakt direkt auf dem Siliziumsteg anzubringen. Dazu ist zunächst eine Planarisierung notwendig, die durch die konforme Deposition eines dicken Deckoxids und anschließendem Polieren mittels CMP durchgeführt werden kann.<sup>3</sup> Diese Planarisierung ermöglicht das Ätzen

<sup>3</sup>Zum CMP siehe auch 4.10.

von Kontaktlöchern, die nur die obere Schicht des Steges öffnen, ohne daß dabei die Polyspacer oder das Substrat kontaktiert werden.

Für das Chemisch-Mechanische Polieren wurde eine „Einspindel-Planarisiermaschine P 200“ der Firma *Peter Wolters* verwendet. Die Anforderungen an den Prozeß waren: von einer 700 nm dicken TEOS-Schicht, die konform über die Strukturen deponiert wurde, über die mittleren 3" eine planare Schicht von mindestens 400 nm Dicke (ab Substrat) mit einer Variation von 30 nm zu erreichen, um für alle Transistoren noch einen genügend großen Abstand zwischen Topkontakt und Gate zu gewährleisten. Die eingesetzten TEOS-Schichten sind systematisch an den Rändern dicker als in der Mitte; bei einer 700 nm-Schicht sind dies etwa 50 nm im wichtigen 3"-Bereich. Der Prozeß mußte auf diese besondere Ausgangsbedingung abgestimmt werden und eine höhere Abtragung am Rand erreichen. Feste, unveränderliche Prozeßparameter zu bestimmen, erweist sich als schwierig, da auch der Zustand der Anlage in die momentane Schleifrate und Homogenität eingeht. So sinkt die Abtragsrate mit der Betriebszeit des Poliertuchs, und vom Zustand des Polierkopfpads hängt das Ausmaß der Rotationsinhomogenität ab, die auch in Abbildung 5.10 sichtbar ist. Genaue Prozeßparameter müssen also jeweils durch Tests unmittelbar vor der Anwendung bestimmt werden. Die Arbeitspunkte für die oben definierte Anwendung wurden in folgenden Bereichen gefunden: Anpreßdruck von 350–425 N, Rückseitendruck von 40–50 kPa, Polierkopffrotation von 130–140 rpm, Polierdauer von 55–70 s. Eine Dickendifferenz von ca. 30 nm ist dabei zu erreichen.

Nach der Planarisierung ist eine genaue Dickenmessung des restlichen Oxids über den Wafer notwendig. Der Vergleich mit den Steghöhen ergibt die zu ätzende Tiefe der Kontaktlöcher über den Stegen. Auf diesen Strukturen können keine Messungen und auch keine Endpunktkontrolle während der RIE gemacht werden. Größere Kontrollpads dagegen verhalten sich während des CMP anders als Nanostrukturen.

Kleine ohmsche Widerstände an den Kontakten sind besonders bei den äußerst kleinen Kontaktflächen wichtig. Bewährt hat sich die Positionierung von Siliziden zwischen Halbleiter und Metall, um die Potentialbarriere zwischen Halbleiter und Metall gering zu halten. Abbildung 5.11 zeigt in einer REM-Aufnahme, daß die Kontaktierung der Siliziumstege mit dieser Prozeßfolge erfolgreich ist.

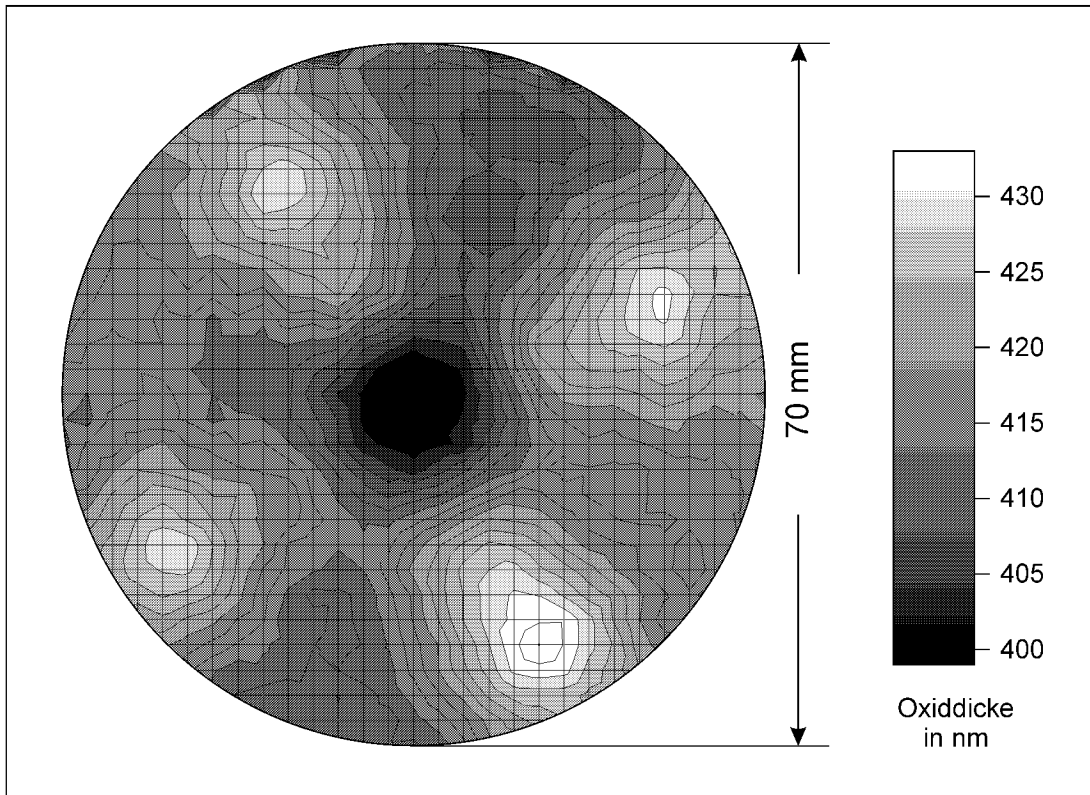


Abb. 5.10: Ellipsometermapping eines polierten Oxids. Dargestellt ist der wichtige mittlere Bereich eines 100 mm-Wafers. Eine TEOS-Schicht (640 nm in der Mitte, 690 nm am Rand) wurde mit 425 N Anpreßdruck und 130 rpm Polierkopfdrehzahl für 55 s poliert. Die Rotationsasymmetrie ist eine Folge der Löcher im Polierkopfpad, durch die der Wafer beim Transport mit Unterdruck festgehalten wird. Sie kann durch einen geeigneten Rückseiten-druck in Grenzen gehalten werden. Die Höhenliniendifferenz beträgt 2 nm, die Kästchenweite 2,5 mm.

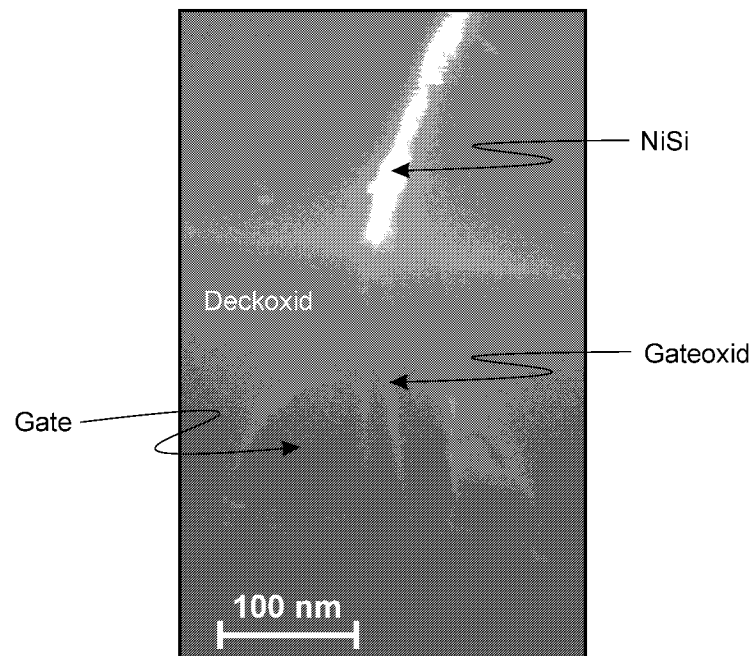


Abb. 5.11: REM-Querschnittsaufnahme des aktiven Bereichs des Transistors vor der Metallisation. Das Nickelmonosilizid (NiSi) auf dem Siliziumsteg ist durchgängig vorhanden, der Kontakt ist also hergestellt. Der Siliziumsteg ist 25 nm breit.

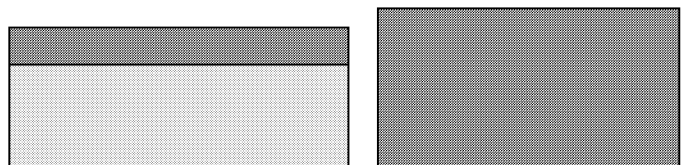
## 5.3 EpiNovaFET

Im Unterschied zur zweiten Variante wird die Schichtfolge p-n-p (bzw. n-p-n) hier durch Epitaxie (mit der Dotierung des Kanalgebietes) und Implantation (Gegendotierung der obersten Schicht) erzeugt. Dies vermeidet die flachen Leitungskanäle zu den Substratkontakten.

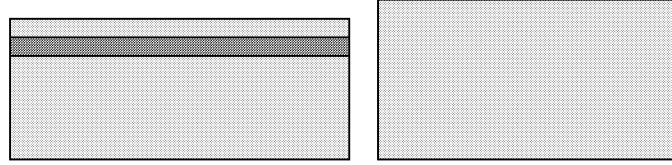
### 5.3.1 Prozeßfolge

Die Prozessierung des Transistors wird hier qualitativ beschrieben. Das genaue Prozeßprotokoll ist im Anhang A.3 dokumentiert.

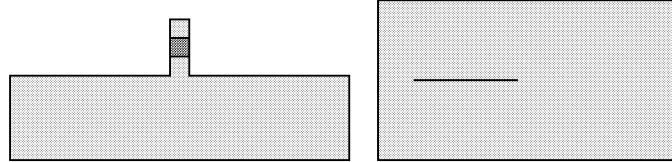
Epitaxie von n-Silizium auf hochdotiertem p-Silizium-Substrat (bzw. p-Silizium auf n-Substrat für n-Kanal-MOSFETs).



Gegendotierung der oberen Schicht durch Implantation mit Bor (bzw. Arsen für n-Kanal-MOSFET) , anschließend erstes Annealing.



E-Beam-Lithographie und RIE-Ätzen der Stege.



Weiter wie bei I-NovaFET ab „Deposition des ersten Deckoxids“.

### 5.3.2 Schlüsselprozesse

Außer der Dotierfolge sind alle Prozesse mit denen der Implantationsvariante identisch.

#### Dotierfolge

Bei dieser Variante ist das Substrat hochdotiert ( $\sim 10^{19} \text{ cm}^{-3}$ ) in der Dotierung des Substratkontakts zu wählen. Darauf wird eine ca. 300 nm dicke, niedrig entgegendotierte Siliziumschicht ( $\sim 10^{17} \text{ cm}^{-3}$ ) epitaktisch aufgewachsen.<sup>4</sup> Großflächig wird die Implantation für den Topkontakt nun durchgeführt, bevor dann die Siliziumstege hergestellt werden. Die in 5.2.2 beschriebene Implantation der Nanostrukturen und die Oxidation der Schutzspacer an den Stegen entfällt dabei.

<sup>4</sup>Wünschenswert wäre eine undotierte Schicht, aber um der Hintergrunddotierung durch die MBE zu begegnen wurde eine niedrige Dotierung gewählt.

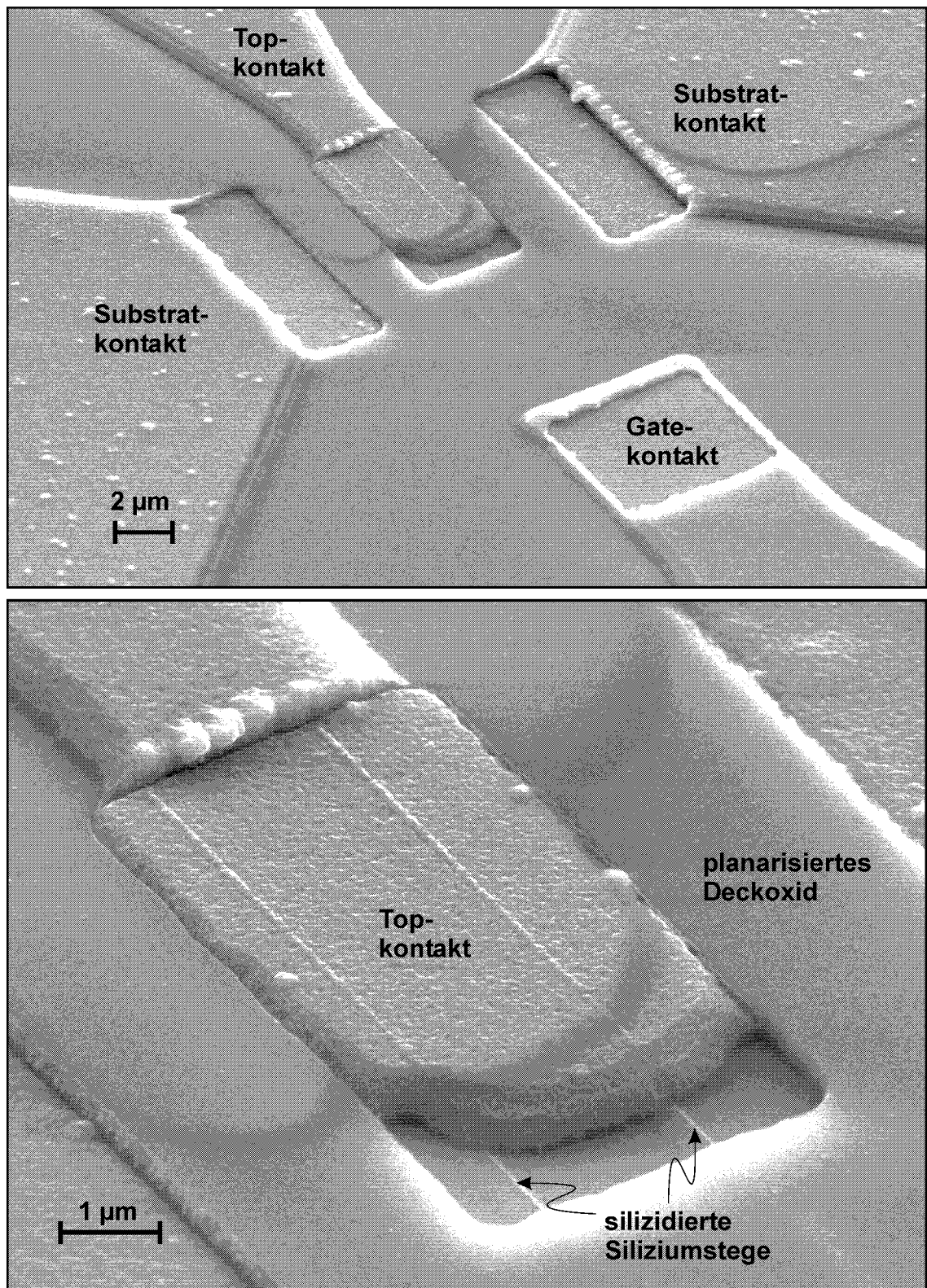


Abb. 5.12: REM-Bild des EpiNovaFETs als 2-Finger-Transistor mit jeweils 10 µm aktiver Steglänge.



# Kapitel 6

## Nanostrukturierung

Ein großer Entwicklungsaufwand wird betrieben, um die Auflösung der optischen Lithographie zu verbessern und auch zukünftige, kleinere Bauelemente mit parallelen Abbildungsverfahren herstellen zu können. Für die Entwicklung und Erforschung von Demonstratoren jedoch ist Elektronenstrahlolithographie (E-Beam-Schreiben) momentan die weitaus wichtigste lithographische Technik, die zu deren Herstellung verwendet wird. Aber auch mit E-Beam-Schreiben ist die Strukturierung im Bereich von unter 100 nm schwierig, insbesondere, wenn ein hohes Aspektverhältnis gefordert ist.

Die Herstellung von Nanostrukturen setzt sich aus den Komponenten „Strukturdefinition“ und „Strukturübertragung“ zusammen. Beide sind in dieser Größenordnung anspruchsvoll und voneinander abhängig. So sollte für eine hohe Auflösung und geringe Linienbreite die Resistdicke minimal sein. Andererseits muß der Resist der Ätzung bei der Strukturübertragung widerstehen und somit eine Mindestdicke aufweisen.

