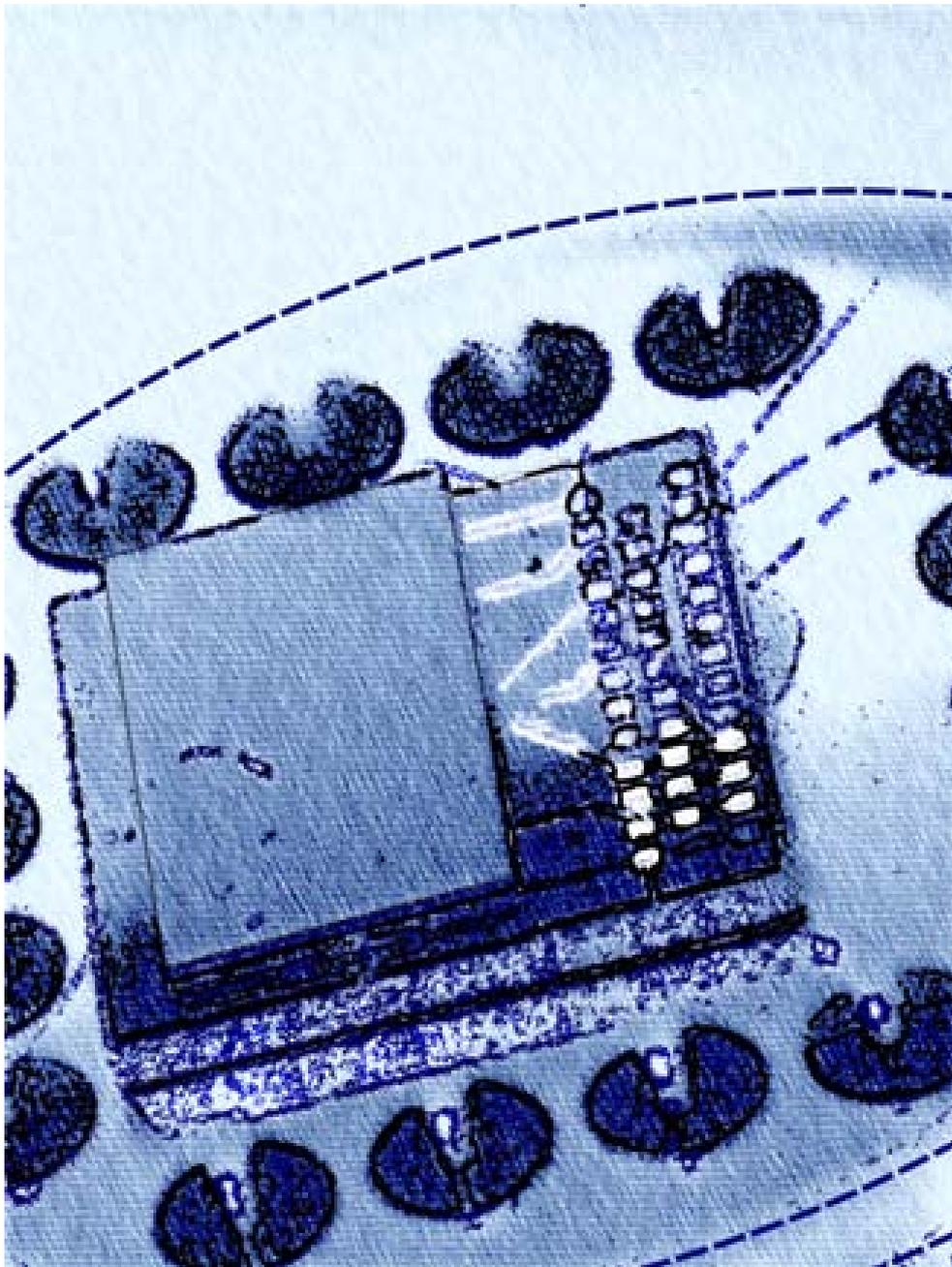


Gunter Freitag

Entwicklung eines neuen Feldeffekt- Gassensors mit hybridem Gate-Aufbau und vertikalem Transistordesign



Cuvillier Verlag Göttingen

UNIVERSITÄT DER BUNDESWEHR MÜNCHEN
Fakultät für Elektrotechnik und Informationstechnik

**Entwicklung eines neuen Feldeffekt-
Gassensors mit hybridem Gate-Aufbau
und vertikalem Transistordesign**

Gunter Freitag

Vorsitzender des Promotionsausschusses: Prof. Dr.-Ing. D. Gerling
1. Berichterstatter: Prof. Dr. rer. nat. I. Eisele
2. Berichterstatter: Prof. Dr. rer. nat. H. Meixner

Tag der Einreichung: 19.10.2004
Tag der Prüfung: 18.3.2005

Mit der Promotion erlangter akademischer Grad:
Doktor-Ingenieur
(Dr.-Ing.)

München, den 25.4.2005

Der Druck der Arbeit wurde durch Haushaltsmittel der Universität der Bundeswehr München gefördert.

Bibliografische Information Der Deutschen Bibliothek

Die Deutsche Bibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.ddb.de> abrufbar.

1. Aufl. - Göttingen : Cuvillier, 2005
Zugl.: München, Univ., Diss., 2005
ISBN 3-86537-448-4

© CUVILLIER VERLAG, Göttingen 2005
Nonnenstieg 8, 37075 Göttingen
Telefon: 0551-54724-0
Telefax: 0551-54724-21
www.cuvillier.de

Alle Rechte vorbehalten. Ohne ausdrückliche Genehmigung des Verlages ist es nicht gestattet, das Buch oder Teile daraus auf fotomechanischem Weg (Fotokopie, Mikrokopie) zu vervielfältigen.

1. Auflage, 2005
Gedruckt auf säurefreiem Papier

ISBN 3-86537-448-4

für meine Frau Selma und meine Familie

Zusammenfassung

Die vorliegende Arbeit stellt ein neues Konzept eines Feldeffekt-Gassensors mit einem floatenden Gate und vertikalem Auslesetransistor für die Messung von Gaskonzentrationen vor und beschäftigt sich mit dessen Verwirklichung und der Entwicklung eines Herstellungsprozess für einen solchen Sensor. Erstmals wird ein Gassensor, der auf dem Prinzip der Feldeffekttransistoren mit Luftspalt basiert, komplett am Institut für Nanotechnologie und Mikrosystemtechnik der Universität der Bundeswehr hergestellt. Messungen an einem ersten Prototyp werden präsentiert.

Beginnend gibt die Arbeit einen Überblick der Ansätze, mit Hilfe eines Mikrosystems Gaskonzentrationen zu messen. Sensoren, die Transistoren direkt als Transducer einsetzen, erweisen sich als äußerst günstig hinsichtlich des geringen Energieverbrauchs, der schnellen Ansprechzeiten, des Potenzials der weiteren Skalierbarkeit, der Fähigkeit der Integration und schließlich der Aussicht auf geringe Herstellungskosten. Das Prinzip, Gaskonzentrationen mittels eines Transistors mit Luftspalt zu messen, besitzt gegenüber anderen GASFETs den Vorteil, nur Oberflächeneffekte zu messen. Das garantiert schnelle Ansprechzeiten, denn das Gas muss nicht erst in das Volumen eines Materials diffundieren.

Das Konzept des entwickelten FGFET (Floating Gate FET) basiert auf dem HSG-FET (Hybrid Suspended Gate Field Effect Transistor), einem Feldeffekttransistor mit Luftspalt. Diese Bauform erlaubt den Einsatz einer großen Menge sensitiver Materialien. Im Vergleich zeigt sich der FGFET gegenüber dem HSG-FET als aussichtsreicherer Kandidat, weil er die größere Empfindlichkeit bei geringerer Einsatzspannung besitzt. Der Grund dafür ist die große Gate-Kapazität, dank einem dünnen Gate-Oxid des Auslesetransistors.

In Kapitel 3 wird das vertikale Konzept des FGFETs entwickelt. Vertikale Transistoren zeichnen sich gegenüber Lateralen durch sehr kurze Kanallängen aus. Daher ist ein großes Verhältnis von Kanalweite zu -länge unkompliziert zu realisieren. Es bestimmt wesentlich den Signalhub des FGFETs. Die Fertigung des Sensors auf einem SOI-Substrat verbessert die Isolierung. Leckströme, die für das typische Driften der GASFETs verantwortlich sind, werden so vermieden. In diesem Abschnitt wird ein Prozessablauf vorgeschlagen, mit dem sich das Konzept leicht umsetzen lässt. Es wird das Layout für diesen Sensor, sowohl für den Chip als auch den Deckel, entworfen. Verschiedene Designs für den Auslesetransistor werden erarbeitet. Die Kanalweiten der in dieser Arbeit erzeugten Transistoren betragen bis zu dem 2000fachen der Kanallänge.

Kapitel 4 zeigt die Vorzüge von SOI-Wafern auf. Nebenbei werden wegen des Einsatzes dieser speziellen Substrate in Verbindung mit den vertikalen Transistoren keine Implantationswannen benötigt. Daher ist zu erwarten, dass der Temperaturbereich des Sensors über 200 °C hinaus erweitert wird.

Die Realisierung des Sensors erfordert zunächst das passende Substrat. Versuche zeigen, dass auf den in dieser Arbeit verwendeten SOI-Wafern Schichten mit der MBE (Molecular Beam Epitaxy) abgeschieden werden können. Die Qualität des kristallinen Filmes entspricht dabei derjenigen, eines auf herkömmlichen Wafern Gewachsenen.

Die zur Herstellung der Sensoren notwendigen Technologien werden in Kapitel 6 entwickelt. Der vorgestellte Sensor wird aus zwei Teilen als Mikrosystem aufgebaut. Der Erste, der Chip, trägt primär den Auslesetransistor und die Verdrahtung. Auf dem Anderen, dem Deckel, befindet sich hauptsächlich die sensitive Schicht. Beide Teile werden mit der Flip-Chip-Technik aufeinander justiert und verklebt. Wegen der leichteren Handhabung befestigt man diesen Sensor auf ein TO-8 Gehäuse und kontaktiert ihn mit einem Bondverfahren.

In dem hierin vorgestellten Fertigungsprozess wird ein Oxid hergestellt und strukturiert, noch bevor mittels MBE kristalline Schichten aufgewachsen werden. Entscheidend ist hier eine perfekte Fotolackmaske. Dort wo die Schicht epitaktisch gewachsen werden soll, dürfen keine Lackreste zurückbleiben, weil sie ein vollständiges Ätzen des Oxides und freilegen der Siliziumoberfläche verhindern. An diesen Stellen kann kein kristallines Silizium abgeschieden werden. Die Folge sind Versetzungslinien, die darüber befindende Bauelemente schädigen.

Die verwendeten vertikalen Transistoren bestehen aus Mesen. Nach deren Herstellung sind die Wafer topologisch geformt. Das stellt besondere Anforderungen an die folgenden Prozessschritte. Nicht jeder Fotolack eignet sich zur Strukturierung über topologisch geformte Substrate. Ferner mangelt es vielen Fotolacken an Beständigkeit gegen die nasschemischen Ätzlösungen für Polysilizium. Verschiedene Fotolacke werden auf ihre Tauglichkeit zur Strukturierung von Polysilizium über Mesen untersucht. Mit dem entwickelten Prozess ist es möglich, ein Gate an den vertikalen Transistor mit einem nur geringen Überlapp von 1 µm zu legen.

Die Mesen werden mit einem RIE-Ätzprozess erzeugt. Für einen zuverlässigen Prozessablauf ist aber eine maßhaltige Übertragung der Strukturen von der Chrommaske bis in das Silizium notwendig. Durch ein Absenken der Temperatur und Reduzieren des Drucks bis zum unteren Limit erreicht man einen hohen Anisotropiefaktor. Die Qualität der Mesaflanken wird auf diese Weise gegenüber früheren Versuchen sogar erhöht.