Zwei Negativresists sind bekannt, deren Auflösung den Anforderungen dieses Transistors gerecht werden: Hydrogen Silsesquioxan (HSQ), ein anorganisches Molekül auf Siliziumbasis, und Calixarene, eine Gruppe von organischen cyclischen Oligomeren.<sup>1</sup> Von Calixarenen ist 1996 zuerst berichtet worden, daß sie als hochauflösender Resist verwendet werden können [30], HSQ ist 1998 erstmals als Resist erwähnt [31].

Für die Anwendung im vertikalen Transistor fiel die Wahl auf HSQ, da ein anorganischer Resist beim Reaktiven Ionen Ätzen bessere Eigenschaften erwarten läßt. Um mit Calixarenen als Maske 300 nm Silizium zu ätzen, müßte die Methode des Dreilagigen-Resists angewendet werden (vergl. [32]). HSQ dagegen kann im Zweilagigen-Resist oder im Direktätzen eingesetzt werden.

---

<sup>1</sup>Polymere und Oligomere sind Moleküle mit einem regelmäßigen, periodisch aufgebauten Molekülskelett oder „Rückgrat“, an dem verschiedene oder gleiche Seitenketten in einer bestimmten Reihenfolge, welche „Sequenz“ genannt wird, angebracht sind. Polymere und Oligomere unterscheiden sich nur in der Zahl der Monomeren, aus denen sie aufgebaut sind. Im allgemeinen wird die Grenze etwa bei fünfzig Bausteinen gezogen.

## 6.1 Eigenschaften des anorganischen Moleküls HSQ

Silsesquioxane sind oligomere oder polymere Moleküle mit der stöchiometrischen Formel  $(\text{RSiO}_{3/2})_{2n}$ , wobei R für einen organischen Rest steht. Besteht dieser Rest allerdings nur aus Wasserstoff, so ergibt sich  $(\text{HSiO}_{3/2})_{2n}$ , das als Hydrogen Silsesquioxan bezeichnet wird. In der Halbleiter-Technologie werden Silsesquioxane als aufschleuderbare Oxide bereits länger verwendet [33, 34] auf Grund der Ähnlichkeit zu  $\text{SiO}_2$  nach einer Temperung. Bei der *Dow Corning Corporation*, die HSQ gelöst in Methylisobutylketon (MIBK) anbietet, ist die Produktbezeichnung *FOx-1x* auf diese Anwendung als *flowable oxide* zurückzuführen. Für diese Arbeit wurde *FOx-12* verwendet, die dünnste Lösung von HSQ bei *Dow Corning*.

HSQ tritt in verschiedenartigen Strukturformeln auf [35]: Ungeordnete Struktur, Leiterstruktur, Käfigstruktur und partielle Käfigstruktur (Abb. 6.1). Im Resist *FOx-12* liegt eine Mischung dieser Strukturen vor.

Beim HSQ-Film werden durch Energiezufuhr – entweder vollständig durch Temperung für die Anwendung als Isolator oder lokal durch E-Beam-Schreiben – die Wasserstoffverbindungen aufgebrochen (Abb. 6.2 a). Über einen chemischen Prozeß kann jeweils zwischen zwei nun nicht abgesättigten Siliziumatomen eine neue Sauerstoffverbindung hergestellt werden (Abb. 6.2 b+c). Diese Vernetzung geschieht dreidimensional und führt zu einer amorphen  $\text{SiO}_2$ -ähnlichen Struktur mit der allgemeinen Summenformel  $\text{H}_x\text{Si}_y\text{O}_z$  [36].

*FOx-12* ergibt beim Aufschleudern nahezu unabhängig von der Drehzahl eine Schichtdicke von ca. 210 nm. Weitere Verdünnungen mit MIBK wurden hergestellt, um mit minimalen Schichtdicken die höchstmögliche Auflösung zu erreichen (Abb. 6.3). Die Dicken wurden mit dem Ellipsometer bestimmt; dabei ist der Brechungsindex von HSQ nach dem Softbake als  $n = 1,40$  bestimmt worden. Dies ist etwas geringer als thermisches  $\text{SiO}_2$  ( $n = 1,46$ ) oder TEOS-Oxid ( $n = 1,43$ ). Unterschiedliche Eigenschaften zwischen HSQ nach dem Softbake und HSQ nach dem Entwickeln konnten mit dem Ellipsometer nicht festgestellt werden.

Die Handhabung von HSQ ist sensibel, da es gekühlt bei  $-5^\circ\text{C}$  bis  $+5^\circ\text{C}$  aufbewahrt werden muß und auch unbelichtet zur Polymerisation neigt. Insbesondere reagiert es mit Luftfeuchtigkeit (weshalb es innerhalb von wenigen Minuten nach der Entnahme aufgeschleudert werden sollte) und mit Bestandteilen aus Glas oder Verunreinigungen im Lösungsmittel [37]. Deshalb wurde HSQ immer in HDPE<sup>2</sup>-Flaschen aufbewahrt und die Verdünnung unmittelbar vor der Nutzung in kleinen Mengen angesetzt.

## 6.2 Strukturdefinition mit HSQ

Ziel ist die reproduzierbare Herstellung von dünnsten Linien. Aufgrund erster Ätztests wurde eine Ausgangsschichtdicke von 100–120 nm angestrebt, die nach Abb. 6.3 durch eine Verdünnung von 1 + 0,75 (*FOx-12* + MIBK) bei 4000 rpm erreicht

<sup>2</sup>HDPE bedeutet *High Density Polyethylen*, also Polyethylen hoher Dichte.

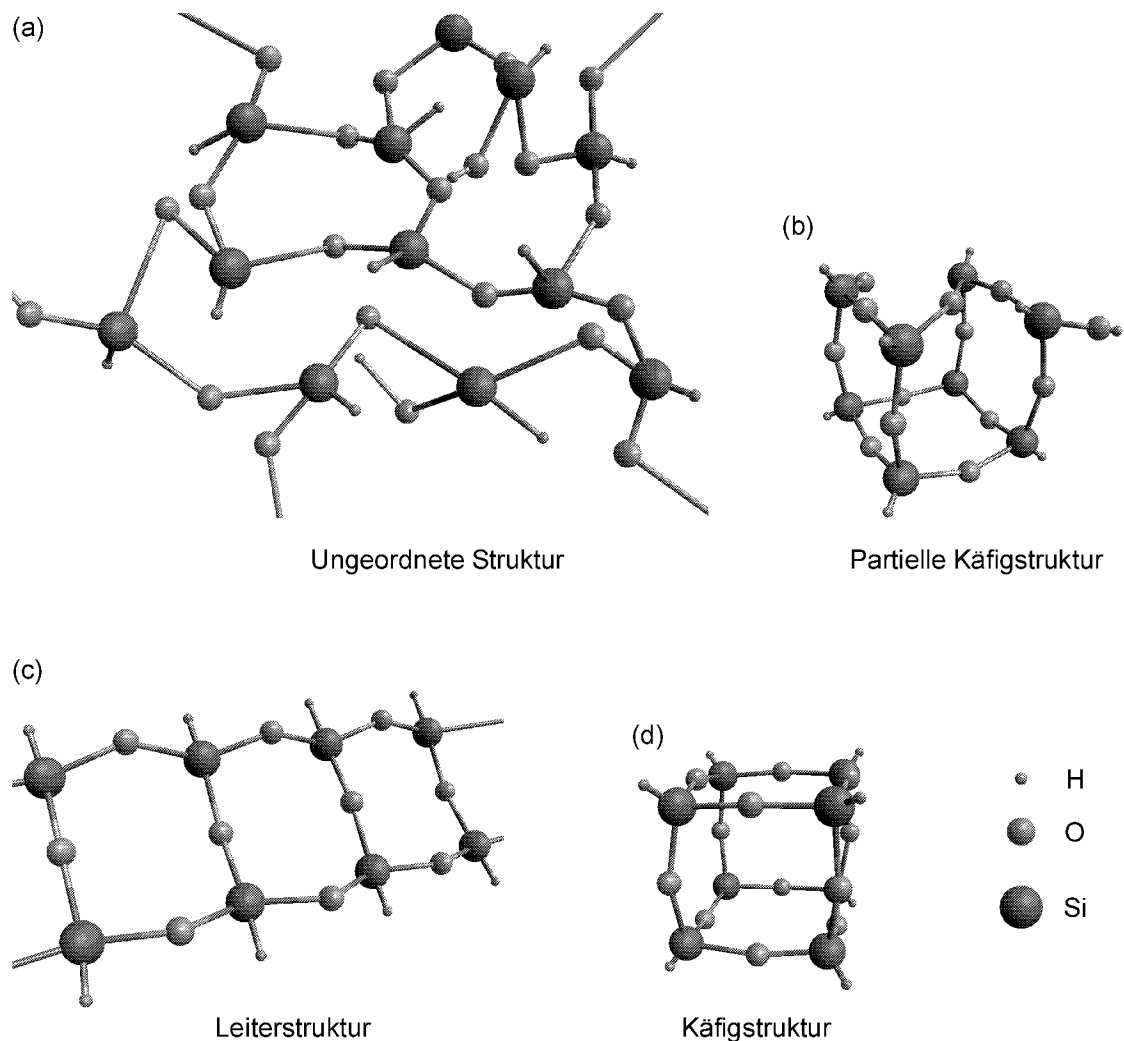


Abb. 6.1: Strukturformeln von HSQ. Als Käfigstruktur ist die einfachste Form dargestellt ( $n=4$ ). Darstellung nach [35].

werden kann. Mit diesen ca. 115 nm dicken HSQ-Schichten wurden sämtliche hier vorgestellten Ergebnisse erzielt. Für die Technik des Elektronenstrahlschreibens und die Erklärung des dabei verwendeten Vokabulars siehe 4.1.1.

Eine Testmaske für die Elektronenstrahlolithographie enthielt Linien verschiedener Breite: sogenannte *Four-Pass-Single-Lines* mit nominell 5 nm Linienbreite sowie Linien mit nominell 25 nm, 37,5 nm, 50 nm, 62,5 nm und 75 nm Breite. Bei den *Single-Lines* wird die Struktur in vier Zügen sukzessive übereinander belichtet, wobei die Breite der Linie lediglich durch den Strahldurchmesser definiert ist – hier 15 nm, dem kleinstmöglichen Strahldurchmesser. Die nominellen Linienbreiten leiten sich aus der CAD-Struktur ab, die mit dem Elektronenstrahl rasterförmig auf die Probe geschrieben wird. Diese Rasterschrittweite (*beam step size*) betrug für die *Single-Lines* 5 nm, für die anderen Linien 12,5 nm. Naturgemäß sind die wirklichen, geschriebenen Linienbreiten größer als die nominellen (siehe hierzu auch Abb. 4.1). Für die Beschreibung wurde dennoch die Bezeichnung „geschriebene Linienbreite“

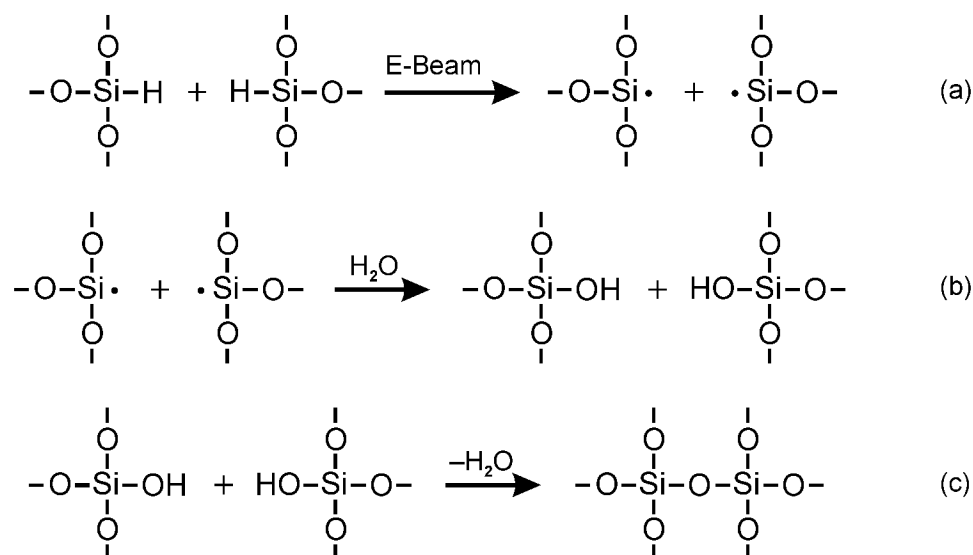


Abb. 6.2: Vernetzung von HSQ beim E-Beam-Prozeß (nach [31]): (a) Die (schwachen) Wasserstoffbindungen werden beim E-Beam-Schreiben aufgebrochen. (b) Durch Feuchtigkeit bilden sich an diesen Stellen Silanole (SiOH-Gruppen). (c) Unter Abgabe von  $\text{H}_2\text{O}$  vernetzen sich zwei Si-Atome über eine Sauerstoffbindung.

(für 25–75 nm-Linien) bzw. *Single-Line* (für die 5 nm-Linien) gewählt.

Nachdem die übrigen Parameter eingestellt waren, blieb die Bestimmung der optimalen Dosis. Da die Dosis einer Struktur wegen des Proximityeffektes sehr von ihrer Form und dem Untergrund abhängt, mußte für nahezu jede Linienbreite die richtige Einstellung gefunden werden. Keinen merklichen Einfluß dagegen hat die Anwesenheit eines 400 nm dicken ausgehärteten organischen Resists unter dem HSQ, sofern das Substrat unverändert Silizium ist. Diese beiden Schichtsysteme – HSQ direkt auf Silizium sowie auf ausgehärtetem Lack – wurden für verschiedene Arten der Strukturübertragung benötigt (siehe Abschnitte 6.3 und 6.4).

Aus Dosisreihen für verschiedene Linienbreiten wurden die Dosen für gleichförmige Linien bei optimalen Profile bestimmt. Unterbelichtete Linien sind – bei nicht voller Resisthöhe – nicht durchgehend oder unregelmäßig, überbelichtete Linien sind insbesondere am Fuß signifikant zu breit. In Abbildung 6.4 sind Beispiele dieser Eigenschaften gezeigt.

In Abbildung 6.5 sind für einige geschriebene Linienbreiten die gemessenen Breiten oben und unten an der Linie für verschiedene Dosen aufgetragen. Abbildung 6.6 zeigt in REM-Bildern die Querschnitte einiger optimierter HSQ-Linien. Die besten Linieneigenschaften wurden erreicht bei:

geschr. Linienbreite:	Single-Line	25 nm	37,5 nm	50 nm	75 nm
Dosis ( $\mu\text{C}/\text{cm}^2$ ):	7500	5000	4250	3500	3000

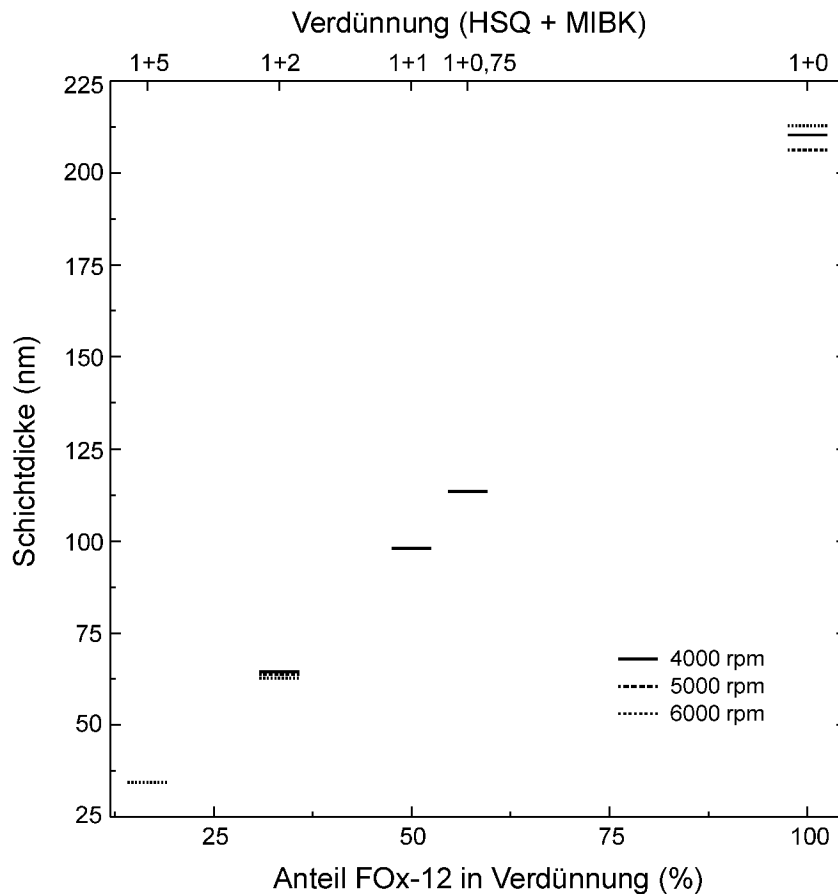


Abb. 6.3: Schichtdicken von HSQ-Verdünnungen. Für verschiedene Verdünnungen und Lackschleuderprogramme wurden die Schichtdicken nach dem Softbake mit dem Ellipsometer bestimmt. Die erzielten Dicken sind weitestgehend linear abhängig vom Lösungsmittelanteil und nahezu unabhängig von der Aufschleuderdrehzahl. Die Länge der Striche entspricht der Fehlerabschätzung bei der Verdünnung.

## 6.3 Strukturübertragung mit Zweilagelack

Der Transistor erfordert eine hochgradig anisotrope Ätzung, d.h. Unterätzungen oder Ablagerungen können nur in sehr geringem Maße toleriert werden. Daher kommt nur Trockenätzen in Betracht, und da insbesondere Reaktives Ionen Ätzen. Bei der ersten verwendeten RIE, einer Standard-RIE, ist ein sehr anisotroper Prozeß zum Siliziumätzen etabliert, bei dem ein Ar/SF<sub>6</sub>-Plasmagemisch zur Anwendung kommt. Die Selektivität von Silizium zu Resist ist auf Grund des hohen Sputteranteils jedoch so klein, daß die Ätzmaske dicker als die gewünschte Ätztiefe sein muß. Für die Anwendung in den vertikalen Transistoren sollte die Ätztiefe mindestens 300 nm betragen, allerdings kann ein E-Beam-Resist für eine vorgegebene Auflösung nicht beliebig dick sein. Die angestrebten Mindestlinienbreiten von 20–30 nm konnten hier mit HSQ der Dicke 115 nm erzeugt werden, geringere Linienbreiten erfordern jedoch dünneren Resist [38]. Eine Methode, diese Strukturen dennoch in Silizium zu übertragen ist die des Mehrlagenresists. Dabei wird ausgenutzt, daß (ausgehärteter)

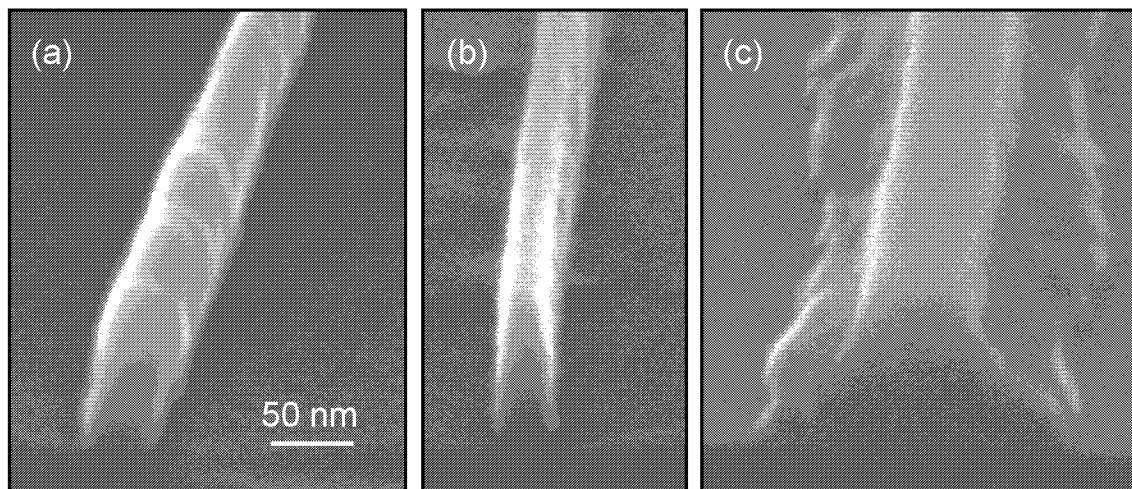


Abb. 6.4: *Linienprofile von HSQ bei (a) unterbelichteten, (b) richtig belichteten und (c) überbelichteten Strukturen. Unterbelichtete Strukturen sind unregelmäßig und haben nicht die volle Resisthöhe. Überbelichtete Strukturen bilden einen breiten, unförmigen Fuß.*

organischer Resist in Sauerstoffplasma anisotrop geätzt werden kann, andere Materialien dabei aber kaum angegriffen werden. So genügt es, eine dünne anorganische Schicht zu strukturieren, um eine dicke organische Ätzmaske zu erzeugen. HSQ als anorganischer Resist kann direkt lithographiert werden, so daß ein sog. Zweilagen- oder „Bilayer“-Resist entsteht (siehe Abb. 6.7).

Verwendet wurde ein Schichtsystem aus AZ 5204 und HSQ. Der 400 nm dicke AZ-Lack wurde eine Stunde bei 250°C ausgehärtet, HSQ anschließend darauf aufgeschleudert und lithographiert (Abb. 6.8 a). Probleme ergaben sich trotz wohlbekannter Ätzparameter beider Plasmen, da beim Sauerstoffplasma eine Unterätzung von ungefähr 15 nm auftrat. Dieser, bei optisch lithographierten Strukturen kaum merkbare isotropische Anteil bewirkte nun, daß die geätzten Strukturen um 30 nm dünner waren als die HSQ-Linien. Was zunächst positiv klingt, war dennoch ein Problem, da die dünneren *Bilayer*-Linien instabil wurden (Abb. 6.8 b). Die schmalsten so hergestellten Siliziumlinien sind zwar nur 30 nm breit, allerdings mit äußerst rauen Seitenflächen (Abb. 6.8 c).

Gleichmäßige Siliziumstege ab 50 nm Breite ließen sich mit dieser Methode für den Einsatz in Transistoren reproduzierbar erzeugen. Dünnsste Linien, wie sie für das Double-Gate-Prinzip benötigt werden, waren hiermit nicht zu erreichen.

## 6.4 Strukturübertragung durch Direktätzen

Eine im Rahmen des Projektes *Extended Silicon MOSFET* angeschaffte neue RIE eröffnete neue Möglichkeiten. Es handelt sich um das *Plasmalab System 100* von *Oxford* mit ICP-Quelle<sup>3</sup>. Die ICP-Technologie in Verbindung mit HBr-basierenden Plasmen ermöglicht das anisotrope Ätzen von Silizium mit sehr hoher Selektivität zu

<sup>3</sup>Siehe hierzu auch 4.1.3.

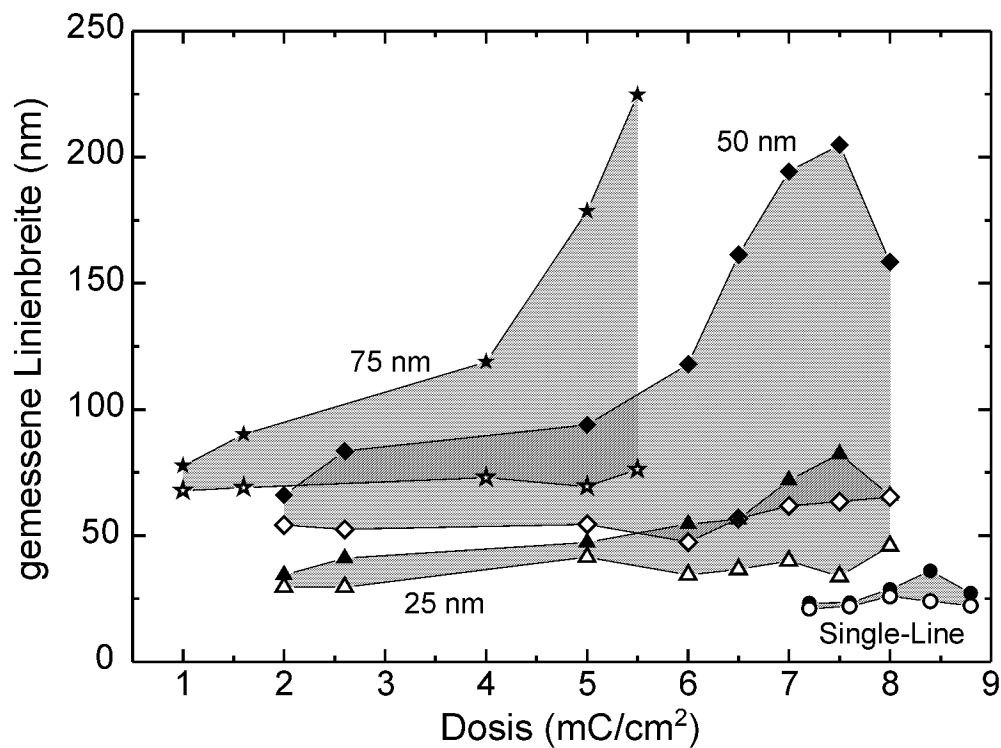


Abb. 6.5: Gemessene Linienbreiten von HSQ in Abhängigkeit der Dosis. Während die Linienbreite oben am Profil (offene Symbole) nahezu gleich bleibt, wird der Fuß der Struktur bei Überbelichtung deutlich breiter (geschlossene Symbole). Die Daten wurden aus Querschnitts-REM-Bildern der Strukturen bestimmt.

SiO<sub>2</sub>. Da HSQ nach der Entwicklung eine SiO<sub>2</sub>-ähnliche Struktur hat, können diese Eigenschaften für das Ätzen dünnster Siliziumstege mit HSQ als direkter Ätzmaske genutzt werden. Auf Grund der hohen Selektivitäten ist es notwendig, vor dem Ätzen des Siliziums das natürliche Oxid vom Substrat zu entfernen. Erreicht wird dies durch einen RIE-Schritt mit CHF<sub>3</sub> (30  $\mu$ bar, 20 ml/min, 300 W) für 5 Sekunden. Bei diesem Schritt wird gleichzeitig ein möglicher HSQ-Schleier am Fuß der Resistlinie entfernt, der sich ansonsten deutlich stärker als bei anderen Ätzprozessen auf das Profil des Siliziumsteiges auswirken würde. Für das Ätzen von Silizium mit der ICP-RIE sind diverse Parameter aufeinander abzustimmen: HBr-Fluß, Bias, ICP-Leistung und Druck. Bei reinen HBr-Plasmen ist die Selektivität zu TEOS-Oxid vom Prozeßdruck abhängig:

HBr-Fluß (sccm)	Bias (W)	ICP (W)	Druck (mTorr)	Si-Ätzrate (nm/min)	Selektivität
50	50	750	7	60	10
50	50	750	30	50	23
50	50	750	60	41	50

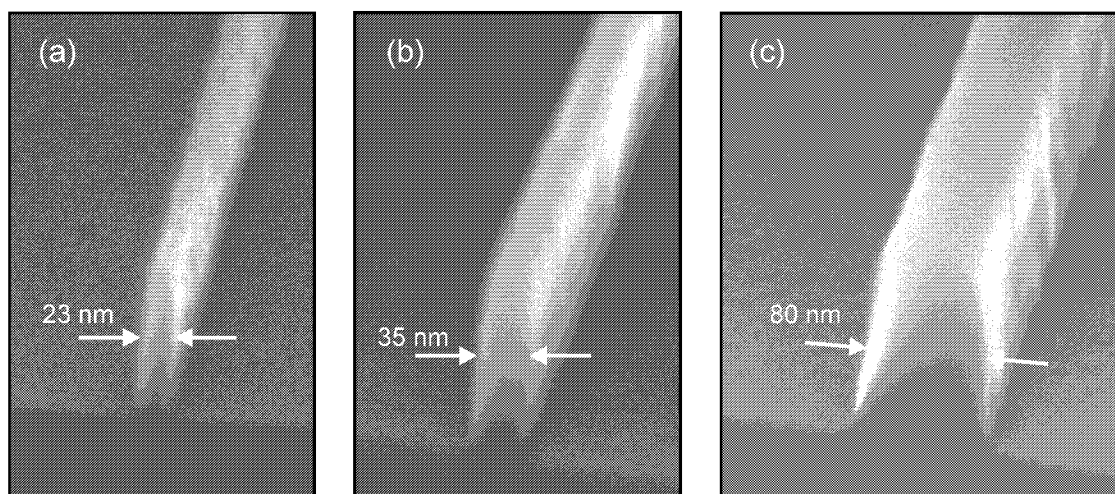


Abb. 6.6: REM-Aufnahmen optimierter HSQ-Querschnitte. (a) Single-Line, (b) 25 nm-Linie, (c) 75 nm-Linie.

Alle diese Prozesse sind weitgehend anisotrop. HSQ jedoch wird mit einer größeren Rate als  $\text{SiO}_2$  geätzt, so daß bessere Selektivitäten erwünscht sind. Erreicht werden können diese durch Hinzufügen von Sauerstoff. Die Reaktionsprodukte und die unpassivierte Siliziumoberfläche reagieren mit den Sauerstoffradikalen, so daß  $\text{SiO}_2$  deponiert wird. Das Gleichgewicht zwischen Deposition und Abtragung durch die physikalische Ätzkomponente muß über die Ätzparameter so eingestellt werden, daß Silizium vertikal geätzt wird, die Seitenwände aber nur passiviert und nicht verbreitert werden. Durch eine Beimischung von 2 sccm  $\text{O}_2$  und der Erhöhung der Bias auf 100 W sowie der ICP-Leistung auf 2000 W konnte dies erreicht werden. Die Ätzrate von TEOS-Oxid ist nun kaum noch meßbar und eine nur wenige Nanometer dicke Oxidschicht kann als Ätzstopp dienen; gleichzeitig sind die Seitenwände senkrecht. Die Ätzrate von Silizium ist dabei abhängig von der Bedeckung durch die Maske: Werden große Siliziumflächen geätzt, so stellt sich ein anderes Gleichgewicht ein als wenn große Flächen von  $\text{SiO}_2$  bedeckt sind. Bestimmt wurden die Ätzraten von ca. 110 nm/min bei einer geringen Ätzfläche und ca. 380 nm/min bei einer großen Ätzfläche.

Ein zweistufiger Prozeß mit beiden beschriebenen Plasmen bringt die besten Ergebnisse: Trotz vorherigem Lackschleierprozeß mit  $\text{CHF}_3$  sorgt der erste Ätzschritt mit reinen HBr-Plasma – bei niedrigem Druck – durch geringe Selektivität für einen Abtrag von möglichen Rückständen. Mit dem zweiten Schritt kann dann die gewünschte Äztiefe erreicht werden. Der optimierte Zweistufenprozeß kann also wie folgt zusammengefaßt werden:

Schritt	Ätzdauer (sek)	HBr-Fluß (sccm)	$\text{O}_2$ -Fluß (sccm)	Bias (W)	ICP (W)	Druck (mTorr)	Si-Ätzrate (nm/min)
1	60	50	–	50	750	7	60
2	variabel	50	2	100	2000	7	verschieden

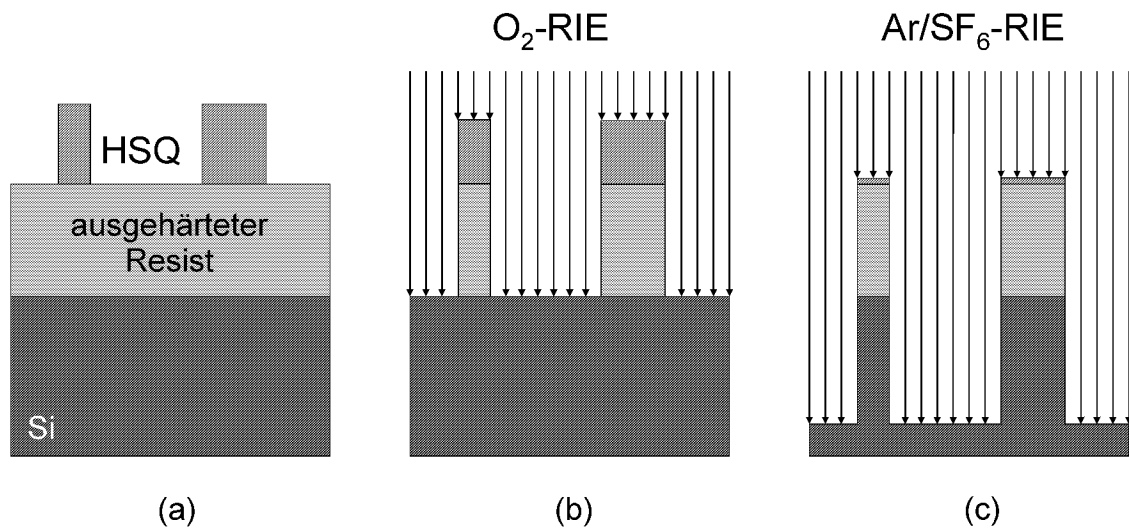


Abb. 6.7: Das Prinzip des Zweilagenlacks. Nach der Lithographie (a) dient HSQ als Maske beim anisotropen Ätzen von ausgehärtetem organischem Resist mit Sauerstoffplasma (b). Dieser Resiststapel ist nun Ätzmaske für das Siliziumätzen mit Ar/SF<sub>6</sub> (c).