Als Metallisierung wird ein Schichtsystem aus Titan, Titannitrid und Aluminium verwendet und mit Sputter-Prozessen abgeschieden. Titan verbessert den elektrischen Kontakt zu n-dotiertem Silizium. Titannitrid verhindert das Phänomen des „Spikings“. Allerdings nur dann, wenn es im richtigen stöchiometrischen Verhältnis und in der richtigen Phase abgeschieden wird. Relevant ist dazu die Abscheidetemperatur. Durch Absenken des Sauerstoffpartialdrucks im Rezipienten, Bias-Sputtern sowie genügend hoch eingekoppelte Leistung können Titan-

nitridschichten mit einer hohen Leitfähigkeit erzeugt werden. Die so erzeugte Metallisierung ist temperaturstabil. Sie erlaubt ein Tempern der Kontakte, was zu einer Verbesserung der Kontaktwiderstände führt. Mit dieser Metallisierung ist der Sensor temperaturstabil bis 450 °C.

Die Strukturierung des Metallisierungssystems erfolgt nasschemisch. Versuche, das Titannitrid mit RCA-1 nasschemisch zu ätzen, zeigen keine Selektivität der Ätzlösung zu Aluminium, wenn es mit ersterem Metall in Berührung steht. Es wird sogar bevorzugt angegriffen, sodass die Metallisierung unbrauchbar ist. Mit Hilfe der APK-Ätze umgeht man das Problem, indem das Aluminium durch Silikatbildung passiviert wird und derart selbst als Ätzmaske herangezogen werden kann. Damit ist es möglich, die Zuleitungen auf dem Chip zuverlässig herzustellen.

Das Design der Deckel hat entscheidenden Einfluss auf die Ausbeute, v. a. während des Klebens. Während des Aufklebens der Deckel ist ein sehr vorsichtiges Aufbringen des Klebstoffs nötig, sonst kann es vorkommen, dass der Kleber in den Luftspalt gezogen wird. Eine Klebersperre wird mit in den Deckel integriert, um das zu verhindern. Ihre Wirkungsweise wird demonstriert. Für den Einsatz eines breiteren Spektrums an Materialien als Deckel wird ein alternativer Kontakt untersucht. Er beweist sich als ausreichend, um eine sensitive Schicht auf ein gewünschtes Potenzial zu legen.

In Kapitel 7 werden drei auf SOI-Wafern realisierte Transistortypen mit unterschiedlichem Layout vermessen. Ihre elektrischen Eigenschaften werden untersucht und verglichen. Wie die Messungen bestätigen, ist es durchaus praktikabel Transistoren mit einem W/L -Verhältnis von 2000 herzustellen. Die Nutzung von Polysilizium als zweite Leiterbahnebene resultiert in einen zu großen Zuleitungswiderstand für die Kontaktierung von Source oder Drain. Dieser Widerstand begrenzt in der Folge die Treiberfähigkeit der Transistoren.

Der fertig gestellte Sensor wird ebenfalls vermessen und charakterisiert. Damit wird der Arbeitspunkt dieses Sensors bestimmt. In diesem besitzt er eine Steilheit an der sensitiven Schicht von $(21,04 \pm 0,13) \mu\text{A}/\text{V}$. Die Messungen belegen dabei eine bemerkenswerte Linearität, denn der Linearitätsfehler ist kleiner als 1 %.

Erste Gaskonzentrationsmessungen wurden mit einem Prototyp durchgeführt. Das Sensorsignal zeigt dabei reversible und reproduzierbare Reaktionen auf die Gaskonzentration. Aufgrund einer schlechten Qualität des Gate-Oxids fließen jedoch Ladungen auf die floatende Gate-Elektrode. Diese Leckströme verändern in unerwünschter Weise ihr Potenzial, worauf das Ausgangssignal des Sensors zurückdriftet. Mit einem verminderten Spannungsabfall über dem Gate-Oxid werden die Leckströme reduziert. Daraufhin werden die für die benutzte sensitive Schicht typischen Signalverläufe aufgezeichnet.

Inhaltsverzeichnis

Zusammenfassung	5
1 Einleitung und Motivation	13
2 Auswahl des Sensorprinzips - Allgemeine Übersicht	17
2.1 Thermische Gassensoren	17
2.1.1 Katalytische Gassensoren	17
2.1.2 Wärmeleitfähigkeitssensoren	18
2.2 Metalloxid-Gassensoren	19
2.3 Festkörperionenleiter	19
2.3.1 Potentiometrischer Sensor	20
2.3.2 Amperometrischer Sensor	20
2.4 Elektrochemische Gassensoren	21
2.5 Optische Gassensoren	21
2.6 Kapazitive Gassensoren	22
2.6.1 Interdigitalkondensatoren	22
2.6.2 MIS-Kapazitäten	23
2.7 Austrittsarbeit	23
2.7.1 Schottky-Dioden	23
2.7.2 GASFET	24
2.8 Sonderformen	24
2.9 Auswahl des Sensorprinzips	25
3 MOS-Transistoren als Gassensoren	27
3.1 Änderung der Einsatzspannung	27
3.2 Der Lundströmsensor	30
3.3 Der HSGFET	32
3.4 Der FGFET	34
3.5 Vergleich der GASFETs	37
4 Entwurf eines Gassensors auf Feldeffektbasis	39
4.1 Architektur des Sensors	39
4.2 Vertikaler Auslesetransistor	39
4.3 SOI-Substrat	43
4.4 Prozessablauf der Herstellung des Gassensors	44

4.5	Layout des Sensors	48
4.5.1	Der Chip	48
4.5.2	Der Deckel	49
4.5.3	Die Transistoren	51
5	FGFET auf SOI	55
5.1	Allgemeine Vorteile der SOI-Substrate für Sensoren	55
5.2	Vorteile der SOI-Substrate für den FGFET	60
5.2.1	Vollständige Isolierung	60
5.2.2	Hochtemperatur-GasFET	60
5.3	Epitaxie auf SOI-Substraten	64
6	Technologieentwicklung	67
6.1	Fotolithografie	67
6.2	Substratreinigung	70
6.3	Oxidherstellung und Strukturierung	73
6.3.1	Oxidherstellung	74
6.3.2	Herstellen der Lackmasken	75
6.3.3	Oxidätzen	77
6.4	Polystrukturierung	80
6.4.1	AR-U 4030	84
6.4.2	XAR-P 3840/2	86
6.4.3	AR-P 3840	89
6.5	Mesaätzen und Freistellen	95
6.6	Metallisierung	100
6.6.1	Entwicklung der Metall-Filme	100
6.6.2	Strukturierung der Metallisierung	112
6.6.3	Kontaktwiderstände	114
6.6.4	Möglichkeiten der weiteren Optimierung	117
6.7	Herstellung der Deckel	118
6.7.1	Ätzen der Klebelaschen	118
6.7.2	Abstandhalter	119
6.7.3	Design der Klebersperre	119
6.7.4	Elektrische Kontaktierung des Deckels	122
7	Charakterisierung	125
7.1	Charakterisierung der Transistoren	125
7.1.1	Transfer- und Ausgangskennlinienfeld	126
7.1.2	Zuleitungswiderstand	126
7.1.3	Volumendiode	131
7.1.4	Einsatzspannung	131
7.1.5	Steilheit	131
7.1.6	Unterschwelligkeit	134
7.2	Charakterisierung des Sensors	136

7.2.1	Transferkennlinien	136
7.2.2	Messung einer Wasserstoffkonzentration	138
8	Ausblick	143
A	Die Abhängigkeit der Lage des Fermi-Niveaus von der Temperatur	147
B	Herstellungsverfahren für SOI-Wafer	149
B.1	Das SIMOX-Verfahren	150
B.2	Gebondete SOI-Wafer	152
B.3	BESOI-Wafer	154
B.4	UNIBOND-Wafer	155
B.5	ELTRAN-Wafer	157
B.6	Vergleich der Herstellungsverfahren	159
C	Prozessablauf der Fotolithografie	163
D	Blende für das Alusputtern	165

1 Einleitung und Motivation

Der Einzug von Elektronik in all unsere Lebensbereiche ist nicht zuletzt durch die fortschreitende Miniaturisierung und Integration begünstigt. Hierfür gibt es verschiedene Gründe. Die Verkleinerung von Geräten aller Art steigert die Marktakzeptanz bei den Endverbrauchern. Zudem lässt sie erst mobile Applikationen zu. Diese Entwicklung wird durch die Integration vereinfacht. Das heißt, mehrere Bauelemente in einem IC (Integrated Circuit) - im Idealfall das gesamte System auf einem Chip - zu realisieren. Dadurch müssen dann weniger Elemente für die entsprechende Schaltung verbaut werden. Das führt nicht nur zur Verkleinerung sondern auch zu einem Sinken der Herstellungskosten.

Neben immer kleineren Geräten wird ein größerer Funktionsumfang verlangt, mit Funktionen, die über den eigentlichen Zweck der Anwendung hinausgehen. Man denke hierbei stellvertretend an Fotohändies oder Höhenmesser in Taschenmessern. Das setzt wiederum voraus, dass die dafür zusätzlich nötigen Sensoren klein genug für den Einbau in das gewünschte System sind. Kaum jemand würde eine Zusatzfunktion des Mobiltelefons akzeptieren, wenn er dieses infolgedessen nicht mehr in die Jackentasche stecken könnte.

Ferner wünscht man sich, alle erdenklichen elektronischen Geräte mobil nutzen zu können. Für deren Energieversorgung stehen aber nur Batterien bzw. Akkus mit einem begrenzten Reservoir zur Verfügung. Die Standortunabhängigkeit erzwingt daher energiesparende Bauteile. So sollte z. B. ein Ozonsensor in einer Uhr dem Träger ständig Informationen über die momentane Umweltsituation liefern, ohne jedoch die Laufzeit der Batterie zu verkürzen.

Wie schon angedeutet, sind von diesem Trend auch Gassensoren betroffen. Nicht nur wegen der Funktionsintegration bedarf es solcher, die miniaturisier- und integrierbarer sind. Gasmessgeräte sollten für einen mobilen und flexiblen Gebrauch an sich schon klein und handlich sein. Eine Lösung für kleine Gassensoren ist die Realisierung als Mikrosystem mit Hilfe der Siliziumtechnologie. Für mobile Anwendungen ist die Auswahl an Gassensoren aber noch nicht groß. Viele heutzutage genutzte Gassensoren, wie z. B. Metalloxid-Gassensoren, müssen geheizt werden und besitzen deshalb einen beträchtlichen Energiekonsum, was dem mobilen Einsatz widerspricht. Optimal wäre ein Sensorprinzip, das mit Hilfe der Siliziumtechnologie hergestellt werden kann und möglichst wenig Energie erfordert. Ein niedriger Preis eröffnet nebenbei weitere Einsatzfelder. Viele Szenarien mobiler Anwendungen sind dabei vorstellbar. Sie reichen von einer quantitativen Erfassung, z. B. in Einweg-Kampfgasmeldern, bis hin zu einer langlebigen qualitativen Auflösung, wie sie in Regelungssystemen benötigt werden.

Diese Zielsetzung fordert die Erforschung und Entwicklung von Gassensoren. Absicht dieser Arbeit ist die mikrosystemtechnische Realisierung eines Sensors für den mobilen Einsatz mit einem hohen Konzentrationsbereich und kurzen Ansprechzeiten bei einem niedrigen Preis.

Diese Arbeit versucht, die Konzepte von Feldeffekttransistoren mit Luftspalt zum Detektieren von Gasen weiterzuentwickeln. In einem neuen Konzept werden dazu vertikale Bauelemente zusammen mit SOI-Substraten verwendet.