Übertragen auf die Herstellung dünner Linien mit HSQ als Ätzmaske führt eine Zeit von 60 sek für den zweiten Prozeßschritt zu einer Ätztiefe von insgesamt ca. 330 nm; dabei wird etwa 60 nm HSQ abgetragen. Eine Verbesserung der Resisteigenschaften kann erreicht werden, indem der HSQ vor der Ätzung für eine Stunde bei mindestens 450°C getempert wird [39, 40]. Damit sind 25 nm breite Siliziumstege hergestellt worden, die den Anforderungen des vertikalen Transistors gerecht werden (siehe Abb. 6.9 a+b und [41]).

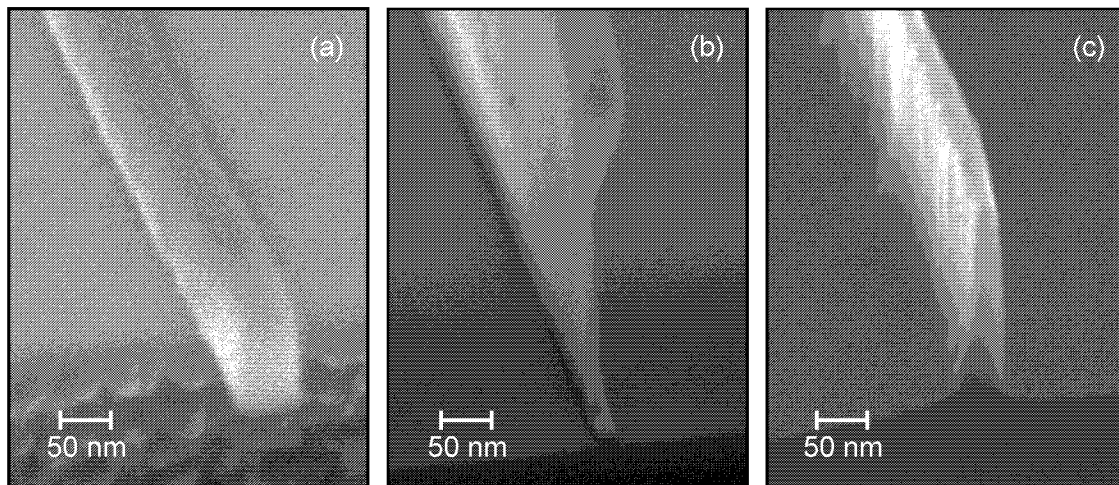


Abb. 6.8: Stationen des Zweilagprozesses: (a) HSQ-Linie auf ausgehärtetem Resist. (b) Bilayer-Steg auf Silizium. Der Fuß der Linie ist nur ca. 10 nm breit und instabil. Diese Linie ließ sich nicht mehr in Silizium übertragen. (c) Dünnschicht mit dieser Methode hergestellter Siliziumsteg (ca. 30 nm). Die Seitenflächen des Stegs und die Linienbreite sind zu unregelmäßig, als daß diese Linien im vertikalen Transistor verwendet werden könnten.

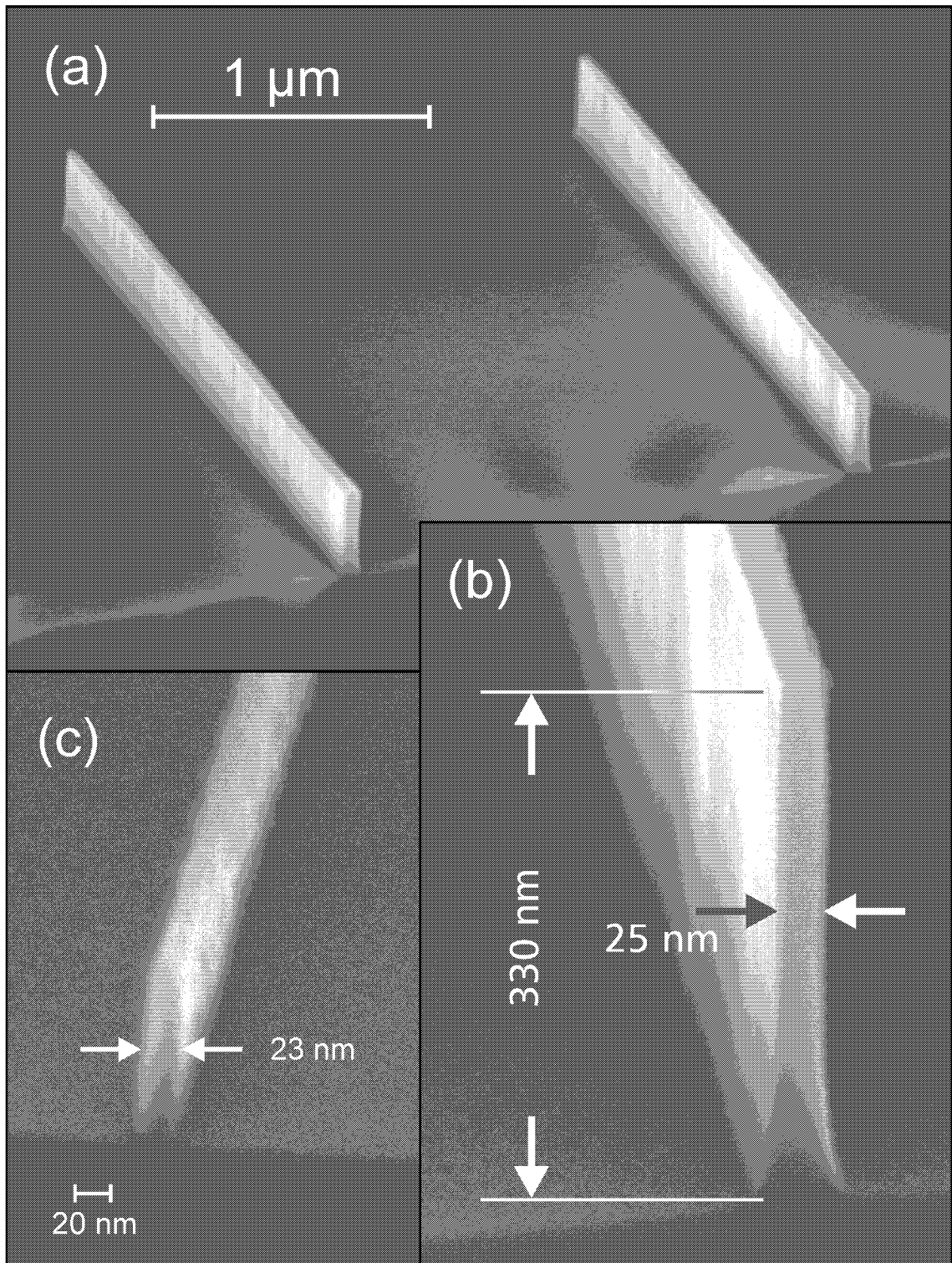


Abb. 6.9: REM-Querschnittsbilder: (a) Mit Direktätzen hergestellte Siliziumlinien sind sehr regelmäßig. (b) Dies sind die dünnsten Siliziumlinien, die reproduzierbar hergestellt werden können. Sie entsprechen genau den Anforderungen des vertikalen Double-Gate-MOSFETs. (c) HSQ-Single-Line, wie sie für die Linie in Bild (b) als Maske diente.



# Kapitel 7

## Bauelementcharakterisierung

Es sind drei verschiedene Layouts des NovaFETs realisiert worden: das ursprüngliche Layout (als p-Kanal-Transistor), die Implantationsvariante (p-Kanal- und n-Kanal-Transistor) und der EpiNovaFET (p-Kanal-Transistor). Auf einem Wafer sind jeweils verschiedene Stegbreiten realisiert, außerdem verschiedene Gateweiten (Steglängen) und Anzahl der parallel geschalteten Stege.

### 7.1 NovaFET (ursprüngliches Layout)

Dieses Transistorlayout hat zwei Schwächen gegenüber den beiden anderen, die sich auch in den Kennlinien widerspiegeln: die Asymmetrie in der Dotierfolge und die hohen Zuleitungswiderstände im Topkontakt (vergl. 5.1.2 und 5.1.2). Abbildung 7.1 zeigt zwei Ausgangskennlinienfelder eines 4-Finger-Transistors mit einer aktiven Länge von je  $3\mu\text{m}$ . Es ist deutlich ein Unterschied zwischen *Source on Bottom* (SOB) und *Source on Top* (SOT) zu erkennen: Für SOB werden eine Steilheit<sup>1</sup> von  $g_m = 82\mu\text{S}/\mu\text{m}$ , eine Schwellenspannungssteigung von  $S = 180\text{mV/dec}$  und eine Schwellenspannung von  $U_{th} = -0,7\text{V}$  gemessen. SOT liefert  $g_m = 22\mu\text{S}/\mu\text{m}$ ,  $S = 188\text{mV/dec}$  und  $U_{th} = -0,95\text{V}$ . Dabei zeigt SOT ein gutes Sättigungsverhalten, während bei SOB keine Sättigung eintritt und ab  $U_{DS} \approx 1,75\text{V}$  bereits *Punch-Through* zu beobachten ist. Beide Konfigurationen zeigen diodenförmiges Anlaufverhalten, welches sich in dem parabelförmigen linearen Bereich des Kennlinienfeldes zeigt. Der Widerstand der oberen Zuleitung zwischen Pad und Kanal kann abgeschätzt werden: Die Dotierung von  $4\text{--}5 \cdot 10^{18}\text{at cm}^{-3}$  führt zu einem spezifischen Widerstand von  $\rho \approx 2 \cdot 10^{-2}\Omega\text{cm}$  [12]. Mit  $R = \rho \cdot l/A$ , wobei die mittlere Länge  $l$  hier  $2,5\mu\text{m}$  ist und die Querschnittsfläche  $A \approx 100 \times 140\text{nm}^2$  beträgt,<sup>2</sup> kann der Gesamtwiderstand des Drahtes zu mehreren  $10\text{k}\Omega$  angenommen werden.

Dieser sehr große Widerstand ist für SOB ein Drainwiderstand  $R_D$  und für SOT

---

<sup>1</sup>Sämtliche hier angegebene Steilheiten sind extrinsisch, d. h. gemessene Werte mit sämtlichen Zuleitungen.

<sup>2</sup>Verlässliche Angaben über die Linienbreiten können nicht gemacht werden, da einerseits diese Linien mittels Dreilagelack und UVN 30 hergestellt wurden, ein Verfahren, das keine befriedigende Auflösung hatte, und andererseits, da wegen des Proximityeffektes die Teststrukturen systematisch dünner waren als die Transistorlinien am Pad.

ein Sourcewiderstand  $R_S$ . In SOT wird die effektive Gatespannung reduziert auf  $U'_G = U_G - I_{Dsat} \cdot R_S$  [15] – und dadurch die Steuerwirkung verringert. Damit können die deutlich geringere Steilheit und die größere Schwellenspannung erklärt werden. Die fehlende Sättigung und der *Punch-Through* gehen auf die unterschiedlichen Dotierverläufe der oberen und unteren p-n-Übergänge zurück. Für die Abhängigkeit der Dicke der Raumladungszone von der Dotierung gilt [42]:

$$t_n = \left( \frac{2\epsilon\epsilon_0 U}{q} \cdot \frac{N_A/N_D}{N_A + N_D} \right)^{1/2}$$

$$t_p = \left( \frac{2\epsilon\epsilon_0 U}{q} \cdot \frac{N_D/N_A}{N_A + N_D} \right)^{1/2}$$

mit  $U$  der anliegenden Spannung,  $N_A, N_D$  die Akzeptor- bzw. Donatorkonzentration im p- bzw. n-Gebiet.

Da die Dotierung in der oberen Schicht größer und der Übergang schärfer ist, weitet sich die Raumladungszone oben tiefer in den Kanal aus als unten. Da der *Punch-Through* wesentlich auf die Ausweitung der Raumladungszone am Drain zurückzuführen ist, tritt er früher auf, wenn das Drain oben ist, also für SOB-Konfiguration des NovaFETs.

In Abbildung 7.2 sind die Kennlinien für verschiedene Gateweiten, also Längen des aktiven Bereiches des Transistors, gegenübergestellt. Da die Kennlinien auf die Gateweite normiert sind, dürften keine signifikanten Unterschiede erkennbar sein. Größere Widerstände mit steigender Gateweite jedoch bewirken eine deutliche Verringerung des Drainstroms und der Steilheit.

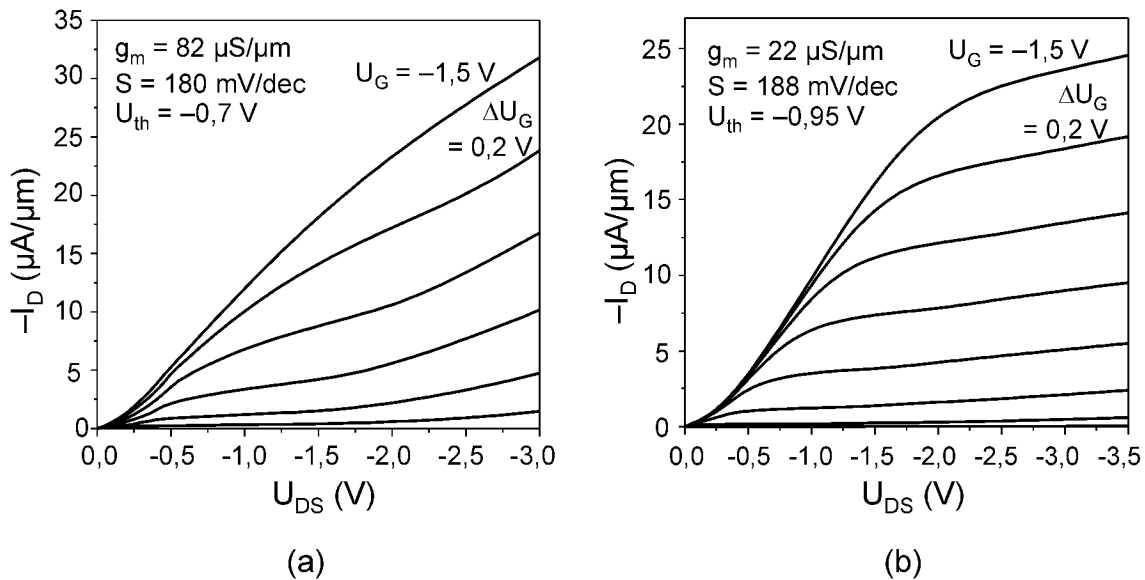


Abb. 7.1: NovaFET (ursprüngliches Layout): Ausgangskennlinienfelder in (a) SOB- und (b) SOT-Konfiguration eines 4-Finger-Transistors mit je  $3 \mu\text{m}$  aktiver Länge. Die Gate-oxiddicke beträgt  $6 \text{ nm}$ , die Kanallänge  $140 \text{ nm}$ .

Mit diesen Transistoren konnte jedoch gezeigt werden, daß dieser Ansatz des vertikalen Double-Gate-MOSFETs grundsätzlich funktioniert.

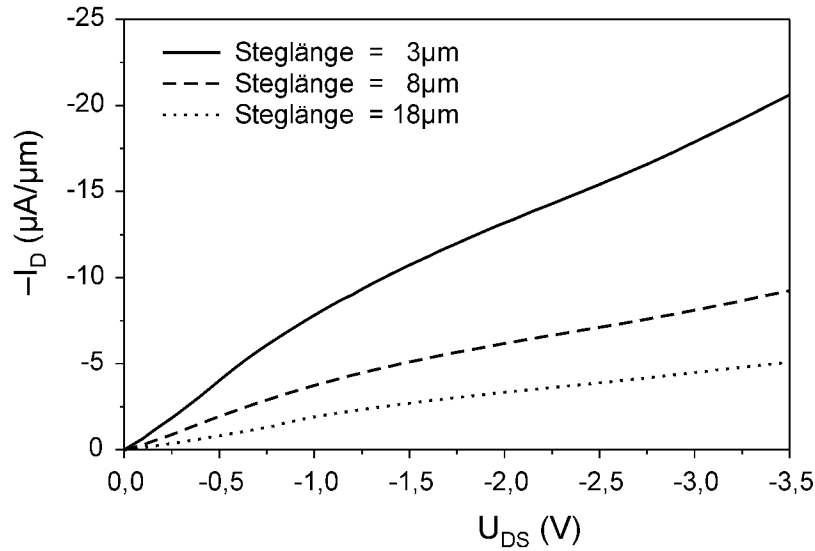


Abb. 7.2: NovaFET (ursprüngliches Layout): Ausgangskennlinien bei  $U_G = 2V$  für verschiedene Gateweiten. Wegen des höheren Widerstandes zeigen Transistoren mit größeren Gateweiten schlechteres Steuerverhalten. Im optimalen Fall müßten die normierten Drainströme identisch sein.