Der Transistor dient in dieser Art der Sensoren als auslesendes Element. Er wandelt eine Potenzialänderung in einen Ausgangsstrom um. Welchen Wertebereich das Signal dabei abdeckt, hängt wesentlich von der Stromverstärkung des Transistors ab. Vertikale Bauelemente zeigen ihre Vorzüge im Bereich kurzer Kanallängen. Im Gegensatz zu den lateralen Bauelementen werden diese nicht durch die Fotolithografie beschränkt. Hier wird die Kanallänge viel mehr durch das Abscheideverfahren vorgegeben. Mit der an der Universität der Bundeswehr eingesetzten Molekularstrahlepitaxie (MBE) erzielt man Kanallängen von einigen wenigen 10 nm. Dadurch ist ein großes Verhältnis von Kanalweite zu -länge erreichbar. Dieses sollte tunlichst groß sein, weil es proportional in die Verstärkung eingeht.

Die Schwierigkeiten vertikaler Bauelemente fallen in dieser Sensorkonzeption nicht ins Gewicht. Da nur ein Transistor als Ausleseelement gebraucht wird, lässt sich eine komplizierte Verdrahtung umgehen. Für den Fall das mehrere Sensoren auf einem Chip benötigt werden, müssen nicht unterschiedliche Kanaltypen verwendet werden.

SOI-Substrate (Silicon on Insulator) werden bereits seit Mitte der 60 Jahre verwendet. Ihr wesentliches Merkmal ist ein unterhalb eines dünnen Siliziumfilmes vergrabenes und ganzflächig geschlossenes Oxid. Daraus erwächst eine Reihe von Vorteilen. Hauptsächlich lassen sich damit Bulk-Effekte unterdrücken. Aufgrund der Einführung von SOI-Substraten in dem vorgestellten Konzept, werden die einzelnen Elemente des Sensors nicht mehr durch Sperrschichten der pn-Übergänge isoliert, sondern durch ein Dielektrikum. Das bedeutet eine Verbesserung der Isolierung und Abnahme der Leckströme innerhalb des Sensors. Neben einem geringeren Stromverbrauch, verspricht es ebenfalls für diese Art der Sensoren, eine Stabilisierung der Kennlinien.

Darüber hinaus ist zu erwarten, dass man durch die Kombination von SOI-Substraten und vertikalen Transistoren zusätzlich den Temperaturbereich der Sensoren nach oben ausdehnt. Ein vertikaler Transistor benötigt keine niedrig dotierten Wannengebiete. Der Spannungsteiler in dem vorgestellten Konzept setzt diese ebenso wenig voraus, weil keine der Elektroden mit einer Sperrschicht isoliert werden muss. Die niedrigsten Dotierkonzentrationen die verwendet werden liegen bei 10^{18} cm^{-3} . In diesem Fall hat der Siliziumkristall noch bei mehr als 200 °C halbleitende Eigenschaften. Folglich eröffnen sich für den entwickelten Sensor Anwendungsfälle mit Umgebungstemperaturen, für die bisher kein Feldeffekt-Gassensor zur Verfügung stand. Daneben können somit chemische oder physikalische Re-

aktionen gemessen werden die erst oberhalb dieser 200 °C stattfinden. So werden neue interessante Materialien als sensitive Schichten zugänglich.

In dieser Arbeit wird ein Sensor nach dem neuen hierin vorgestellten neuen Konzept entworfen, der mit Hilfe der Nano- und Mikrosystem-Technologie aufgebaut wird. Dazu werden das Wissen und die Erkenntnisse der Gassensorik und der Nanotechnologie des Instituts für Nanotechnologie und Mikrosystemtechnik der Universität der Bundeswehr zusammengeführt. Der Schwerpunkt dieser Arbeit befasst sich mit den Schwierigkeiten und Problemen bei der Herstellung des Sensors. Es ging nicht darum, jeweils das Optimum für die einzelnen Arbeitsschritte zu finden. Viel mehr dürfen sie sich nicht widersprechen, müssen alle aufeinander abgestimmt sein und im Zusammenspiel funktionieren.

2 Auswahl des Sensorprinzips - Allgemeine Übersicht

Sensoren zur Gasanalyse sind Bauelemente, die in Abhängigkeit eines elektrochemischen oder physikalischen Messeffektes ein elektrisches Messsignal liefern. Dieses Signal steht mit der Konzentration als Messgröße in einem eindeutigen Zusammenhang. Ein Gassensor besteht in der Regel aus einem sensitiven Material, welches mit der zu messenden Substanz eine Wechselwirkung eingeht und dadurch einen Parameter ändert, sowie einem Ausleseelement, dem sog. Transducer, der diesen Parameter in ein elektrisches Signal umsetzt. Von diesen Systemen abzugrenzen sind Analysesysteme, denen das Testgas zugeführt wird. Letztere werden oft schon von den Dimensionen her nicht als Bauelemente sondern als Geräte bezeichnet. Im Gegensatz dazu befindet sich ein Sensor in dem zu untersuchenden Medium. Die zu analysierende Substanz wird durch Diffusion zum sensitiven Material transportiert.

2.1 Thermische Gassensoren

Bei den thermischen Gassensoren gibt es zwei unterschiedliche Typen: Katalytische und Wärmeleitfähigkeitssensoren.

2.1.1 Katalytische Gassensoren

Katalytische Gassensoren eignen sich für den Nachweis brennbarer Gase wie Kohlenmonoxid¹, Wasserstoff, Alkohole und andere Kohlenwasserstoffe. Diese Sensoren messen die Temperaturerhöhung an Katalysatoroberfläche bei der Reaktion brennbarer Gase an ihr.

Typische Vertreter sind die so genannten Pellistoren. Der Aufbau eines solchen Sensors besteht typischerweise aus einem Platindraht als Heizwendel kugelförmig oder perlenartig umgeben von einem keramischen Material. In diesem sind Metalle als Katalysatoren eingelagert. Der Heizdraht erwärmt die Anordnung auf 300 °C bis 600 °C. Durch diese hohen Temperaturen verbrennen die vorhandenen brennbaren Gase an der Oberfläche des Sensors. Die hierbei frei werdende thermische Energie erhöht wiederum die Temperatur der Perle. Diese Temperaturänderung lässt sich aufgrund der Temperaturabhängigkeit des Widerstandes des Platindrahtes elektrisch messen.

¹CO

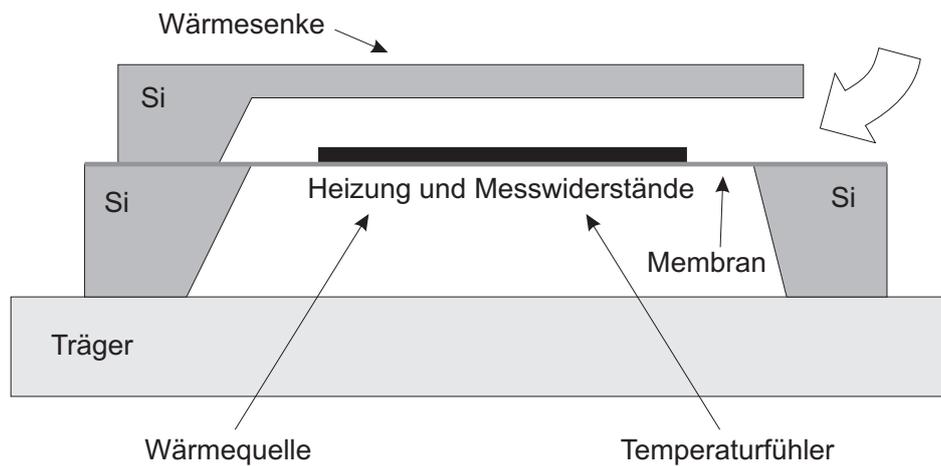


Abbildung 2.1: Aufbau eines mikromechanisch gefertigten Wärmeleitfähigkeitssensors

Über Wärmekonvektion in der Umgebungsluft sowie Wärmeleitung der Luft und des Drahtes wird die Energie des Pellistors abgegeben. Die Temperatur der Perle wird aber nicht allein durch die Reaktion der Gase bestimmt, sondern von Parametern, wie Strömungsgeschwindigkeit, Druck und Temperatur des umgebenden Mediums. Deshalb werden in der Regel zwei Elemente in einer Referenzmessung eingesetzt, z. B. in einer Wheatstone'schen Brücke.

2.1.2 Wärmeleitfähigkeitssensoren

Wärmeleitfähigkeitssensoren werden auch als kalorimetrische Sensoren bezeichnet. Mit diesen wird die thermische Leitfähigkeit von Gasgemischen bestimmt. Deshalb besitzen sie keine sensitive Schicht und es findet keine chemische Reaktion statt. Sie sind daher physikalische Sensoren. Diese Sensoren eignen sich besonders zur Bestimmung des Verhältnisses zweier Gase zueinander in binären Gasgemischen.

Der Aufbau eines Wärmeleitfähigkeitssensors besteht aus einer Wärmesenke, einer Wärmequelle sowie einem Temperaturfühler. Es existieren von diesem Typ auch mikromechanisch gefertigte Bauelemente, deren Aufbau in Abbildung 2.1 beispielhaft gezeigt wird. Auf einer sehr dünnen Membran, meistens aus Siliziumnitrid, werden ein Heiz- und ein Temperatur-Messwiderstand strukturiert. Oft dient der Heizmäander analog wie beim Pellistor gleichzeitig als Temperaturfühler. Die thermische Energie wird über die Wärmeleitung der Luft an den Steg, der hier als Wärmesenke funktioniert, abgeben. Ändert sich die Wärmeleitfähigkeit des eintretenden Gasgemisches, ändert sich auch die Temperatur auf der Membran, was als elektrisches Signal ausgelesen wird. In [PDO93] und [SA02] werden als Mikrosystem aufgebaute Wärmeleitfähigkeitssensoren beschrieben und untersucht.

2.2 Metalloxid-Gassensoren

Metalloxidgassensoren dienen zum Nachweis von reduzierenden oder oxidierenden Gasen. Sie nutzen als Sensoreffekt die Änderung der elektrischen Leitfähigkeit von Metalloxiden in Abhängigkeit des Sauerstoff-Partialdruckes an ihrer Oberfläche. Mit einer Widerstandsmessung wird dieser Effekt als elektrisches Signal ausgelesen.

In einem typischen Aufbau wird ein Metalloxidfilm in Sinter-, Dickschicht-, oder auch Dünnschichttechnik auf eine elektrisch isolierende Schicht aufgetragen, unter der sich ein Heizwiderstand befindet. Um den Widerstand zu messen sind, an der Oxidschicht elektrische Kontakte angebracht. Metalloxide liegen gewöhnlich nicht im stöchiometrischen Verhältnis vor, sondern sind eher sauerstoffarm. Diese Sauerstofffehlstellen wirken wie Donatoren und geben Elektronen an das Leitungsband ab. Diese Oxidfilme verhalten sich also wie n-Leiter. Abhängig vom Sauerstoffpartialdruck des umgebenden Mediums werden diese Fehlstellen und so die Konzentration der Ladungsträger im Leitungsband erniedrigt. Sind reduzierende oder oxidierende Gase vorhanden, wird nun Sauerstoff von der Oberfläche aufgebraucht oder ihr zugeführt. Erstere Gase erhöhen und letztere verringern so die freien Ladungsträger und verändern den Widerstand.

Um die Adsorption unerwünschter Gas zu verhindern, wird der Sensor auf eine Temperatur zwischen 120 °C und 600 °C gebracht. Für eine weitere Erhöhung der Selektivität, werden dem Metalloxidfilm katalytisch wirkende Metalle beigemischt. Zudem ist neben der Wahl des Materials auch die Abscheidetechnik entscheidend für die Sensitivität auf ein bestimmtes Gas. Mikrosystemtechnische Realisierungen sind in [BTBv⁺99, PVC⁺02] beschrieben.

2.3 Festkörperionenleiter

Der Stromfluss durch Keramiken aus Zirkoniumoxid² und Yttriumoxid³ erfolgt bei Temperaturen oberhalb von 350 °C durch Sauerstoffionen⁴. Der Transportmechanismus ähnelt also dem in einem flüssigen Elektrolyten. Auf zwei Oberflächen werden poröse, sauerstoffdurchlässige Elektroden angebracht, die dann das elektrische Signal liefern. Damit die Ionen gut diffundieren können, wird die Keramik auf 450 °C bis 950 °C geheizt. Dieser Effekt lässt sich potentiometrisch oder amperometrisch zur Bestimmung der Sauerstoffkonzentration verwenden. Festkörperionenleiter sind die weitverbreitetsten Gassensoren. In den letzten Jahren wird immer mehr versucht diese Sensoren mit Herstellungsmethoden der Mikroelektronik zu fertigen. Einen genaueren Überblick über diese Aktivitäten ist in [Dub03] zu finden.

²ZrO₂
³Y₂O₃
⁴O²⁻

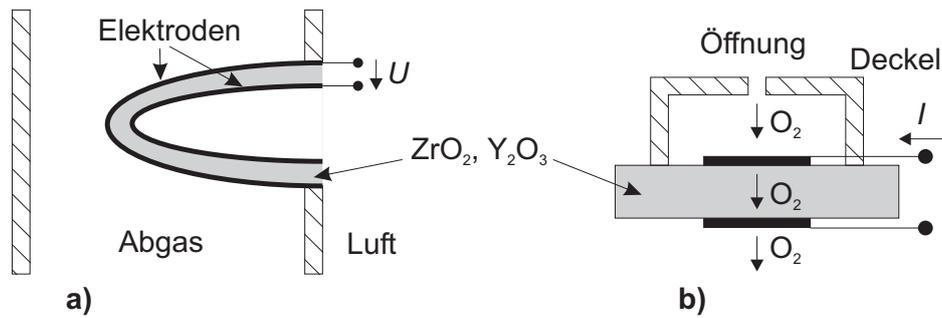


Abbildung 2.2: Prinzipskizze von Festkörperionenleitern als Gassensoren: a) potentiometrischer und b) amperometrischer Sensor

2.3.1 Potentiometrischer Sensor

Der wohl bekannteste potentiometrische Gassensor ist die Lambda Sonde, die zur Regelung des Kraftstoff-Luft-Gemisches für Verbrennungsmotoren genutzt wird. Ihr schematischer Aufbau ist in Abbildung 2.2a zu sehen. Eine Elektrode befindet sich im zu messenden Medium, die andere in einem Referenz-Medium, im Fall der Lambda Sonde in Umgebungsluft. Durch die unterschiedlichen Sauerstoffkonzentrationen kommt es zu einem Sauerstoff-Diffusionsstrom und es bildet sich die Spannung

$$U = U_0 + \left(\frac{R T}{n F} \right) \ln(p_{ref}/p_{gas}) \quad (2.1)$$

aus, wobei R = allgemeine Gaskonstante, T = Temperatur, n = Wertigkeit, F = Faraday Konstante, U_0 = Nullpunktspannung, p_{ref} = Sauerstoffpartialdruck im Referenzgas und p_{gas} = Sauerstoffpartialdruck im Testgas sind. Anstatt des Referenzgases kann auch eine Referenzelektrode, die entweder hermetisch vor dem zu analysierenden Gas versiegelt ist oder aus geeignetem Material besteht, genutzt werden. Dadurch lassen sich die Sensoren sehr einfach in Siliziumtechnologie herstellen.

2.3.2 Amperometrischer Sensor

Das Prinzip eines amperometrischen Sensors zeigt Abbildung 2.2b. Beide Elektroden befinden sich im zu untersuchenden Gas. Über der Kathode wird ein kleiner Hohlraum konstruiert, um den Raum der an Sauerstoffionen verarmt konstant zu halten. Damit möglichst alle Sauerstoffmoleküle, die diese Elektrode erreichen, in Ionen zerlegt werden, bietet sich hier als Elektrodenmaterial Platin an. Legt man eine konstante Spannung an die Elektroden, fließt ein Strom in Abhängigkeit der Konzentration an Sauerstoffionen, weil diese für den Ladungstransport sorgen.

2.4 Elektrochemische Gassensoren

Elektrochemische Gassensoren besitzen einen flüssigen oder gelartigen Elektrolyt, in den das zu messende Gas durch eine semipermeable Membran gelangt. Diese Sensoren haben Empfindlichkeiten im ppm-Bereich. Deshalb eignen sie sich besonders toxische Gase zu detektieren, zur Lecksuche oder Feststellung von Betriebsstörungen.

In dem üblichen Aufbau eines solchen Sensors befinden sich eine Mess- und eine Gegenelektrode in dem Elektrolyt. In ihm läuft eine chemische Reaktion ähnlich wie in einer Batterie ab. Während aber in einer Batterie alle Reaktionspartner vorhanden sind, fehlt in diesem Sensor einer, der für die Reaktion erforderlich ist. Das ist in diesem Fall das zu messende Gas. Es kann durch eine Membran in den Elektrolyten eindiffundieren. Dieses wird an der Messelektrode elektrochemisch umgesetzt, sodass Elektronen über die Elektroden fließen. Der Strom kann ausgelesen und als elektrisches Messsignal erfasst werden, weil die Reaktionsgeschwindigkeit abhängig vom herrschenden Partialdruck ist. Als Gegenelektrode muss ein Material gewählt werden, welches die Messelektrode so polarisiert, dass die elektrochemische Reaktion stattfindet und eine möglichst hohe Selektivität erlaubt. Da dieses nicht immer möglich ist, wird durch das Anlegen einer Spannung von außen die Messelektrode auf ein bestimmtes Potenzial gelegt. Weil Letzteres aber mit der Änderung des fließenden Stromes variiert, wird die Spannung zwischen Messelektrode und einer Referenzelektrode eingestellt.

Viel Aufwand wird beim Herstellen der Diffusionsbarriere betrieben, um lineare Sensoren zu realisieren. Ist die Diffusion durch diese geschwindigkeitsbestimmend für die Reaktion, so ist nach dem Fick'schen Gesetz der messbare Strom proportional zur Konzentration des gesuchten Gases. Da der Elektrolyt oder die Elektroden mit der Zeit verbraucht werden, haben diese Sensoren nur eine begrenzte Lebensdauer, meistens im Bereich von einem Jahr.

2.5 Optische Gassensoren

Optische Sensoren nutzen die Änderung der Absorption bzw. Transmission oder des Brechungsindex einer sensitiven Schicht. Sie haben eine Lichtquelle, die eine solche Schicht als Transducer bestrahlt.

In dem typischen Aufbau wird auf einem elektrischen Fotodetektor ein Glaspättchen mit der gassensitiven Schicht angebracht und dieses System von einer Lichtquelle bestrahlt. Abhängig von der optischen Transmission ändert sich das elektrische Signal des Photosensors. So ein System eignet sich sehr für eine mikrosystemtechnische Fertigung [LFL⁺00].

Ein anderer Ansatz führt das Licht in einem Lichtwellenleiter, dessen Oberfläche sensitiv beschichtet ist. Durch die Variation der optischen Parameter verändert sich die Intensität oder das Spektrum des im Wellenleiter geführten Lichtes. In anderen Anordnungen wird das Ende der Faser beschichtet. Das Licht in ein und

demselben Lichtleiter eingeleitet, am Ende reflektiert, zurückgeführt und das zurückkommende Spektrum ausgewertet. Oftmals wird das Signal nicht elektrisch ausgegeben sondern direkt das optische Signal über einen Wellenleiter weitergeleitet [BTV⁺00].

Eine besondere Art der optischen Sensoren arbeitet mit der Oberflächenplasmonresonanz. Trifft monochromatisches Licht auf eine Grenzfläche von einem optisch dichtem zu einem optisch dünnen Medium unter einem kleineren Winkel als der einer Totalreflexion, wird zwar das gesamte Licht reflektiert, aber es bildet sich trotzdem ein elektrisches Feld in dem optisch dünneren Medium aus. Diese nimmt stark exponentiell mit der Entfernung von Grenzfläche ab. Wird diese Grenzfläche nun mit einem nicht-magnetischen Metall sehr dünn beschichtet durchdringt dieses Feld die leitende Schicht und es werden Oberflächenplasmonen erzeugt. Nur wenn diese Plasmonen in Resonanz mit den Photonen stehen, tritt jetzt eine Reflexion auf. Diese Resonanz ist abhängig von der Wellenlänge und des Auftreffwinkels des Lichtes und auch äußerst stark von dem Brechungsindex an der Oberfläche der Grenzfläche. Der Reflexionswinkel reagiert daher äußerst sensibel auf die Änderung des Brechungsindex einer gassensitiven Schicht. In [MA96] wird von einer Messung dieses Winkels zur Bestimmung von Gaskonzentrationen berichtet.

2.6 Kapazitive Gassensoren

Kapazitive Sensoren messen die Variation der dielektrischen Eigenschaften einer sensitiven Schicht, wenn Gas an der Oberfläche adsorbiert oder im Volumen absorbiert wird. Die Permittivität ändert sich, weil die Schicht elektrische Dipole oder Ionen ein- oder anlagert.

2.6.1 Interdigitalkondensatoren

Eine Form der Realisierungen ist v. a. die Ausführung eines Interdigitalkondensators. Über diese Struktur wird die sensitive Sicht abgeschieden. Die ineinander greifenden Finger dieses Kondensators sind dabei von dem gassensitiven Material umgeben. Das Auslesen der Impedanz liefert das elektrische Signal. Bei Gasbeaufschlagung ändert sich dieses, da sich die Parameter des Dielektrikums verändern. Typischerweise detektieren diese Sensoren Schwefeldioxid und Kohlendioxid.⁵ Diese Bauform lässt sich sehr leicht mit Dünnschichttechniken herstellen und in Mikrosystemen verwenden oder sogar in die Siliziumtechnologie integrieren [EMD92, EDH94].

⁵SO₂ bzw. CO₂

2.6.2 MIS-Kapazitäten

Eine andere häufig angewendete Methode, die dielektrischen Eigenschaften zu ermitteln, ist das Messen der Kapazität einer MIS (Metal Insulator Semiconductor)-Struktur. Der Aufbau besteht, wie der Name schon sagt, aus einem Schichtstapel mit einem Halbleiter (meistens Silizium), einem Isolator (z. B. Siliziumdioxid oder Siliziumnitrid) sowie einem Metall und lässt sich sehr leicht mikrosystemtechnisch realisieren. Als Metallisierung werden gerne Übergangsmetalle (hier typischerweise Palladium) verwendet, da sie das molekulare Gas katalytisch zerlegen. Das zu messende Gas diffundiert durch das Metall und lagert sich an der Metall-Isolator-Grenzfläche an. Diese angelagerten Dipole beeinflussen die Anzahl der Ladungsträger in der Inversionsschicht und Verarmungszone. Das kann anhand der Strom-Spannungskennlinie oder der Kapazitäts-Spannungskennlinie gemessen werden. Typischerweise werden mit solchen Sensoren Wasserstoff oder Wasserstoff enthaltende Gase detektiert, weil dieser sehr leicht durch Palladium wandert. Beispiele dafür finden sich z. B. in [LGE⁺99, ÅCS⁺02]. Um die Diffusion zu erleichtern, versucht man das Metall als poröse Schicht abzuscheiden. So können auch Konzentrationen anderer Gase, wie z. B. Stickstoffdioxid⁶ gemessen werden [KLK⁺01]. Da ebenfalls die Dicke und die spezifische katalytische Wirkung des Materials der Metallisierung ausschlaggebend für die Selektivität sind, werden auch andere Metallisierungssysteme untersucht [ÅCS⁺02].