## 7.2 I-NovaFET

Transistoren der Implantationsvariante wurden mit einer Gateoxiddicke von  $t_{ox} = 3,7$  nm hergestellt. Die Kanallänge beträgt 75 nm (p-Kanal-) bzw. 130 nm (n-Kanal-Transistor). Abbildung 7.3 zeigt Ausgangs- und Transferkennlinienfeld eines p-Kanal-MOSFETs in SOT-Konfiguration. Die gegenüber dem ursprünglichen Layout verbesserte Steilheit von jetzt  $150 \mu S/\mu m$  und die geringere Unterschwellenspannungssteigung von 80 mV/dec sind im wesentlichen auf die verkleinerten Zuleitungswiderstände zurückzuführen. DIBL ist zu 30 mV/V, die Schwellenspannung zu  $U_{th} = 0,24$  V bestimmt worden. Für SOB wurden dagegen deutlich abfallenden Werte ( $g_m = 90 \mu S/\mu m$ ,  $S = 133$  mV/dec,  $DIBL = 120$  mV/V und  $U_{th} = 0,27$  V) bestimmt.

SOB- und SOT-Messungen vergleichbarer n-Kanal-Transistoren sind in Abbildung 7.4 gegenübergestellt. In beiden Konfigurationen sind  $S = 80$  mV/dec und  $U_{th} = -0,17$  V gleich, Steilheit und DIBL dagegen sind für SOT ( $g_m = 195 \mu S/\mu m$ ,  $DIBL = 32$  mV/V) besser als für SOB ( $g_m = 145 \mu S/\mu m$ ,  $DIBL = 55$  mV/V). Alle Transistoren dieses Typs haben sehr gute Sperreigenschaften. Die Sperrströme liegen im Bereich von  $1$  pA/ $\mu m$  und teilweise unter der Auflösungsgrenze des Meßgerätes.

Die Werte für die Schwellenspannung sind nicht auf den geeigneten Arbeitspunkt optimiert. So sollte für p-Kanal-Transistoren  $U_{th} \approx 0,1$  V und für n-Kanal-Transistoren  $U_{th} \approx -0,1$  V sein. Durch die Wahl geeigneter Gatematerialien kann die Schwellenspannung jedoch eingestellt werden. Die Verbesserungen in Steilheit

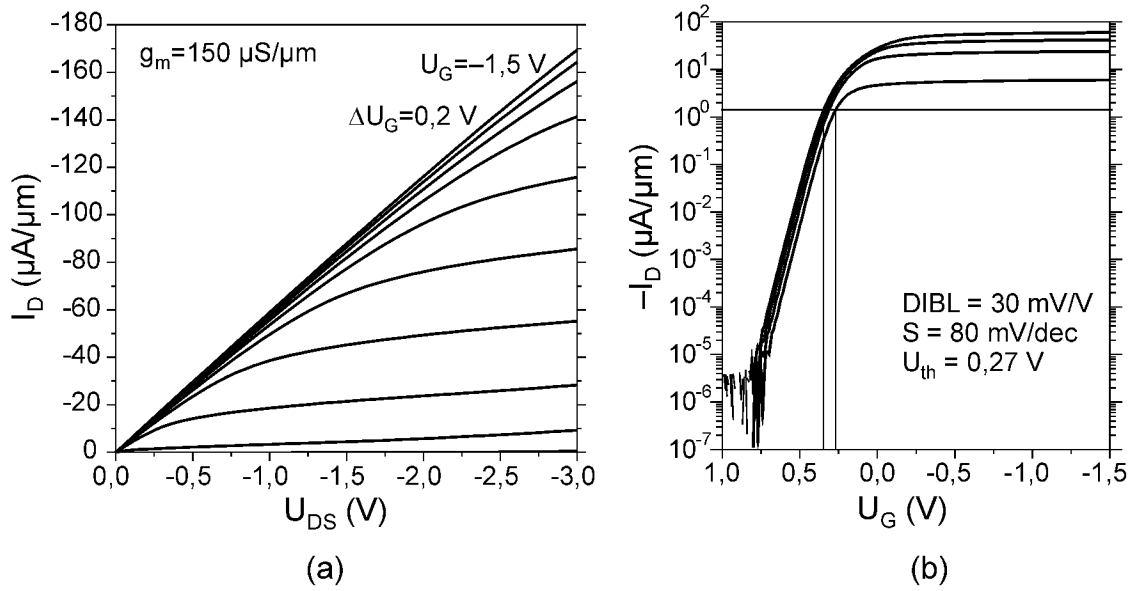


Abb. 7.3: *p*-Kanal-I-NovaFET: (a) Ausgangs- und (b) Transferkennlinienfeld. Die Kanaldotierung ist  $< 10^{13} \text{ cm}^{-3}$ .

und Unterschwellenspannungssteigung sind im wesentlichen auf die verringerten Zuleitungswiderstände zurückzuführen, die verringerten DIBL-Werte weisen auf eine Unterdrückung von Kurzkanaleffekten hin. Daß bei dieser Variante die SOT-Konfiguration die besseren Charakteristika hat, ist Hinweis auf ein neues „Nadelöhr“: Die Substratzuleitung ist jetzt limitierend, da der Kanal und damit das Substrat für bessere Transistoreigenschaften niedrig dotiert sein sollte (vergl. 3.6.2). Lediglich 130 nm tiefe implantierte Schichten führen zum Kanal. Obwohl die Zuleitungen mehrere Mikrometer breit sind, ist der Widerstand doch größer als beim Topkontakt. Beim EpiNovaFET ist dieses Problem durch Entkopplung von Substrat- und Kanaldotierung behoben. In Tabelle 7.1 sind die erreichten Kenngrößen für die verschiedenen Transistortypen abschließend gegenübergestellt.

### 7.3 EpiNovaFET

Diese Transistoren sind als *p*-Kanal-MOSFETs realisiert worden mit 50 nm Gatelänge und 6,6 nm Gateoxid. Abbildung 7.5 zeigt die Ausgangskennlinien zweier Transistoren, deren Geometrie sich lediglich in der Länge des Topkontaktes unterscheidet (Abb. 7.5 c): Der erste Transistor ist bei einer aktiven Steglänge von  $5 \mu\text{m}$  auf  $4 \mu\text{m}$  kontaktiert, der zweite auf nur  $2,5 \mu\text{m}$ . Das bedeutet, daß ein unterschiedlich großer Anteil des Topkontaktes wie im ursprünglichen Layout durch die obere Schicht geführt wird. Dieser Anteil hat zwar – wegen der höheren Dotierung – einen geringeren Widerstand als beim ursprünglichen NovaFET, allerdings einen deutlich höheren als der topkontaktierte Anteil.

Die Kennlinien 7.5 a+b sind für einen besseren Vergleich identisch skaliert. Ist  $U_G$  die Gatespannung bei der mit  $U_{DS} > U_{DSsat}$  der Sättigungsstrom  $I_{Dsat}$  fließt, so

	I-NovaFET				EpiNovaFET
	p-Kanal		n-Kanal		p-Kanal
Konfiguration	SOT	SOB	SOT	SOB	SOT
$g_m(\mu\text{S}/\mu\text{m})$	150	90	195	145	480
$S(\text{mV}/\text{dec})$	80	133	80	80	150
DIBL (mV/V)	30	120	32	55	210
$U_{th}(\text{V})$	0,27	0,24	-0,17	-0,17	-0,34

Tab. 7.1: Übersicht der erreichten Kenngrößen für die verschiedenen Konfigurationen des NovaFETs.

gilt für die durch den Sourcewiderstand  $R_S$  verringerte effektive Gatespannung  $U'_G$  nach [15]:

$$U'_G = U_G - I_{Dsat} \cdot R_S$$

Vergleicht man die Kennlinien bei gleichen Sättigungsströmen, so ergibt sich aus der Differenz der effektiven Gatespannungen die Differenz der Sourcewiderstände:

$$\begin{aligned}
 U'_{G1} - U'_{G2} &= I_{Dsat} \cdot R_{S1} - I_{Dsat} \cdot R_{S1} \\
 &= I_{Dsat}(R_{S1} - R_{S1}) \\
 R_{S1} - R_{S1} &= \frac{U'_{G1} - U'_{G2}}{I_{Dsat}}
 \end{aligned}$$

Damit läßt sich die Differenz der Widerstände zu  $200\,\Omega \pm 50\,\Omega$  bestimmen. Daß Widerstandsdifferenzen in dieser Größenordnung bereits solche Auswirkungen auf die Kennlinien haben, ist ein Zeichen dafür, daß die hohen Widerstände im  $\text{k}\Omega$ -Bereich erfolgreich vermieden werden konnten.

In Bild 7.6 a+b werden Ausgangs- und Transferkennlinien eines Transistors mit einem 90 nm breiten Steg und  $5\,\mu\text{m}$  aktiver Steglänge gezeigt. Obwohl DIBL-Wert (200 mV/V) und Unterschwellenspannungssteigung (150 mV/dec) nicht optimal sind, werden hohe Drainströme und ein gutes Sättigungsverhalten erreicht. Die Steilheit beträgt  $g_m = 480\,\mu\text{S}/\mu\text{m}$ . Der Einfluß der Stegbreiten ist in 7.6 c+d beschrieben. Dünnere Stege führen zu besserem Sättigungsverhalten, kleinerer Unterschwellenspannungssteigung und geringerem DIBL. Dies sind Zeichen für die Kopplung der beiden Gates, d. h. man kann im sub-100 nm Bereich nicht mehr von zwei parallel geschalteten vertikalen Single-Gate-Transistoren sprechen; es kann eindeutig Double-Gate-Verhalten konstatiert werden.

## 7.4 Einordnung

Diverse Gruppen forschen weltweit auf dem Gebiet der neuen MOSFET-Konzepte. Einen breiten Raum nimmt dabei der FinFET ein – lediglich eine Arbeit, die sich unter anderem auch mit vertikalen Double-Gate-MOSFETs beschäftigt, konnte gefunden werden. In der folgenden Tabelle werden die wichtigsten Kenngrößen der Transistoren dieser Arbeit und die anderer Gruppen gegenübergestellt.

	Typ	$S$ (mV/dec)	DIBL (mV/V)	$I_{on}$ ( $\mu\text{A}/\mu\text{m}$ )	$I_{off}$ (A/ $\mu\text{m}$ )	$U_{dd}$ (V)	$t_{ox}$ (nm)	$L$ (nm)	Zitat
1.	n-NovaFET	80	32	170	1 n	1,5	3,7	130	diese Arbeit
2.	p-NovaFET	80	30	77	4 p	1,5	3,7	75	diese Arbeit
3.	p-EpiNova	150	210	380	1 n	1,5	6,6	50	diese Arbeit
4.	n-FinFET	86	$\sim 100$	650	25 p	1,2	3	80	Infineon [43]
5.	n-FinFET	100	$\sim 350$	1100	25 p	1,2	3	40	Infineon [43]
6.	p-FinFET	69	k. A.	820	$\sim 150$ n	1,2	2,5	45	Berkeley [44]
7.	n-FinFET	125	71	892	$\sim 1 \mu$	1,2	1,7	10	AMD [45]
8.	p-FinFET	101	120	712	$\sim 4 \mu$	1,2	1,7	10	AMD [45]
9.	n- $\Omega$ -FET	75	11	1140	8 n	1	1,8	25	[46]
10.	p- $\Omega$ -FET	80	108	780	400 p	1	1,8	25	[46]
11.	SOI p-FET	71	24	328	186 n	1,2	1,2	14	IBM [47]
12.	vert. n-DG	120	k. A.	24	2 p	1,5	3	70	[4]

„k. A.“ bedeutet „keine Angabe“, ungefähre Werte in der Tabelle sind in der Veröffentlichung nicht explizit angegeben, sondern wurden aus den Kennlinien ausgelesen. Der  $\Omega$ -FET ist eine Weiterentwicklung des FinFETs, bei dem das Gate in  $\Omega$ -Form auch oben am Steg und teilweise unter dem Steg anliegt. Wegen der unterschiedlichen gebräuchlichen Definitionen der Gateweite bei Double-Gate-MOSFETs sind die Ströme bei den FinFETs auf die Steghöhe normiert, bei den vertikalen DG-MOSFETs auf die einfache aktive Steglänge.

Da bei den Transistoren dieser Arbeit die Schwellenspannung nicht eingestellt waren, sind zur Bestimmung von  $I_{on}$  und  $I_{off}$  nicht die Spannungen 0 V und  $U_{dd}$ , sondern  $U_0$  und  $U_0 + U_{dd}$  verwendet worden, wobei  $U_0$  geeignet gewählt wurde. Diese Methode wurde ebenfalls von Infineon (4.+5.) und AMD (7.+8.) angewandt. Während die I-NovaFETs (1.+2.) gute Unterschwellenspannungssteigungen und kleine DIBL-Werte aufweisen, sind die Ströme wegen großer Substratwiderstände recht gering. Beim EpiNovaFET (3.) dagegen sind deutlich höhere Ströme erreicht, allerdings sind Unterschwellenspannungssteigung und DIBL schlechter. Dies liegt am dicken Gateoxid, da  $t_{ox}$ , wie in 3.4 beschrieben, in die Unterschwellenspannungssteigung eingeht. Ein auf ein Drittel der Dicke reduziertes Gateoxid ließe eine Unterschwellenspannungssteigung von 90 mV/dec erwarten und  $I_{on}$  auf das Dreifache anwachsen: Der Sättigungsstrom ist im idealen Fall

$$I_{Dsat} = \frac{W C_{ox} \mu_{eff}}{2 L} (U_G - U_{th})^2.$$

Sättigung der Ladungsträgergeschwindigkeit könnte das Anwachsen des Stromes jedoch vorher beschränken. Dennoch kann man erwarten, daß mit dieser Änderung, der Reduzierung des Gateoxides auf ca. 2 nm, der p-EpiNovaFET mit den anderen p-Kanal-Transistoren in allen Punkten vergleichbar wäre, in der vorliegenden Form nur in einzelnen Werten. Die Werte aus den Studien des vertikalen n-DG-Transistors von Schulz können vom n-I-NovaFET jedoch übertroffen werden.

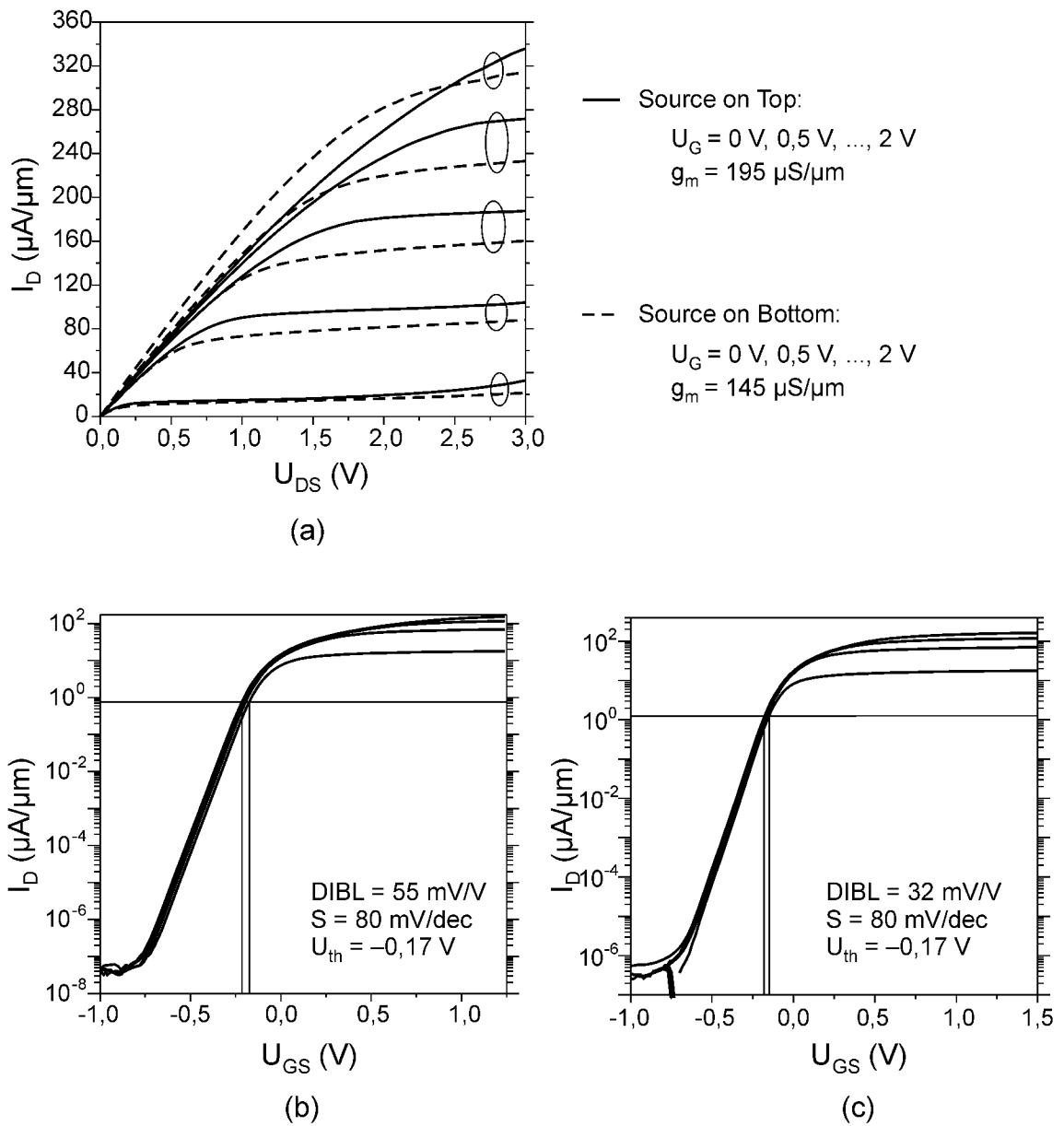


Abb. 7.4: *n*-Kanal-I-NovaFET: (a) Vergleich der Ausgangskennlinien für SOB und SOT. Die Transferekennlinien für (b) SOB und (c) SOT sind vergleichbar.