2.7 Austrittsarbeit

Lagern sich an einer Metalloberfläche Ionen oder Dipole an, wird die Austrittsarbeit des Materials verändert. Passiert das Gleiche an einer Halbleiteroberfläche, verschiebt sich dessen Fermi-Niveau. Dieser Effekt tritt auch unter Gasbeaufschlagung auf. In diesem Fall sorgen Gasmoleküle oder -ionen für die Änderung. Materialien, die nur selektiv die Adsorption von bestimmten Gasen erlauben, können zur Bestimmung von Gaskonzentrationen in Sensoren eingesetzt werden.

2.7.1 Schottky-Dioden

Kommt ein Metall in direkten Kontakt mit einem Halbleiter, entsteht die Schottky-Barriere. So ein Bauelement hat eine gleichrichtende Wirkung. Die Höhe der Barriere ergibt sich maßgeblich durch die Differenz der Austrittsarbeit aus dem Metall und dem Fermi-Niveau des Halbleitermaterials und bestimmt die Strom-Spannungskennlinie der Diode. Diese variiert folglich bei einer Abhängigkeit der Austrittsarbeit mit der Gaskonzentration. Das kann anhand des Stromflusses durch die Diode beobachtet werden. Diese Sensoren werden z. B. in [CBL00] und [KLL⁺00] beschrieben. Diese Sensoren werden meistens monolithisch gefertigt und lassen sich von daher schon sehr einfach integrieren.

⁶NO₂

2.7.2 GASFET

Ein Gassensor, der einen Transistor als Signalumsetzer nutzt, wird im Allgemeinen als GASFET (Gas Sensitive Field Effect Transistor) bezeichnet. Er nutzt ganz ähnlich, wie die MIS-Kapazität, die Adsorption von Dipolen oder Ionen an einer Metall-Halbleiter-Grenzfläche. Allerdings wird mit diesem Sensortyp die Auswirkung der Gasanlagerung auf die Leitfähigkeit im Kanal des Transistors gemessen. Der Aufbau gleicht sehr dem eines herkömmlichen MOS-Transistors (Metal Oxide Semiconductor) und unterscheidet sich nur durch den Gate-Schichtstapel. Diesem Sensor gibt es in zwei unterschiedlichen Varianten. Die ursprüngliche Anordnung, der ISFET (Ion Sensitive Field Effect Transistor) geht auf Lundström zurück [LSSL75]. Über dem Kanal liegen ein Dielektrikum und darüber eine katalytisch wirkende Gate-Metallisierung. Wie auch bei den MIS-Kapazitäten und Schottky-Dioden muss das Gas durch die Metallisierung diffundieren, um sich an der Grenzfläche zwischen Isolator und Metall anlagern zu können. Diesen Nachteil versucht der SGFET (Ssuspended Gate Field Effect Transistor) zu umgehen. Diese Bauform geht auf Janata [Jan83] zurück. Hier befindet sich ein sehr kleiner Luftspalt zwischen dem Isolator und der Metallisierung. In diesen kann das Gas direkt eintreten, was schnellere Ansprechzeiten ermöglicht. Da diese Sensoren im Grunde schon Halbleiter-Bauelemente sind, liegt es nahe sie mit Methoden der Mikrosystemtechnik zu fertigen.

2.8 Sonderformen

Neben den bisher vorgestellten Methoden Gaskonzentrationen zu messen gibt es Versuche weitere physikalische und chemische Effekte für mikrosystemtechnisch hergestellte oder miniaturisierte Sensoren zu nutzen.

Schon seit einiger Zeit werden massensensitive Gassensoren untersucht. Sie detektieren die Massenzunahme sensitiver Schichten. Als Mikrosystem hergestellte Sensoren nutzen Schwingquarze, SAW (Surface Acoustic Wave) Elemente oder in Siliziumtechnologie gefertigte Kantelever. Bei den Schwingquarzen wird die sensitive Schicht auf den Quarz aufgebracht. Eine Variation der Gaskonzentration verändert die Masse und beeinflusst dadurch die Federkonstante des Quarzes und somit die Frequenz des Oszillators. In einer Untersuchung eines solchen Sensors wird diese Frequenzänderung direkt als elektrisches Signal ausgelesen werden [EMD92].

Ein anderer Weg versucht den Zusammenhang zwischen Gaskonzentrationen und dem Volumen von bestimmten Stoffen zu nutzen. So dehnt sich z. B. Palladium aus, wenn man es Wasserstoff aussetzt, weil dieser in das Metall hinein diffundiert. In [BFK⁺03] werden Mikrokantelever mit diesem Metall beschichtet und die Verbiegung infolge der Scherkräfte durch die Volumenänderung der obersten Schicht kapazitiv ausgewertet. Ganz ähnlich verändern Polymere ihr Volumen mit dem pH-Wert. Bringt man dieses Material mit einer Chemikalie zusammen, die

selektiv mit einem Gas reagiert und dadurch den pH-Wert verschiebt, ist eine Gaskonzentration messbar. Verkapselt man das Volumen, steigt der Druck in diesem an. [HOBV04] stellt einen Gassensor für Kohlendioxid⁷ vor, der mit Hilfe eines Mikro-Drucksensors diese Änderung ausliest.

2.9 Auswahl des Sensorprinzips

Die vorrangigen Anforderungen, die an einen Gassensor gestellt werden, sind Sensitivität, Selektivität und Stabilität. Generell sollen diese drei Eigenschaften so gut wie möglich sein. Je nach Einsatz und Verwendung des Sensors treten noch weitere Merkmale hinzu, die für oder gegen ein bestimmtes Sensorprinzip sprechen. Im Allgemeinen werden folgende Kriterien herangezogen, um einen Sensor zu bewerten:

- Sensitivität
- Selektivität
- Stabilität
- Lebensdauer
- Querempfindlichkeit
- Ansprechzeiten
- Temperaturbereich
- Robustheit
- Energieverbrauch
- Anschaffungs- und Betriebskosten

Bisher existiert kein Prinzip, das in allen Punkten allen anderen überlegen ist. Daher müssen sie für jede einzelne Anwendung nach den individuellen Anforderungen gewichtet werden. Eine allgemein gültige Bewertung ist nicht zulässig.

Ausgehend von der Idee einen Sensor für mobile Anwendungen zu schaffen, ist es entscheidend, ein Prinzip auszuwählen, das nur wenig Energie verbraucht. Damit scheidet die Sensoren aus, die permanent geheizt werden müssen, also thermische Gassensoren, Metalloxid-Gassensoren sowie Festkörperionenleiter. Ferner sollte das Prinzip ein weiteres Miniaturisieren zulassen. Optische Sensoren benötigen z. B. eine Lichtquelle, die mit viel Aufwand justiert und eingestellt werden muss. Es ist schwierig diese so zu verkleinern, dass sich der gesamte Aufbau in anderen Applikationen (wie z. B. Händies, Uhren) integrieren lässt. Um weiterhin kurze Ansprechzeiten zu realisieren, eignen sich Oberflächenreaktionen besser als

⁷CO₂

Volumenreaktionen. Das prädestiniert das Prinzip der Messung der Änderung der Austrittsarbeit mittels eines GASFETS mit Luftspalt. Im folgenden Abschnitt wird dieser Sensortyp noch genauer betrachtet werden.

3 MOS-Transistoren als Gassensoren

In diesem Kapitel werden die verschiedenen Wege aufgezeigt mit einem MOS-Transistor Gaskonzentrationen zu messen. Diese verschiedenen Ansätze entstanden im Verlauf einer Entwicklung angefangen vom Lundströmsensor bis hin zum FGFET (Floating Gate FET). Anhand dieses historischen Fortschritts werden die GASFET-Varianten mit jeweils ihrem Aufbau, ihrer Funktionsweise und ihren speziellen Schwierigkeiten dargestellt.

3.1 Änderung der Einsatzspannung

Befindet sich eine Metall-Isolator-Halbleiter-Struktur eines MOSFET (Metal Oxide Semiconductor Field Effect Transistor) im thermodynamischen Gleichgewicht, sind die Austrittsarbeit W_A des Metalls und das Fermi-Niveau W_F im Volumen des Halbleiters (sog. Bulk) gleich. Abbildung 3.1a stellt das dazugehörige Banddiagramm am Beispiel eines p-dotierten Halbleiters dar. Die Differenz φ_{MS} der Austrittsarbeiten von Metall und Halbleiter fällt nicht nur über dem Isolator, sondern auch über einen kurzen Bereich im Halbleiter ab. Darum sind die Bänder des Halbleiters in diesem Zustand zur Grenzfläche hin verbogen. Solange das Fermi-Niveau W_F überall unter dem intrinsischen Energieniveau W_i liegt, ist das Bulk p-leitend und der Transistor sperrt.

Der Ausgangstrom eines MOSFET hängt maßgeblich von der Einsatzspannung U_T ab. Man definiert sie allgemein als die kleinste Spannung, die am Gate anliegen muss, damit sich im Kanal eine Inversionsschicht ausbildet. D. h. die Anzahl der Minoritätsladungen im Kanal entspricht derjenigen, der Majoritätsladungen im Volumen des Substrates. In Abbildung 3.1b ist für diesen Fall das Banddiagramm gezeigt. An der Grenzfläche befindet sich das Fermi-Niveau W_F über dem Intrinsischen W_i . In diesem Bereich ist der Halbleiter nun n-leitend. Erhöht man die Spannung, vergrößert man nur die Ladungsträger in der Inversionsschicht, nicht aber die Raumladungszone [Gro67]. Gleichung (3.1) gibt die Einsatzspannung U_T für einen Transistor analytisch an.

$$U_T = U_{FB} + 2\varphi_F + \frac{1}{C_i'} \sqrt{2\varepsilon\varepsilon_0 q N_{A,D} (2\varphi_F - U_{SB})} \quad (3.1)$$

- C_i' - Flächenkapazität des (gesamten) Isolators
- $N_{A,D}$ - Majoritätsladungsträgerdichte
- q - Elementarladung
- U_{FB} - Flachbandspannung

- U_{SB} - Substratvorspannung
 ϵ - relative Dielektrizitätskonstante
 ϵ_0 - absolute Dielektrizitätskonstante
 φ_F - Fermipotenzial im Volumen

Die zwei letzten Summanden der Gleichung (3.1) zeigen die Abhängigkeit der Einsatzspannung von den Materialeigenschaften des Halbleiters und der Spannung zwischen Bulk und dem Source-Kontakt. Die Flachbandspannung U_{FB} beschreibt den Einfluss der Eigenschaften des Gate-Kondensators auf die Einsatzspannung und wird durch die Differenz φ_{MS} der Austrittsarbeiten aus Metall und Halbleiter sowie den Ladungen Q_i in den einzelnen Schichten des Dielektrikums mit der Kapazität C_i bestimmt:

$$U_{FB} = \varphi_{MS} - \sum_{i=0}^n \frac{Q_i}{C_i}. \quad (3.2)$$

Liegt die Spannung U_{FB} an der Gate-Elektrode an, herrscht die Flachbandbedingung, d. h. dieses Potenzial kompensiert genau die unterschiedlichen Austrittsarbeiten der Materialien und die Ladungen im Dielektrikum. Außerdem sind die Ladungsträgerkonzentrationen im Halbleiter allorts gleich. Die Bänder sind nun flach. Abbildung 3.1c zeigt das Banddiagramm für genau diese Situation. Aus dieser Betrachtung wird klar, dass Änderungen der Austrittsarbeiten, der Ladungen zwischen Elektrode und Halbleiter, oder auch eine Änderung des Dielektrikums eine Potenzialänderung $\Delta\phi$ zur Folge haben, die sich direkt auf die Einsatzspannung auswirkt:

$$\Delta U_T = \Delta\phi = \Delta\varphi_{MS} + \Delta\left(\sum_{i=0}^n \frac{Q_i}{C_i}\right). \quad (3.3)$$