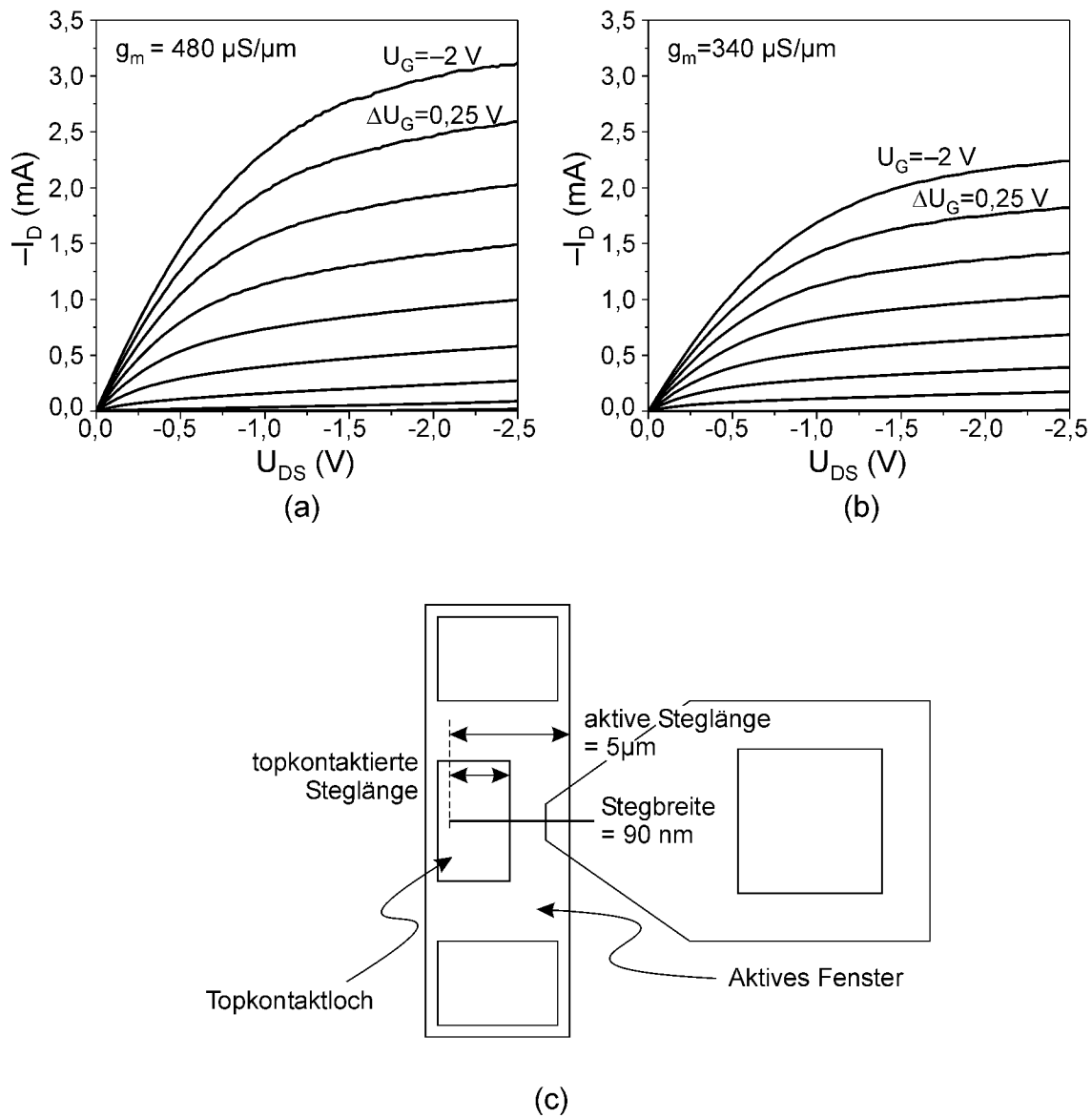


Abb. 7.5: EpiNovaFET: Vergleich der Ausgangskennlinienfelder zweier Transistoren mit gleicher Geometrie außer der topkontaktierten Steglänge, welche bei (a)  $4 \mu\text{m}$ , bei (b)  $2,5 \mu\text{m}$  beträgt. Der höhere Widerstand bei (b) wirkt sich deutlich auf die Kennlinie aus. (c) Geometrie der Transistoren.

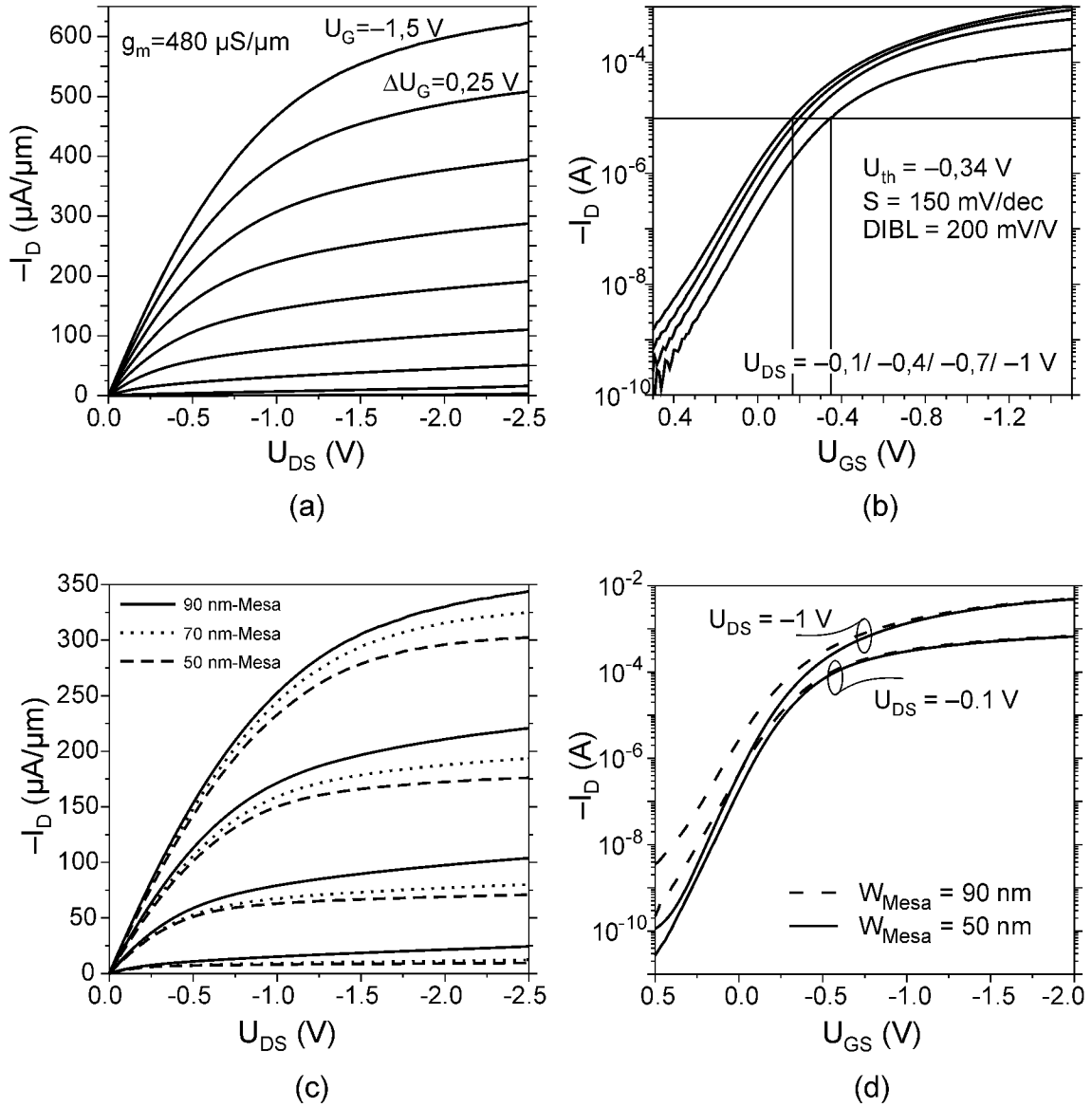


Abb. 7.6: EpiNovaFET: (a) + (b) zeigen die Kennlinienfelder eines Transistors mit einem 90 nm Steg der aktiven Länge  $5 \mu\text{m}$ . (c) + (d) vergleichen die Kennlinien für verschiedene Stegbreiten bei vier Stegen mit je  $5 \mu\text{m}$  aktiver Länge. Dünnere Linien führen zu niedrigeren Strömen, aber auch zu kleinerem DIBL und geringerer Unterschwellenspannungssteigung.



# Kapitel 8

## Zusammenfassung und Ausblick

In dieser Arbeit wurde die Entwicklung eines vertikalen Double-Gate-MOSFETs vorgestellt. Double-Gate-Transistoren werden von der ITRS-Roadmap als *non-classical CMOS*-Bauelemente gehandelt, als mögliche Nachfolger der heute üblichen planaren MOSFETs, wenn diese nicht weiter skaliert werden können. Das Prinzip des Double-Gates verspricht durch seine Symmetrie eine bessere Gatekontrolle des Kanalgebietes und damit die Unterdrückung von Kurzkanaleffekten, ferner ein verbessertes Skalierungspotential. Vertikale Bauelemente verlangen eine lithographieunabhängige Definition der Gatelänge und ermöglichen zudem eine hohe Packungsdichte.

Der hier beschriebene, „NovaFET“ genannte Transistor entstand im Rahmen des HGF-Projektes „Extended Silicon MOSFET“, ein Projekt am Institut für Schichten und Grenzflächen (ISG-1) des Forschungszentrums Jülich, das die Erforschung, Realisierung und Charakterisierung von Feldeffekttransistoren mit Kanallängen unter 100 nm zum Ziel hatte. Von der ursprünglichen Version ausgehend, wurde der NovaFET in zwei weiteren Layouts realisiert, dem „I-NovaFET“ und dem „EpiNovaFET“, die jeweils durch die Erfahrungen mit den vorausgegangenen Ergebnissen verbessert wurden. Kern aller NovaFETs ist ein dünner Siliziumsteg mit der vertikalen Schichtfolge p-n-p für einen p-Kanal-Transistor bzw. n-p-n für einen n-Kanal-Transistor. Nach der Gateoxidation werden an diesen Steg beidseitig Gatespacer angebracht. Zwischen dem Substratkontakt und dem Kontakt an der oberen Schicht wird die Source-Drain-Spannung angelegt.

Die Nanostrukturierung ist vielleicht die größte Herausforderung an diesem Transistorkonzept, da sehr regelmäßige, ca. 300 nm hohe und bis zu 30 nm dünne Stege produziert werden müssen. Der Einsatz von *Hydrogensilsesquioxan* (HSQ) als hochauflösender Elektronenstrahlresist bildet die Grundlage für diese Strukturierung. 23 nm breite Linien mit 100 nm Höhe konnten aus HSQ mit Elektronenstrahlschreiben hergestellt werden. Da HSQ anorganisch und nach dem Schreiben  $\text{SiO}_2$ -ähnlich ist, widersteht es der RIE-Ätzung besser als organische Resists. Durch den Einsatz einer RIE mit *inductive coupled plasma*- (ICP)-Technologie und HBr-basierenden Plasmen kann Silizium mit sehr hoher Selektivität zu  $\text{SiO}_2$  geätzt werden. Dadurch kann HSQ, obwohl nur 100 nm dick, als Maske zum Direktätzen von 300 nm Silizium dienen. Siliziumstege mit 25 nm Breite und 330 nm Höhe konnten dabei produziert werden, die genau den kleinsten Strukturen entsprechen, die für die Realisierung des

NovaFETs gefordert sind.

In der ursprünglichen Version, die als p-Kanal-MOSFET hergestellt wurde, ist die obere Zuleitung durch den Steg herausgeführt und über ein Pad kontaktiert, die Schichtfolge ist mittels LPCVD epitaktisch hergestellt. Niedrige Ströme ( $20\text{--}30\ \mu\text{A}/\mu\text{m}$ ) und starke Asymmetrie zwischen *Source on Top* (SOT) und *Source on Bottom* (SOB) belegen den zu großen Widerstand in der oberen Zuleitung und einen zu flachen unteren p-n-Übergang. Die zweite Version kontaktiert den Steg direkt von oben und stellt die Schichtfolge durch Implantation in die Nanostege und Diffusion her. Das Substrat ist hierbei in der (niedrigen) Dotierung des Kanalgebietes zu wählen. Der Topkontakt verlangt eine Planarisierung, die hier mit Chemisch-Mechanischem Polieren (CMP) durchgeführt wurde, und das hochselektive Ätzen von Silizium zu  $\text{SiO}_2$ , das mit den gleichen Plasmen wie die Siliziumstege erfolgt. Diese Ätzung wird hier für die definierte Rückätzung der Gatespacer aus Polysilizium benötigt, damit am Topkontakt keine Kurzschlüsse entstehen. Das Gateoxid wurde bei diesen Transistoren durch Niedrigtemperaturfeuchtoxidation hergestellt, einem recht neuen Verfahren zur langsamen Oxidation im Konvektionsofen.

Mit diesen Veränderungen wurde der I-NovaFET als p- und n-Kanal-Transistor prozessiert. Sie zeigen deutlich besseres Verhalten: Der p-NovaFET hat eine Steilheit von  $g_m = 150\ \mu\text{S}/\mu\text{m}$ , eine Unterschwellenspannungssteigung von  $S = 80\ \text{mV}/\text{V}$  und ein DIBL von  $30\ \text{mV}/\text{V}$ . Bei den n-Kanal-NovaFET wurde  $g_m = 195\ \mu\text{S}/\mu\text{m}$ ,  $S = 80\ \text{mV}/\text{V}$  und  $\text{DIBL} = 32\ \text{mV}/\text{V}$  gemessen. Diese Werte zeigen bereits unterdrückte Kurzkanaleffekte. Die Ströme sind jedoch mit  $77\ \mu\text{A}/\mu\text{m}$  (p-NovaFET) bzw.  $170\ \mu\text{A}/\mu\text{m}$  (n-NovaFET) am Arbeitspunkt bei  $1,5\ \text{V}$  immer noch nicht optimal. Auch bei dieser Variante sind Unterschiede zwischen SOT- und SOB-Konfiguration vorhanden, die diesmal jedoch auf hohe Zuleitungswiderstände im Substratkontakt hinweisen. Die implantierten Wannen sind zu flach für mehrere  $\mu\text{m}$ -lange Zuleitungen.

Mit der dritten Variante wurde dieser Engpaß beseitigt, indem auf ein hochdotiertes Substrat das Kanalgebiet (niedrig entgegengdotiert) epitaktisch aufgewachsen wurde. Die obere Schicht wurde durch Gegenimplantation definiert. Mit diesem Epi-NovaFET, der als p-Kanal-Transistor hergestellt wurde, sind  $380\ \mu\text{A}/\mu\text{m}$  bei  $1,5\ \text{V}$  Arbeitsspannung,  $g_m = 480\ \mu\text{S}/\mu\text{m}$ ,  $S = 150\ \text{mV}/\text{V}$  und  $\text{DIBL} = 210\ \text{mV}/\text{V}$  gemessen worden. Die Verschlechterung der Unterschwellenspannungssteigung und des DIBL sind auf das mit  $6,6\ \text{nm}$  noch nicht optimale Gateoxid zurückzuführen.

Der Einsatz von dünnerem Gateoxid ließe  $S$  und DIBL verkleinern und die Ströme weiter ansteigen, so daß dieser Transistortyp als prinzipiell gleichwertig den anderen neuen MOSFET-Konzepten gelten kann. Verwendet man undotiertes Silizium im Kanalgebiet, so können mit Hilfe von *buried layers* n- und p-Transistoren auf einem Substrat prozessiert und somit CMOS-Schaltungen realisiert werden. Infineon zeigt Interesse daran, diesen vertikalen Double-Gate-MOSFET – möglicherweise als Speichertransistor – einzusetzen.