Die Ausgangskenngröße eines Transistors, der Strom I_{DS}

$$I_{DS} = \begin{cases} \mu C_i' \frac{W}{L} \left((U_{GS} - U_T) U_{DS} - \frac{U_{DS}^2}{2} \right) & \text{für } U_{GS} - U_T > U_{DS} \\ \mu C_i' \frac{W}{2L} (U_{GS} - U_T)^2 & \text{für } U_{GS} - U_T \leq U_{DS} \end{cases} \quad (3.4)$$

ist wiederum von der Einsatzspannung U_T abhängig. Hierbei stellen U_{GS} , U_{DS} die Spannungen von Gate nach Source bzw. von Drain nach Source dar. Im Falle der hier besprochenen GASFETs interessiert nur der obere Zweig der Gleichung (3.4), der Widerstandsbereich. Denn dort zeigt der Transistor seine höchsten Steilheiten, was eine große Signalverstärkung des Sensors bedeutet. Für Sensoren fordert man darüber hinaus Linearität. Folglich betreibt man den Sensor bei einer geringen Drain-Source-Spannung U_{DS} , d. h., der quadratische Term kann vernachlässigt werden. Es gilt also die Näherung

$$I_{DS} \approx \mu C_i' \frac{W}{L} (U_{GS} - U_T) U_{DS} \quad \text{für } U_{DS} \ll U_{GS} - U_T, \quad (3.5)$$

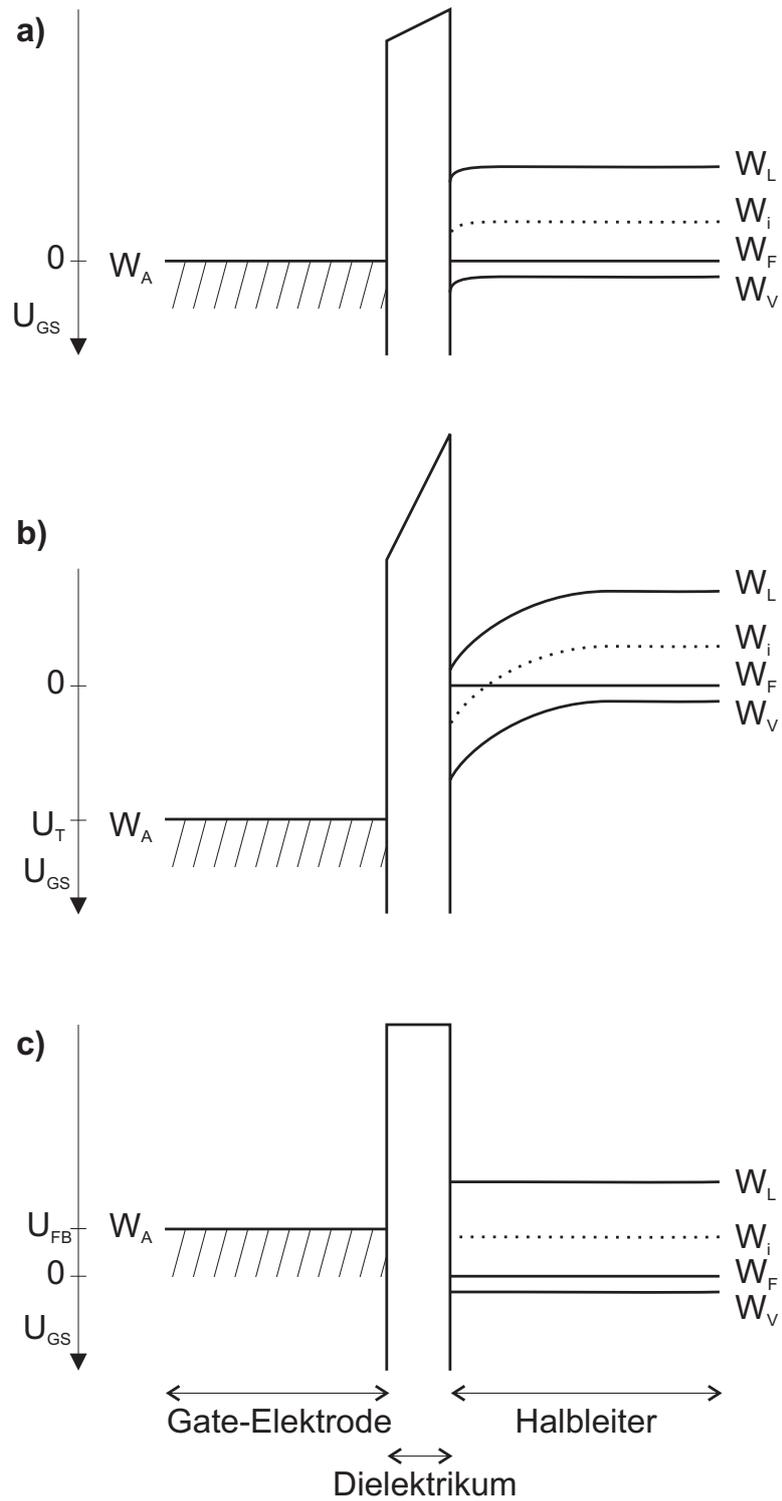


Abbildung 3.1: Bänderdiagramme eines MOS-Transistors in verschiedenen Zuständen: a) im thermodynamischen Gleichgewicht; b) bei anliegender Einsatzspannung; c) bei anliegender Flachbandspannung

wenn die Drain-Source-Spannung U_{DS} viel kleiner als die Differenz aus U_{GS} und U_T ist. Um die Einsatzspannungsänderung ΔU_T zu messen, bietet es sich an den Transistor in einem Regelkreis mit einer Festwertregelung zu betreiben. Dabei wird die Gate-Elektrode so angesteuert, dass der Ausgangsstrom I_{DS} konstant gehalten wird:

$$I_{DS} = const. = \mu C_i' \frac{W}{L} (U_{GS} + \Delta U_{GS} - U_T - \Delta U_T) U_{DS} . \quad (3.6)$$

Wie aus Gleichung (3.6) hervorgeht, muss dazu die Einsatzspannungsänderung ΔU_T durch eine gleich große Spannungsänderung ΔU_{GS} kompensiert werden. Aus dem Regelkreis erhält man mit der Regelgröße ΔU_{GS} eben diese Änderung, die mit der Potenzialänderung $\Delta \phi$ identisch ist:

$$\Delta U_{GS} = \Delta U_T = \Delta \phi . \quad (3.7)$$

3.2 Der Lundströmsensor

Wie vorher bereits im Abschnitt 2.7 erwähnt, eignen sich MOSFETs zur Messung von Gaskonzentrationen. Der erste Sensor dieser Art wurde von Lundström beschrieben [LSSL75]. Man bezeichnet ihn als ISFET (Ion Sensitive Field Effect Transistor) oder oft auch als Lundströmsensor. Abbildung 3.2 zeigt den Aufbau dieses Sensors, der ziemlich stark dem eines Transistors in integrierten Schaltungen angelehnt ist. Er basiert auf einem n-Kanal-Transistor mit einem 10 nm starken Gate-Dielektrikum aus Siliziumdioxid. Als Gate-Elektrodenmaterial wird Palladium verwendet. Diese Art der Sensoren wurde auch in ähnlichen Konfigurationen untersucht. Das Substrat besteht im Allgemeinen aus einem p- oder n-leitenden Silizium. Er besitzt dotierte Source- und Drain-Gebiete und einen dazwischenliegenden Kanalgebiet. Darüber befindet sich ein Isolator, der in der Regel aus Siliziumdioxid hergestellt wird. Um die Stabilität des Bauteiles zu verbessern, wurden aber auch schon andere Materialien, wie Aluminiumoxid und v. a. Siliziumnitrid untersucht. Als Material der Gate-Elektrode finden in den meisten Fällen Platinmetalle statt des in der CMOS-Technologie (Complementary Metal Oxide Semiconductor Technologie) üblichen Polysilizium Verwendung.

Um das genaue Funktionsprinzip des Sensors zu erörtern, müssen die Vorgänge in der Gate-Elektrode und an der Grenzfläche zwischen dieser und dem Isolator betrachtet werden. Gelangen Wasserstoff-Moleküle an die Grenzfläche zwischen Luft und Metall, so werden diese dissoziiert. Der Wasserstoff diffundiert anschließend an die andere Grenzfläche des Gate-Isolators und bildet dort eine Dipolschicht. Diese Ladungen werden im Halbleiter durch Gegenladungen kompensiert. Je nachdem, ob sich der Transistor in Verarmung oder in Inversion befindet, geschieht dieses durch Änderung der Raumladungszoneweite bzw. durch eine Erhöhung oder Verringerung der Ladungsträger in der Inversionsschicht. Wasserstoffionen im Inneren der Palladium-Gate-Elektrode werden durch eine Elektronenwolke neutralisiert und abgeschirmt. Die Debye-Hückel-Länge gibt an, wie weit die Felder einer

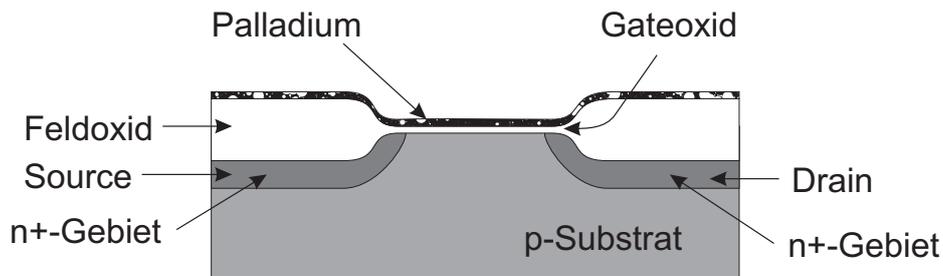


Abbildung 3.2: Schematische Darstellung eines ISFET („Lundström-Sensor“)

Ladung in einem Festkörper wirken und damit, wie groß diese Elektronenwolke ist. Diese ist eine Materialkonstante. Sie liegt für Metalle ungefähr im Bereich von einem Nanometer. Deshalb spielen Ladungen im Inneren der Gate-Elektrode keine Rolle. Die Dipolschicht an der Grenzfläche zwischen Metall und Halbleiter verändert die Flachbandspannung und somit auch die Einsatzspannung (siehe Gleichung (3.3)). Diese Kenngrößen sind proportional zum Bedeckungsgrad. Somit erhält man also mit diesem Sensor einen linearen Zusammenhang zwischen der Wasserstoffkonzentration und der Einsatzspannung.

Das Gate-Material des ISFETs muss für die zu detektierenden Gase permeabel sein. Speziell für die Wasserstoffdetektion bietet sich Palladium an, da es die Eigenschaft besitzt, für Wasserstoffionen durchlässig zu sein. Zudem ist die Verwendung von Platinmetallen vorteilhaft, weil sie das Dissoziieren der Gase katalytisch unterstützen. Damit die Ionen leichter durch das Palladium wandern können, wird dieser Sensortyp meistens leicht geheizt, denn die Diffusionsgeschwindigkeit steigt mit der Temperatur. Auf diese Weise kann man die Ansprechzeiten reduzieren. Der von Lundström beschriebene Sensor wird z. B. bei einer Temperatur von 150 °C betrieben [LSSL75]. Auch andere Materialien wurden untersucht, ob sie als sensitive Schicht nutzbar sind und wie sie mit speziellen Verfahren für Gase permeabel gemacht werden können [HEÅL00]. Möglichkeiten sind unter anderem spezielle Sputterverfahren zur Erzeugung von porösem Metall, plötzliche thermische Belastungen, um die Bildung von Mikro-Rissen zu erzielen sowie auch die Strukturierung von kleinen Löchern in die Elektrode. Auch poröses Silizium wurde zur Erhöhung der Durchlässigkeit verwendet.

Die Diffusion wird unter Umständen durch die ungewollte Anlagerung von Gasen, die sich ebenfalls in dem umgebenden Medium befinden, behindert. In [EE97] wird z. B. die Empfindlichkeit eines solchen Sensors bei gleichzeitigem Vorhandensein von Kohlenmonoxid untersucht. Es wird berichtet, dass das adsorbierte Kohlenmonoxid an der Grenzfläche zwischen Luft und Palladium die Adsorption und Desorption von Wasserstoff unterdrückt und die Diffusion dessen durch die Gate-Elektrode verhindert. Das kann zu einem vollständigen Verlust der Sensitivität des Sensors führen.

Ein Vorteil dieses Sensors ist der geringe Einfluss von an der Oberfläche der sensitiven Schicht adsorbierten Gasen. Wie oben bereits erwähnt, wirken Ladungen

in Metallen nur ungefähr ein Nanometer. Deshalb werden die Ladungen der von der Gate-Elektrode selbst abgeschirmt.

3.3 Der HSGFET

Viele Probleme des ISFETS ergeben sich daraus, dass die Ionen des zu detektierenden Gases erst durch ein Material diffundieren müssen. Das legt es nahe, die Oberfläche der sensitiven Schicht frei zugänglich zu gestalten.

Das führt zu einem Aufbau mit einem Luftspalt zwischen dieser und dem Gate-Dielektrikum. Die Gasmoleküle oder -Ionen werden so direkt von der Oberfläche aus der Gasphase heraus adsorbiert. Die angelagerten Teilchen bilden auf oder an dieser eine Dipolschicht aus, die dazu einer Potenzialänderung $\Delta\phi$ führt. Man bezeichnet diese Sensoren allgemein als MAISFET (Metal Airgap Insulator Semiconductor Field Effect Transistor). Zum ersten Mal wurde dieses 1983 in Form eines monolithisch hergestellten SGFET (Ssuspended Gate FET) vorgestellt [Jan83]. Bei diesem Sensortyp wird das Gate als Brücke über dem Isolator und Kanal gefertigt, um den Luftspalt zu realisieren. Für die Herstellung dieser „Gate-Brücke“ wird meistens ein sog. „Spacer“ (ein Platzhalter) verwendet, der nach dem Aufbringen des sensitiven Materials wieder herausgelöst wird. Das begrenzt allerdings die Auswahl der Materialien, denn der Isolator und die sensitive Schicht dürfen nicht wesentlich geätzt werden und keine chemische Verbindung mit dem Spacer eingehen. Die Probleme dieser Technologie werden in [Fli93] beschrieben.

Der hybrid gefertigte HSGFET (Hybrid Ssuspended Gate FET) unterliegt diesen Einschränkungen nicht. Dieser wurde an der Universität der Bundeswehr in München entwickelt. In diesem Konzept werden Transistor und Gate-Elektrode getrennt gefertigt und erst im letzten Arbeitsschritt so zusammen gefügt, dass ein Luftspalt entsteht. Dieses Vorgehen ist vorteilhaft, weil auf diese Weise Materialien als sensitive Schichten zum Einsatz gelangen können, die bestimmte komplizierte Abscheideverfahren erfordern oder nicht CMOS-kompatibel sind, obwohl der Transistor selbst in CMOS-Technologie hergestellt wird. Abbildung 3.3 zeigt den Aufbau schematisch. Der Transistor im unteren Teil gleicht sehr dem ISFET. Gas, das in den Luftspalt eintritt, wird von der sensitiven Schicht adsorbiert. Ebenso kann aber auch das Gate-Oxid dieses adsorbieren, was genauso zu einer Änderung der Einsatzspannung führt. Demzufolge muss eine Passivierung verwendet werden, die chemisch inert ist. In der Praxis hat sich dafür Siliziumnitrid als günstig erwiesen. Deshalb besteht das Gate-Dielektrikum aus einem Schichtstapel aus Siliziumoxid und -nitrid. Die Gate-Elektrode, oft als Deckel bezeichnet, wird unabhängig davon gefertigt und zum Schluss über dem Transistor mit Hilfe der Abstandhalter positioniert. Damit er stabil sitzt, wird er verklebt. Es gibt aber auch Konstruktionen, die ihn mit einem Drahtbügel oder einer Schraube festklemmen. Neben dem Tragen der sensitiven Schicht, schützt der Deckel gleichzeitig den Kanal vor Lichteinfall.

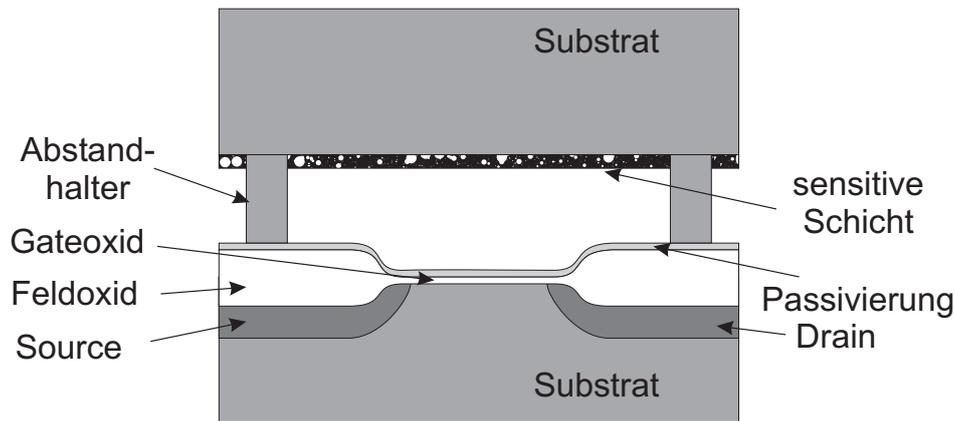


Abbildung 3.3: Schematisch Darstellung eines HSGFET

Wie aus Abbildung 3.3 ersichtlich ist, ergibt sich die Gate-Kapazität C_{ges} hier aus den Kapazitäten des Gate-Dielektrikums (einschließlich der Passivierung) C_i , des Luftspaltes C_L und der sensitiven Schicht C_S als serielle Schaltung:

$$\frac{1}{C_{ges}} = \frac{1}{C_i} + \frac{1}{C_L} + \frac{1}{C_S} . \quad (3.8)$$

Setzt man

$$C = \epsilon_0 \epsilon \frac{A}{d} \quad (3.9)$$

in Gleichung (3.8) ein (dabei ist ϵ die relative Dielektrizitätskonstante, ϵ_0 die absolute Dielektrizitätskonstante und A die Fläche), so erhält man nach Umformung die Flächenkapazität

$$C' = \frac{C_{ges}}{A} = \frac{\epsilon_0}{\frac{d_{Ox}}{\epsilon_{Ox}} + \frac{d_S}{\epsilon_S} + d_L} . \quad (3.10)$$

Die Weite des Luftspaltes d_L ist mit 1 bis 2 μm wesentlich größer als die Schichtdicken des Dielektrikums d_{Ox} und der sensitiven Schicht d_S . Die relative Dielektrizitätskonstante der Luft ist ungefähr 1 und damit kleiner als die der anderen beiden Kapazitäten ϵ_{Ox} sowie ϵ_S . Deshalb bestimmt hauptsächlich der Luftspalt die Gesamtkapazität C_{ges} . Diese wird immer klein bleiben, selbst wenn man die Dicke des Gate-Oxides noch wesentlich reduziert. Das geht auch aus Gleichung (3.10) hervor. Den Luftspalt selbst sollte man nicht weiter verkleinern, weil sonst der Austausch mit der Umgebungsluft durch zu lange Laufzeiten der Gasmoleküle behindert wird.

Für einen Kondensator gilt allgemein die Beziehung:

$$Q = C \cdot U \quad (3.11)$$

für Ladung Q , Kapazität C und Spannung U . So beeinflusst auch die Gate-Kapazität eines Transistors die Ladung in dessen Kanal. Da diese bei einem HSGFET durch

den großen Luftspalt sehr klein ist, resultiert daraus eine hohe Einsatzspannung U_T (Letztere liegt hierbei zwischen 20 und 30 V, siehe auch [Bög98]) und eine geringe Steilheit

$$g_m = \mu C' \frac{W}{L}, \quad (3.12)$$

die als Proportionalitätskonstante in die Ausgangskennlinie Gleichung (3.4) und (3.5) des Sensors eingeht.

3.4 Der FGFET

Mit dem sog. FGFET (Floating Gate FET) wird das Konzept des HSGFET erweitert. Hier sind der Transistor und der gassensitive Kondensator räumlich voneinander getrennt. Der Aufbau wird in Abbildung 3.4 skizziert. Über den dotierten Wannen für die Source- und Drain-Gebiete sowie der Referenzelektrode liegt ein Dielektrikum. Dieses ist unterschiedlich dick, am dünnsten über dem Kanal des Transistors. Die Gate-Elektrode wird aus den vorher genannten Gründen wieder von einer Passivierung bedeckt. Auch bei dieser Konzeption ist die sensitive Schicht für Gase frei zugänglich.

Bei diesem Sensor wird im Gegensatz zu HSGFET und ISFET die Spannungsänderung der sensitiven Schicht nicht direkt an den Transistor weitergegeben. Hier wird das Gate des Transistors an den Mittelabgriff eines kapazitiven Spannungsteilers kontaktiert. Abbildung 3.5 gibt das Ersatzschaltbild für diese Anordnung an.