# Anhang A

## Prozeßfolgen

### A.1 Prozeßprotokoll NovaFET (ursprüngliches Layout)

Bei dieser ursprünglichen Variante wird der aktive Bereich als p-n-p-Schichtfolge epitaktisch gewachsen. Der obere Kontakt wird entlang des Steges herausgeführt und über ein Pad kontaktiert.

1. **Substrat:** 3"-Silizium-Wafer, p/Bor-dotiert,  $0,005 - 0,02 \Omega \text{ cm} < 100 >$

2. **Epitaxie:**

Untere Schicht: 130 nm p-Si,  $3 \cdot 10^{17} - 1 \cdot 10^{18} \text{ cm}^{-3}$

Mittlere Schicht: 140 nm n-Si, ca.  $3 \cdot 10^{17} - 1 \cdot 10^{18} \text{ cm}^{-3}$

Obere Schicht: 130 nm p-Si, ca.  $3 \cdot 10^{18} \text{ cm}^{-3}$

3. **Dreilagenlack:**<sup>1</sup>

Unterer Resist: ausgehärteter AZ 5204

SiO<sub>2</sub>-Zwischenschicht

Oberer Resist: UVN 30 (verd., ca. 100 nm), Softbake 60 s bei 100°C

4. **Elektronenstrahlolithographie:**

---

<sup>1</sup>Dieser Dreilagenlack zur Herstellung der Siliziumstege war ein Ergebnis einer Diplomarbeit [48]. Nach der Einführung des HSQ als Resist wurden die Stege mit Zweilagenlack und später mit Direktätzen hergestellt.

Layer	Dosis 1 ( $\mu\text{C}/\text{cm}^2$ )	Dosis 2 ( $\mu\text{C}/\text{cm}^2$ )	Dosis 3 ( $\mu\text{C}/\text{cm}^2$ )	Step-Size (nm)	Beam-Size (nm)
UV Mesa	200	150	150	250	250
Mesa 25	700	700	800	12,5	15
Mesa 50	500	500	400	12,5	15
Mesa 75	400	400	350	12,5	15
Mesa 100	250	250	200	12,5	15
Mesa 300	250	250	200	12,5	15

Post-Exposure-Bake: 60 s 110°C

Entwicklung: 15 s MF CD-26

#### 5. Siliziumätzen: RIE 1

Lackschleier:  $\text{O}_2$ -Plasma (20 sccm  $\text{O}_2$ , 30  $\mu\text{bar}$ , 200 W Bias)

$\text{SiO}_2$ -Zwischenschicht:  $\text{CHF}_3$  20 sccm, 30  $\mu\text{bar}$ , 200 W Bias

Unterer Resist:  $\text{O}_2$ -Plasma

Silizium: 13 min 40 sccm Ar + 1 sccm  $\text{SF}_6$ , 20  $\mu\text{bar}$ , 150 W Bias

Resist entfernen:  $\text{O}_2$ -Plasma

Dektak-Messung: 360–370 nm-Stufen

#### 6. Reinigung: 10 min Piraña ( $\text{H}_2\text{O}_2 + \text{H}_2\text{SO}_4$ (2+3)) im Ultraschallbad

#### 7. Deposition 1. Deckoxid: 300 nm $\text{SiO}_2$

#### 8. Lithographie „Aktiv“:

Dehydrieren: 2 min 115°C

HMDS, UV 6 aufschleudern

Softbake: 60 s 130°C

Belichtung: 0,5 s (neuer Belichter)

Post-Exposure-Bake: 90 s 135°C

Entwicklung: 45 s MF 84-MX

#### 9. Öffnen des aktiven Bereichs:

Lackschleier: 20 s Plasmaverascher, Programm 1

Aushärten des Resists: 20 min 125°C

$\text{SiO}_2$ -Ätzen: AF 91-09

#### 10. RCA-Reinigung: (80°C)

$\text{H}_2\text{O}_2 + \text{H}_2\text{SO}_4$  (2+3), 10 min, anschl. Spülen in DI-Wasser

$\text{H}_2\text{O}_2 + \text{NH}_4\text{OH} + \text{H}_2\text{O}$  (1+1+5), 10 min, anschl. Spülen in DI-Wasser

$\text{H}_2\text{O}_2 + \text{HCl} + \text{H}_2\text{O}$  (1+1+5), 10 min, anschl. Spülen in DI-Wasser

11. **Gateoxidation:**

60 s 0,5%ige Flußsäure, je 200 ml 5 Tr. Propanol

RTO: 3 min bei 950°C (Programm *qttrox No 2*)

12. **Deposition Polysilizium:** direkt im Anschluß an die Gateoxidation 200 nm p-Polysilizium (560°C)

13. **Temperung:** RTP 20 s 1050°C (Programm *JBor-a*)

14. **Lithographie „Poly“:** (wie Lithographie „Aktiv“)

15. **Polysilizium zurückätzen:** RIE 1

Lackschleier:  $\text{O}_2$ -Plasma

Polysilizium: 40 sccm Ar + 1 sccm  $\text{SF}_6$ , 20  $\mu\text{bar}$ , 150 W Bias

Resist entfernen:  $\text{O}_2$ -Plasma

16. **Reinigung:** 10 min Piraña im Ultraschallbad

17. **Deposition 2. Deckoxid:** 300 nm  $\text{SiO}_2$

18. **Lithographie „Substrat/Gate“:** (wie Lithographie „Aktiv“)

19. **Substrat-/Gatekontaktlöcher ätzen:**

Hardbake: 3 min 140°C

RIE 1:

Lackschleier:  $\text{O}_2$ -Plasma

$\text{SiO}_2$ :  $\text{CHF}_3$  20 sccm, 30  $\mu\text{bar}$ , 200 W Bias

Reinigung: Aceton + Propanol

Reinigung: 10 min Piraña im Ultraschallbad

20. **Lithographie „Top“:** (wie Lithographie „Aktiv“)

21. **Topkontaktlöcher ätzen:**

Hardbake: 3 min 140°C

RIE 1:

Lackschleier:  $\text{O}_2$ -Plasma

$\text{SiO}_2$ :  $\text{CHF}_3$  20 sccm, 30  $\mu\text{bar}$ , 200 W Bias

22. **RCA-Reinigung**

23. **Nickel aufdampfen:** nach 5 s AF 91-09 und 30 s DI-Wasser 20 nm Nickel aufdampfen
24. **Silizidierung:** RTP 560°C
25. **Nickel entfernen:**  $\text{H}_2\text{O}_2 + \text{HCl}$  (1+3)
26. **Lithographie „Metall“:**
  - HMDS, UVN 30 aufschleudern
  - Softbake: 60 s 90°C
  - Belichtung: 0,5 s (neuer Belichter)
  - Post-Exposure-Bake: 60 s 95°C
  - Entwicklung: 30 s MF CD-26
27. **Aluminium aufdampfen:** 300 nm, mit vorherigem Ar-Sputtern
28. **Lift-Off**

## A.2 Prozeßprotokoll I-NovaFET (Implantationsvariante)

Das hier verwendete Layout unterscheidet sich in zwei wesentlichen Punkten vom vorherigen: Zum einen wird die p-n-p-Abfolge für einen p-Kanal-MOSFET (bei n-Kanal-MOSFET n-p-n) hier alleine mit Implantation realisiert, und zum anderen wird die obere Schicht direkt oben an der Mesa kontaktiert. Dies erfordert eine Planarisierung, die hier mit CMP durchgeführt wird.

1. **Substrat:** 100 mm-Silizium-Wafer <100> mit gewünschter Dotierung des Kanalgebietes
2. **Elektronenstrahlolithographie:** „Moers-Maske-5.dxf“

Belackung mit HSQ

Layer	Step-Size (nm)	Beam-Size (nm)	Anfangsdosis ( $\mu\text{C}/\text{cm}^2$ )	Erhöhung ( $\mu\text{C}/\text{cm}^2$ )
UV Mesa E-Beam Grob	250	300	300	(fest)
Mesa 5	5	15	7100	+ 200
Mesa 25	12,5	15	4600	+ 200
Mesa 37	12,5	15	3850	+ 200
Mesa 50	12,5	15	3100	+ 200
Mesa 75	12,5	15	2600	+ 200

Entwicklung: 1 min MF CD-26

3. **Hardbake:** 1 h 450°C in N<sub>2</sub>

4. **Siliziumätzen:**

HSQ-Schleier: 5 s CHF<sub>3</sub> 20 sccm, 30  $\mu\text{bar}$ , 300 W Bias (RIE 1)

Plasma 1: 60 s HBr (50 sccm, 7 mTorr, 50 W Bias, 750 W ICP) (RIE 3)

Plasma 2: 60 s HBr + O<sub>2</sub> (50 sccm + 2 sccm, 7 mTorr, 100 W Bias, 2000 W ICP) (RIE 3)

Resist entfernen: 45 s 2%ige Flußsäure

5. **Reinigung:** 10 min Piraña (H<sub>2</sub>O<sub>2</sub> + H<sub>2</sub>SO<sub>4</sub> (2+1)) im Ultraschallbad

6. **Deposition Implantationsoxid:** 100 nm TEOS-SiO<sub>2</sub> (LPCVD)

7. **Chrom aufdampfen:** 100 nm (rotierend)

8. **Lithographie „ImpOx“:**

Dehydrieren: 90 s 115°C

HMDS (VPO, Progr. 1)

UV 6 aufschleudern

Softbake: 60 s 130°C

Belichtung: 2,1 mJ/cm<sup>2</sup>

Post-Exposure-Bake: 90 s 140°C

Entwicklung: 45 s MF 84-MX

9. **Öffnen des Implantationsbereichs:**

Lackschleier: O<sub>2</sub>-Plasma (RIE 1)

Chrom-Etch ca. 100 s

SiO<sub>2</sub>-Ätzen: 0,5%ige Flußsäure (ca. 7 min)

Reinigung: Aceton + Propanol, O<sub>2</sub>-Plasma (RIE 1), Chrom-Etch

**10. RCA-Reinigung (60°C):**

H<sub>2</sub>O<sub>2</sub> + H<sub>2</sub>SO<sub>4</sub> (2+1), 10 min, (alter RR,) anschl. Spülen in DI-Wasser + Waschmaschine

H<sub>2</sub>O<sub>2</sub> + H<sub>2</sub>SO<sub>4</sub> (2+1), 10 min, (neuer RR,) anschl. Spülen in DI-Wasser  
HF-Dip: 15 s 0,5%ige Flußsäure

H<sub>2</sub>O + H<sub>2</sub>O<sub>2</sub> + NH<sub>4</sub>OH (20+4+1), 10 min, anschl. Spülen in DI-Wasser

H<sub>2</sub>O + H<sub>2</sub>O<sub>2</sub> + HCl (20+1+1), 10 min, anschl. Spülen in DI-Wasser + Waschmaschine

**11. Oxid-Spacer:**

15 nm thermisches Oxid

RIE 1: CHF<sub>3</sub> 20 sccm, 30 μbar, 200 W

**12. Reinigung:** 10 min Piranha (H<sub>2</sub>O<sub>2</sub> + H<sub>2</sub>SO<sub>4</sub> (2+1)) im Ultraschallbad

**13. Implantation:**

Bor, 1 keV, 7°/21°,  $2 \cdot 10^{15} \text{ cm}^{-2}$  (für p-Kanal-MOSFET)

Arsen, 10 keV, 7°/21°,  $2 \cdot 10^{15} \text{ cm}^{-2}$  (für n-Kanal-MOSFET)

**14. Annealing + Diffusion 1:** 20 s 950°C (Die erste Temperung soll Implantationsschäden ausheilen und die Mesa so ätzresistenter machen. Der wichtige Diffusionsschritt zur Bildung des Kanalbereichs erfolgt zusammen mit der Aktivierung des Polysiliziums und dem Annealing der Kontakt-Implantation.)

**15. Oxid entfernen:** 0,5%ige Flußsäure

**16. Deposition 1. Deckoxid:** 120 nm TEOS-SiO<sub>2</sub> (LPCVD)

**17. Chrom aufdampfen:** 100 nm (rotierend)

**18. Lithographie „Aktiv“:** (wie Lithographie „ImpOx“)

**19. Öffnen des aktiven Bereichs:**

Lackschleier: O<sub>2</sub>-Plasma

Chrom-Etch ca. 100 s

SiO<sub>2</sub>-Ätzen: 0,5%ige Flußsäure

Reinigung: Aceton + Propanol, O<sub>2</sub>-Plasma, Chrom-Etch

**20. RCA-Reinigung (60°C)**

**21. Gateoxidation:** Niedrigtemperatur-Feuchtoxidation, 1 h 600°C

22. **Deposition Polysilizium:** direkt im Anschluß an die Gateoxidation 150 nm p-Polysilizium (650°C) (n-Polysilizium für n-Kanal-MOSFET)
23. **Hartmaske:** Deposition von 110 nm TEOS-SiO<sub>2</sub>
24. **Lithographie „Poly“:** (wie Lithographie „ImpOx“)
25. **Hartmaske ätzen:**
  - Lackschleier: 30 s Plasmaverascher, Programm 1
  - Oxidätzung: AF 91-09
  - Resist entfernen: O<sub>2</sub>-Plasma
26. **Polysilizium zurückätzen:**
  - nat. Oxid entfernen: 10 s AF 91-09
  - Polysilizium: 30 s HBr (50 sccm, 7 mTorr, 50 W Bias, 750 W ICP); anschl. HBr + O<sub>2</sub> (50 sccm + 2 sccm, 7 mTorr, 100 W Bias, 2000 W ICP), 100 nm überätzen (ca. 30 s)
  - Resist entfernen: O<sub>2</sub>-Plasma
27. **Reinigung:** 10 min Piraña im Ultraschallbad
28. **Deposition 2. Deckoxid:** 700 nm TEOS-SiO<sub>2</sub>
29. **Chemisch-Mechanisches Polieren:** 60 s, 425 N, 40 kPa, 130 rpm; vorher Parameter durch Test überprüfen, da Drift möglich.
30. **RCA-Reinigung (60°C)**
31. **Lithographie „Gate“:** (wie Lithographie „ImpOx“)
32. **Gatekontaktlöcher ätzen:**
  - Lackschleier: O<sub>2</sub>-Plasma
  - SiO<sub>2</sub>: CHF<sub>3</sub> 20 sccm, 30  $\mu$ bar, 200 W Bias
  - Resist entfernen: O<sub>2</sub>-Plasma
33. **Lithographie „Drain“:** (wie Lithographie „ImpOx“)
34. **Topkontaktlöcher ätzen:**
  - Lackschleier: O<sub>2</sub>-Plasma
  - SiO<sub>2</sub>: CHF<sub>3</sub> 20 sccm, 30  $\mu$ bar, 200 W Bias
  - Ätzdauer vorher bestimmen aus Ätzrate, Oxiddicke nach CMP und Steghöhen.
  - Resist entfernen: O<sub>2</sub>-Plasma
35. **Lithographie „Source“:** (wie Lithographie „ImpOx“)

**36. Substratkontaktlöcher ätzen:**

Lackschleier: O<sub>2</sub>-Plasma

SiO<sub>2</sub>: CHF<sub>3</sub> 20 sccm, 30 μbar, 200 W Bias

Resist entfernen: O<sub>2</sub>-Plasma

**37. RCA-Reinigung (60°C)****38. Kontakt-Implantation:**

2 keV Bor, 7°/21°,  $2 \cdot 10^{15} \text{ cm}^{-2}$  (für p-Kanal-MOSFET)

15 keV Arsen, 7°/21°,  $2 \cdot 10^{15} \text{ cm}^{-2}$  (für n-Kanal-MOSFET)

**39. Annealing + Diffusion 2:**

15 s 850°C + 10 s 1050°C (für p-Kanal-MOSFET)

20 s 850°C + 20 s 1050°C (für n-Kanal-MOSFET)

**40. RCA-Reinigung (60°C)**

**41. Nickel aufdampfen:** nach 60 s 0,5%ige Flußsäure und 30 s DI-Wasser 15 nm Nickel aufdampfen

**42. Silizidierung:** RTP 25 s 560°C

**43. Nickel entfernen:** H<sub>2</sub>O<sub>2</sub> + HCl (1+3) für 2:30 min

**44. Aluminium aufdampfen:** 350 nm drehend, mit vorherigem Ar-Sputtern (15 s)

**45. Lithographie „Metall“:**

Dehydrieren: 90 s 115°C

HMDS (VPO, Progr. 1)

ma-N 2405 aufschleudern

Softbake: 90 s 90°C

Belichtung: 13,5 s (neuer Belichter)

Entwicklung: 30 s ma-N 532

**46. Aluminium ätzen:** H<sub>3</sub>PO<sub>4</sub> + HNO<sub>3</sub> + H<sub>2</sub>O (10+1+4), ca 3:30 min, anschl. Resist entfernen mit Aceton + Propanol; ggf. vor der Ätzung Hardbake.

## A.3 Prozeßprotokoll EpiNovaFET

Im Unterschied zur zweiten Variante wird die Schichtfolge p-n-p (bzw. n-p-n) hier durch Epitaxie (mit der Dotierung des Kanalgebietes) und Implantation (Gegendotierung der obersten Schicht) erzeugt. Dies vermeidet die flachen Leitungskanäle zu den Substratkontakten.

1. **Substrat:** 100 mm-Silizium-Wafer  $\langle 100 \rangle$ , hoch p-dotiert
2. **Epitaxie:** 300 nm n-Si,  $5 \cdot 10^{17} \text{ cm}^{-3}$  Sb
3. **Implantation:** 2 keV Bor,  $7^\circ/21^\circ$ ,  $2 \cdot 10^{15} \text{ cm}^{-2}$
4. **Annealing + Diffusion 1:** RTP 20 s  $950^\circ\text{C}$
5. **Elektronenstrahlolithographie + Silizium ätzen:** Wie beim I-NovaFET
6. **Reinigung:** 10 min Piraña ( $\text{H}_2\text{O}_2 + \text{H}_2\text{SO}_4$  (2+1)) im Ultraschallbad

Weiter wie bei I-NovaFET ab 16.



# Anhang B

## Physikalische Größen und Abkürzungen

$C_{\text{high-}\kappa}$	Gatekapazität mit alternativem Gateoxid als Dielektrikum
$C_{ox}$	Gatekapazität mit $\text{SiO}_2$ als Dielektrikum
CMP	Chemisch-Mechanisches Polieren
$EOT$	<i>equivalent oxide thickness</i>
$g_m$	auf Gateweite normierte Steilheit eines Transistors
HSQ	<i>Hydrogen Silsesquioxan</i> , anorganisches Polymer, als <i>flowable oxide</i> und Elektronenstrahlresist einsetzbar
$I_D$	Drainstrom eines Transistors
$I_{off}$	Leck-Drainstrom bei $U_G = 0$ und $U_{DS} = U_{dd}$
ITRS	<i>International Technology Roadmap of Semiconductors</i>
$k$	$= \varepsilon$ , relative Dielektrizitätskonstante
$k$	Boltzmannkonstante
$L$	geometrische Gatelänge
$L_{eff}$	effektive Gatelänge
LPCVD	<i>low-pressure chemical vapor deposition</i>
MFS	<i>minimal feature size</i>

MIBK	<i>Methylisobutylketon</i> , Lösungsmittel
MOS	<i>metal-oxide-semiconductor</i> , Metall-Oxid-Halbleiter-Schichtfolge
MOSFET	<i>metal-oxide-semiconductor field-effect transistor</i> , Metall-Oxid-Halbleiter-Feldeffekttransistor
$N_A$	Akzeptorkonzentration in p-Halbleiter
$N_D$	Donatorkonzentration in n-Halbleiter
$n_i$	intrinsische Ladungsträgerdichte
$N_{sub}$	Fremdatomkonzentration im Kanal
RIE	<i>reactive ion etching</i> , Reaktives Ionenätzen
$S$	Unterschwelligkeitssteigung
SIMS	<i>secondary ion mass spectrometry</i>
$t_{dep}$	Dicke der Raumladungszone unter dem Gate
$t_{ox}$	Dicke des Gateoxides
$t_{si}$	Dicke des Siliziums unter dem Gate bei SOI- oder Double-Gate-Transistoren
TEOS	<i>Tetraethylorthosilicat</i> , $(OC_2H_5)_4$ , Flüssigkeit, aus der in der LPCVD Oxid deponiert wird
$U_D$	Source-Drain-Spannung
$U_{dd}$	Arbeitsspannung eines Transistors
$U_{FB}$	Flachbandspannung eines Transistors
$U_G$	Gatespannung
$U_{th}$	Schwellenspannung, Einsatzspannung
$q$	Elementarladung
W	Gateweite

# Literaturverzeichnis

- [1] G. Moore: *Cramming more components onto integrated circuits*, Electronics **38**, 4 (1965), 114-7
- [2] L. Risch: *The end of the CMOS roadmap – new landscape beyond*, Materials Science and Engineering **C 19** (2002), 363-8
- [3] The International Technology Roadmap for Semiconductors: *Emerging Research Devices*, <http://public.itrs.net/Files/2001ITRS/PIDS.pdf>, (2001) 31
- [4] T. Schulz: *Konzepte zur lithographieunabhängigen Skalierung von vertikalen Kurzkanal-MOS-Feldeffekt-Transistoren und deren Bewertung*, Dissertation an der Ruhr-Universität Bochum (2001)
- [5] J. Moers: *Entwicklung und Charakterisierung vertikaler Silizium-Feldeffekttransistoren auf der Basis selektiver Epitaxie*, Berichte des Forschungszentrums Jülich 3535, ISSN 0944-2952
- [6] P. Kluth: *Selbstassemblierte Nanostrukturierung ultradünner Silizidschichten und Entwicklung von Nano-MOSFET-Bauelementen*, Berichte des Forschungszentrums Jülich 3994, ISSN 0944-2952
- [7] S. Mantl, M. Dolle, St. Mesters, P. F. P. Fichtner, H. L. Bay: *Patterning method for silicides based on local oxidation*, Appl. Phys. Lett. , **67** (1995), 3459-61
- [8] P. Kluth: *persönliche Mitteilung*
- [9] S. M. Sze: *Semiconductor devices, physics and technologie*, Wiley, New York (1985)
- [10] H. Schaumburg: *Werkstoffe und Bauelemente der Elektrotechnik 2, Halbleiter*, Teubner, Stuttgart (1991)
- [11] K. Heime: *Elektronische Bauelemente*, Vorlesungsskript, RWTH Aachen
- [12] S. M. Sze: *Physics of Semiconductor Devices*, Wiley, New York (1981)
- [13] R. H. Dennard, F. H. Gaensslen, H.-N. Yu, V. L. Rideout, E. Bassous, A. R. LeBlanc: *Design on ion-implanted MOSFET's with very small physical dimensions*, IEEE J. Solid-State Circuits, **SC-9** (1974), 256-68
- [14] Y. Taur, D. A. Buchanan, W. Chen, D. J. Frank, K. E. Ismail, S.-H. Lo, G. A. Sai-Halas, R. G. Viswanathan, H.-J. C. Wann, S. J. Wind, H.-S. Wong: *CMOS Scaling into the Nanometer Regime*, Proc. IEEE **85** (1997), 486-503

- [15] C. H. Wann, K. Noda, T. Tanaka, M. Yoshida: *A Comparative Study of Advanced MOSFET Concepts*, IEEE Transactions on Electron Devices **43** (1996), 1742-52
- [16] D. J. Frank, Y. Taur: *Design considerations for CMOS near the limits of scaling*, Solid-State Electronics **46** (2002), 315-20
- [17] Y. Taur: *CMOS design near the limit of scaling*, IBM J. Res. & Dev. **46** (2002), 213-22
- [18] D. A. Muller, T. Sorsch, S. Moccio, F. H. Baumann, K. Evans-Lutterodt, G. Timp: *The electronic structure at the atomic scale of ultrathin gate oxides*, nature **399** (1999), 758-61
- [19] The International Technology Roadmap for Semiconductors: *2002 Update*, <http://public.itrs.net> (2002)
- [20] B. Hoppe: *Mikroelektronik 1, Prinzipien, Bauelemente und Werkstoffe der Silizium-technologie*, Vogel, Würzburg (1997)
- [21] H.-S. Wong, D. J. Frank, P. M. Solomon, C. H. J. Wann, J. J. Welser: *Nanoscale CMOS*, Proceedings of the IEEE **87** (1999), 537-70
- [22] L. F. Thompson, C. G. Grant, M. J. S. Bowden: *Introduction to microlithography*, ACS symposium series, Salem (1983), 48
- [23] H. Ryssel, H. Glawischnig *Ion Implantation Techniques*, Springer-Verlag, Berlin, Heidelberg, New York (1982)
- [24] G. Carter, W. A. Grant: *Ionenimplantation in der Halbleitertechnik*, Carl Hanser Verlag, München, Wien (1981)
- [25] A. Benninghoven, F. G. Rüdenauer, H. W. Werner: *Secondary Ion Mass Spectrometry*, John Wiley & Sons, Inc., New York (1987)
- [26] S. Sivaram, H. Bath, R. Leggett, A. Maury, K. Monning, R. Tolles: *Planarizing Inter-level Dielectrics by Chemical-Mechanical Polishing*, Solid State Technologic **35** (May 1992), 87-91
- [27] S. Kluth: *persönliche Mitteilung*
- [28] J. Appenzeller, J. A. del Alamo, R. Martel, K. Chan, P. Solomon: *Ultrathin 600° C Wet Thermal Silicon Dioxide*, Electrochemical and Solid-State Letters **3** (2000), 84-6
- [29] M. Goryll, J. Moers, St. Trellenkamp, L. Vescan, M. Marso, P. Kordoš, H. Lüth: *Thin low-temperature gate oxides for vertical field-effect transistors*, Proceedings of the 4<sup>th</sup> International Conference on Advanced Semiconductor Devices and Microsystems 2002, 275-7
- [30] J. Jujita, Y. Ohnishi, Y. Ochiai, S. Matsui: *Ultrahigh resolution of calixarene negative resist in electron beam lithography*, Appl. Phys. Lett. **68** (1996), 1297
- [31] H. Namatsu, Y. Takahashi, K. Yamazaki, T. Yamaguchi, M. Nagase, K. Kurihara: *Three-dimensional siloxane resist for the formation of nanopatterns with minimum linewidth fluctuations*, J. Vac. Sci. Technol. B **16** (1998), 69

- [32] J. Kretz, L. Dreeskornfeld, J. Hartwich, W. Rösner: *20 nm electron beam lithography and reactive ion etching for the fabrication of double gate Fin FET devices*, Microelectronic Engineering **67-68** (2003), 763-8
- [33] A. D. Butherus, T. W. Hou, C. J. Mogab, H. Schonhorn: *O<sub>2</sub> plasma-converted spin-on-glass for planarization*, J. Vac. Sci. Technol. B **3** (1985), 1352
- [34] The International Technology Roadmap for Semiconductors: *Interconnect*, <http://public.itrs.net/Files/2002Update/2001ITRS/Interconnect.pdf> (2001), 14
- [35] R. H. Baney, M. Itho, A. Sakakibara, T. Suzuki: *Silsesquioxanes*, Chem. Rev. **95** (1995), 1409
- [36] M. J. Loboda, C. M. Grove, R. F. Schneider: *Properties of a-SiO<sub>x</sub>:H Thin Films Deposited from Hydrogen Silsesquioxane Resins*, J. Electrochem. Soc. **145** (1998), 2861
- [37] H. Meynen, Senior Application Engineer for Semiconductor Fabrication Materials, Dow Corning Corp., Belgien: *persönliche Mitteilung*
- [38] B. E. Maile, W. Henschel, H. Kurz, B. Rienks, R. Polman, P. Kaars: *Sub-10 nm Line-width an Overlay Performance Achieved with a Fine-Tuned EBPG-5000 TFE Electron Beam Lithography System*, Jpn. J. Appl. Phys. **39** (2000), 6836-42
- [39] J. Appenzeller, IBM T. J. Watson Research Center, Yorktown Heights NY: *persönliche Mitteilung*
- [40] H.-C. Liou, J. Pretzer: *Effect of curing temperature on the mechanical properties of hydrogen silsesquioxane thin films*, Thin Solid Films **335** (1998), 186-91
- [41] St. Trellenkamp, J. Moers, A. van der Hart, P. Kordoš, H. Lüth: *Patterning of 25-nm-wide silicon webs with an aspect ratio of 13*, Microelectronic Engineering **67-68** (2003), 376-80
- [42] H. Ibach, H. Lüth: *Festkörperphysik*, Springer-Verlag, Berlin, Heidelberg, New York, 5. Aufl. (1999)
- [43] Infineon Technologies – L. Risch (Projektleiter): *Verbundprojekt: Extended CMOS. Abschlußbericht* (2003)
- [44] X. Huang, W.-C. Lee, C. Kuo, D. Hisamoto, L. Chang, J. Kedzierski, E. Anderson, H. Takeuchi, Y.-K. Choi, K. Asano, V. Subramanian, T. J. King, J. Bokor, C. Hu: *Sub-50 nm P-Channel FinFET*, IEEE Transactions on Electron Devices **48** (2001), 880-6
- [45] B. Yu, L. Chang, S. Ahmed, H. Wand, S. Bell, C.-Q. Qang, C. Tabery, C. Ho, Q. Xiang, T.-J. King, J. Bokor, C. Hu, M.-R. Lin, D. Kyser: *FinFET Scaling to 10 nm Gate Length*, International Electron Devices Meeting (IEDM 2002) – Technical Digest, 10.02
- [46] F.-L. Yang, H.-Y. Chen, F.-C. Chen, C.-C. Huang, C.-Y. Cahng, H.-K. Chiu, C.-C. Lee, C.-C. Chen, H.-T. Huang, C.-J. Chen, H.-J. Tao, Y.-C. Yeo, M.-S. Liang, C. Hu: *25 nm CMOS Omega FETs*, International Electron Devices Meeting (IEDM 2002) – Technical Digest, 10.03

- [47] B. Doris, M. Jeong, T. Kanarsky, Y. Zhang, R. A. Roy, O. Dokumaci, Z. Ren, F.-F. Jamin, L. Shi, W. Natzle, H.-J. Huang, J. Mezzapelle, A. Mocuta, S. Womack, M. Gribelyuk, E. C. Jones, R. J. Miller, H.-S. P. Wong, W. Haensch: *Extreme Scaling with Ultra-Thin Si Channel MOSFETs*, International Electron Devices Meeting (IEDM 2002) – Technical Digest, 10.06
- [48] St. Trellenkamp: *Entwicklung spezieller Prozeßschritte zur Herstellung vertikaler Double-Gate-MOSFETs*, Diplomarbeit im ISI des FZ Jülich (2000)

# Danke

Mein Dank gilt all denen, die es mir ermöglicht haben, diese Arbeit durchzuführen. Ihre Mitarbeit, Unterstützung und Ratschläge haben wesentlich zu deren Entstehung beigetragen. Besonderer Dank gilt dabei

Herrn Prof. Dr. H. Lüth für die wissenschaftliche Betreuung dieser Arbeit und die Möglichkeit, diese Arbeit am Institut für Schichten und Grenzflächen durchzuführen.

Herrn Prof. Dr. S. Mantl für die freundliche Bereitschaft, das Korreferat zu übernehmen, und für interessanten Austausch im Rahmen der „X-MOS-Runden“.

Herrn Prof. Dr. P. Kordoš für die Unterstützung und Motivierung.

Herrn Dr. J. Moers für die intensive Begleitung der Arbeit, fachlichen Rat und Teilhabe an seinen großen Erfahrungen in Siliziumtechnologie.

Herrn Dr. A. van der Hart für die gute und erfolgreiche Zusammenarbeit beim Elektronenstrahlschreiben.

Herrn Dipl.-Ing. H. P. Bochem für viele hervorragende REM-Aufnahmen. Ohne die aus diesen Bildern gewonnenen Informationen wäre die Arbeit nicht durchführbar gewesen.

Herrn Dr. M. Goryll für die Charakterisierung der Gateoxide.

Frau Dr. S. Kluth für die Untersuchungen der Implantationsprofile in Nanostrukturen sowie für sorgfältiges Korrekturlesen von Veröffentlichungen.

Herrn Dipl.-Ing. A. Steffen, Herrn J. Müller und dem gesamten Reinraumteam für die immerwährende Unterstützung bei technischen Problemen.

Herrn Dipl.-Ing. F. Ringelmann für die erfolgreiche Instandhaltung der RIE.

Herrn Dipl.-Ing. A. Steffen, Herrn J. Müller, Herrn J. Zillikens, Frau A. Pracht, Herrn H. Wingens und Herrn K. H. Deussen für Dutzende Depositionen, Aufdampfungen und Oxidationen.

Herrn W. Michelsen und Frau K. Panitz für die Implantationen.

Frau Dr. L. Vescan und Herrn Dr. H. Bay für die Epitaxien.

Herrn Dipl.-Ing. A. Mück für die SIMS.

M. Marso, M. Mikulics, M. Wolter, M. Goryll, P. Javorka, S. Montanari und J. Bernát für die gute Stimmung beim Besuch von Konferenzen.

M. Mikulics, C. Varini sowie F. Semjank für die angenehme Atmosphäre im Büro.

allen Kolleginnen und Kollegen für die guten Arbeitsbedingungen.

meinen Eltern, die mir das Studium ermöglichten und damit den Grundstein zu dieser Arbeit legten.

meiner Frau Sabine Ley, nicht nur für die sorgfältige sprachliche Durchsicht der Arbeit.



Forschungszentrum Jülich  
*in der Helmholtz-Gemeinschaft*



Jül-4139  
Juli 2004  
ISSN 0944-2952