Die beiden Kapazitäten C_1 und C_2 nutzen kollektiv das Floating Gate als Kondensatorplatte. Diese bildet vereint mit der Passivierung, dem Luftspalt, der sensitiven Schicht und schließlich der Referenz-Elektrode 1 den Kondensator C_1 und zusammen mit der Referenz-Elektrode 2 den Kondensator C_2 . Elektrisch wirken sie wie zwei Kondensatoren in Serienschaltung. Aus Gründen der Ladungsgleichheit gilt für diese

$$C_1(U_1 + \Delta\phi - U_G - \Delta U_G) = C_2(U_G + \Delta U_G - U_2) \quad (3.13)$$

unter der Voraussetzung, dass die Gate-Kapazität vernachlässigbar klein ist. Daher liegt am Gate des Transistors die Spannung

$$U_G = \frac{C_1 U_1 + C_2 U_2}{C_1 + C_2} \quad \text{für} \quad \Delta\phi = 0 \quad \text{und} \quad \Delta U_G = 0 \quad (3.14)$$

an, unter der Bedingung, dass noch keine Änderung an der Schicht eingetreten ist. (Damit hat dann auch ΔU_G den Wert 0, weil das ja die Änderung des Potentials repräsentiert.) In dieser Schaltung führt eine Potenzialänderung $\Delta\phi$ an der sensitiven Schicht zu einer Änderung der Gate-Spannung

$$\Delta U_G = \frac{C_1}{C_1 + C_2} \Delta\phi. \quad (3.15)$$

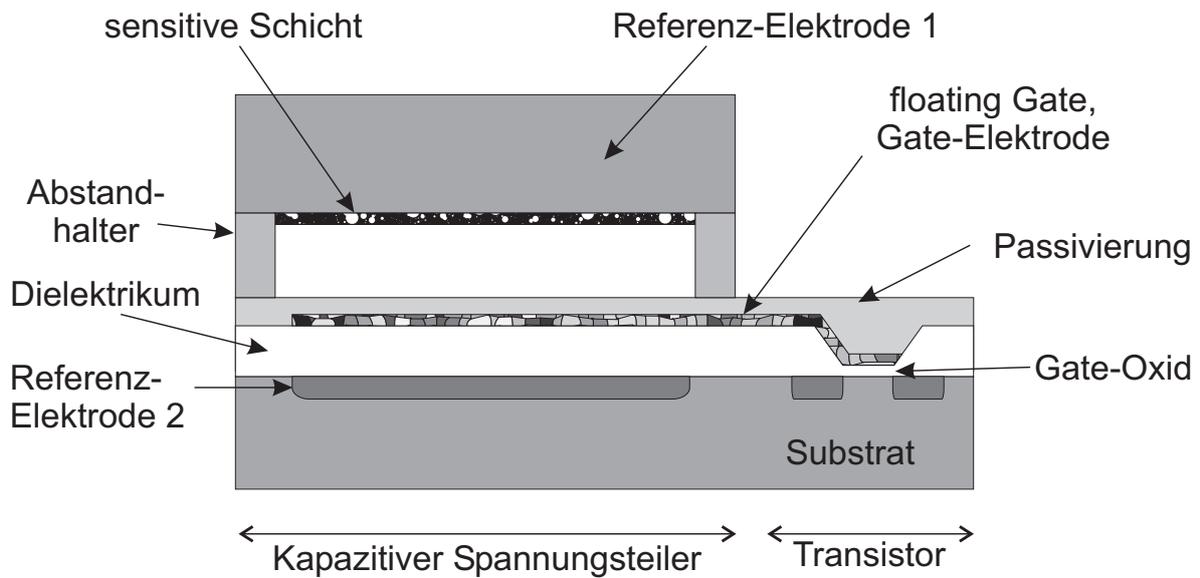


Abbildung 3.4: Schematische Darstellung eines FGFET

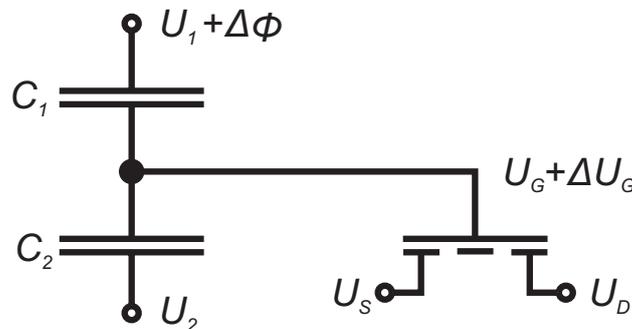


Abbildung 3.5: Ersatzschaltbild eines FGFET

Analog zu den Kennlinien des ISFET und HSGFET (Gleichung (3.6)) kann man für diesen Sensor den Strom I_{DS} des Transistors als Ausgangsgröße beschreiben:

$$I_{DS} = \mu C_i' \frac{W}{L} \left(\left(\frac{C_1 U_1 + C_2 U_2}{C_1 + C_2} + \frac{C_1}{C_1 + C_2} \Delta\phi - U_T \right) U_{DS} - \frac{U_{DS}^2}{2} \right), \quad (3.16)$$

in dem man die Gleichungen (3.14) und (3.15) entsprechend einsetzt. Eine Potenzialänderung $\Delta\phi$ an der sensitiven Schicht hat keinen Einfluss auf die Flachbandspannung U_{FB} oder Einsatzspannung U_T des Transistors, da alle Größen, die in die Gleichungen (3.1) und (3.2) eingehen, durch den Herstellungsprozess festgelegt sind. Trotzdem lässt sich hier genauso der oben erwähnte Regelkreis zur Bestimmung einer Potenzialänderung $\Delta\phi$ nutzen. Als Stellgröße kann dazu U_1 oder U_2 dienen, um wie oben, den Ausgangsstrom I_{DS} konstant zu halten, indem die Spannungsänderung $\Delta\phi$ an der Kapazität ausgeglichen wird.

Die Steuerung dieser beiden Größen erweist sich jedoch als ungünstig. Sobald an dem Spannungsteiler und im Bereich des Luftspaltes unterschiedliche Spannungen angelegt werden, fangen Ladungen auf der Oberfläche an, entlang der Feldlinien zu wandern. Infolgedessen beginnt die Kennlinie zu driften, d. h., sie entfernt sich von einer stabilen Basislinie. Besser ist aus diesem Grunde die Wannensteuerung [KBF⁺03]. Hierzu werden alle Spannungen an dem kapazitiven Teiler gleich gewählt. Man legt sie auf das definierte Massepotenzial. Als Stellgröße werden nun nicht die Spannungen U_1 oder U_2 verändert, sondern die Potenziale der Wanne sowie die Spannungen an Source und Drain um den gleichen Betrag mit entsprechend umgekehrtem Vorzeichen. (Die Spannung U_{DS} bleibt dabei trotzdem konstant.) Das Verändern der letzteren drei Potenziale um $-\Delta\phi$ bewirkt das Gleiche, wie die Beaufschlagung von $\Delta\phi$ an der Gate-Spannung, weil U_{GS} , U_{DS} und die Source-Bulk-Spannung U_{SB} dieselben Werte annehmen.

Zunächst nachteilig erscheint, dass eine Potenzialänderung der sensitiven Schicht nicht direkt, sondern nur in einem Verhältnis entsprechend der Gleichung (3.14) auf die Gate-Elektrode übertragen wird. Dieses Verhältnis ist stets viel kleiner als eins, weil die Kapazität des Luftspaltes C_1 kleiner ist als die des Oxides C_2 . Dafür ist es aber möglich, die Gate-Kapazität des FGFET um ein Vielfaches größer zu gestalten als die eines HSGFET. Dadurch erhält man Steilheiten, welche um zwei Größenordnungen größer sind.

Die Empfindlichkeit E der Sensoren ist ein Maß für Signalhöhe auf eine Änderung der Messgröße. Man gewinnt sie durch Differenziation der Ausgangskennlinie nach der Eingangsgröße im Arbeitspunkt. Mit Gleichung (3.4) ergibt sich so für den HSGFET:

$$E_{HSGFET} = \frac{\partial I_{DS}}{\partial \Delta\phi} = \mu C' \frac{W}{L} U_{DS} \quad (3.17)$$

und für den FGFET aus Gleichung (3.16):

$$E_{FGFET} = \frac{\partial I_{DS}}{\partial \Delta\phi} = \mu C_i' \frac{W}{L} \frac{C_1}{C_1 + C_2} U_{DS} . \quad (3.18)$$

Für einen Vergleich dieser beiden Empfindlichkeiten E_{HSGFET} und E_{FGFET} sollen zwei realistische Beispiele für Sensoren mit dem gleichen Luftspalt und auch sonst gleichen Transistoren herangezogen werden. Der Luftspalt mache 2 μm aus, die Dicke des Oxides ($\epsilon_{Ox}=3,9$) der Kapazität 2 sei 300 nm. Der Transistor des FGFET habe ein 20 nm dickes Gate-Oxid (ebenfalls $\epsilon_{Ox}=3,9$). Die daraus resultierenden Flächenkapazitäten werden in Tabelle 3.1 angegeben. Bei diesem FGFET wird eine Potenzialänderung Δ entsprechend Gleichung (3.15) nur in einem Verhältnis von 1 zu 27 auf das Gate des Transistors übertragen. Aber dafür hat der Transistor das 390fache der Steilheit (Gleichung (3.12)) des HSGFET mit ansonsten gleichen Parametern. Damit besitzt der FGFET eine um den Faktor 14,4 höhere Empfindlichkeit gegenüber dem anderen Sensor.

Sensor	Kondensator	ϵ	Schichtdicke	Flächenkapazität in $\frac{\text{nF}}{\text{cm}^2}$
HSGFET	Gate-Kapazität C'	1	2000 nm	0,443
FGFET	Spannungsteiler C_1	1	2000 nm	0,443
	Spannungsteiler C_2	3,9	300 nm	11,5
	Gate-Kapazität C_i'	3,9	20 nm	172,7

Tabelle 3.1: Werte der Kapazitäten für Beispiele von HSGFET und FGFET

3.5 Vergleich der GASFETs

Auch für GASFETs müssen zur Bewertung wieder die schon in 2.9 angeführten Kriterien herangezogen und einzeln, spezifisch für die Anwendung, gewichtet werden.

Im Falle der GASFETs hängen die Selektivitäten, die mit diesen Systemen erreicht werden, maßgeblich von der verwendeten selektiven Schicht ab. Versuche, mehrere dieser Bauelemente oder gar Arrays mit unterschiedlichen Schichten zu verwenden, zeigen, dass damit die Selektivität erhöht und die Querempfindlichkeiten unterdrückt werden können [KBF⁺03]. Da diese Gassensoren zurzeit noch nicht industriell gefertigt werden, können noch keine Aussagen über die Lebensdauer und Robustheit im realen Einsatz getroffen werden. Nebenbei hängen die letzten zwei Punkte stark von der Einsatzumgebung ab.

Ihre Vorzüge zeigen GASFETs aber auf dem Gebiet der Sensitivität. Wenn es gelingt, selektive Schichten zu finden, mit denen es möglich ist, reine Oberflächeneffekte zu messen, wird es nicht nötig sein, den Sensor zu heizen, sondern man kann ihn bei Raumtemperatur betreiben und den Sensor energiesparend einsetzen. Obendrein sind dadurch geringe Ansprechzeiten zu realisieren, weil nur kurze Zeiten notwendig sind, den Oberflächenbedeckungsgrad zu verändern.

Die meisten Transistoren, die in der Sensorik zur Analyse von Gaskonzentrationen untersucht werden ähneln denen, die in CMOS-Technologie hergestellt werden. Das bietet sich an, da so niedrige Herstellungskosten garantiert werden können, weil diese Technologie von Anfang an auf Massenfertigung ausgelegt ist.

Die meisten der eingesetzten selektiven Schichtsysteme basieren auf reversiblen Effekten. Diese verbrauchen sich von daher nicht aufgrund der Messung, wodurch auch niedrige Betriebskosten zu erwarten sind. Zusätzlich ist eine Integration des Sensors bei genügender CMOS-Kompatibilität mit der Auswerteschaltung auf einem Chip denkbar, weil beides in Siliziumtechnologie hergestellt wird, was die Kosten noch weiter senken könnte. Alles in allem haben GASFETs das Potenzial zu sehr billigen Sensoren.

Ein großes Problem liegt in der Stabilisierung des Sensorsignals und ist Gegenstand aktueller Forschung. Da ein Gas kein leitfähiges Medium darstellt, können sich unter Umständen undefinierte Potenzialverhältnisse einstellen. Geeignete

Schaltungen können das verhindern. Stellt man den Sensor in eine reale Testumgebung, so wird auf der Oberfläche des Sensors die Luftfeuchte adsorbiert und es bildet sich auf dieser ein Feuchtefilm. Über diesen fließen Ladungen zwischen den unterschiedlichen Potenzialen, im ungünstigsten Fall z. B. auf die Gate-Elektrode. Hierdurch driftet das Signal mit der Zeit auch ohne Änderung der Gaskonzentration. Mit Guard-Ringen wurde schon recht erfolgreich gezeigt, dass diese Effekte umgangen werden können [Bur03]. Eine andere Möglichkeit ist ein kurzzeitiges Heizen oder Erwärmen auf eine Temperatur, die leicht über der Umgebungstemperatur liegt. Da die Transistorkennlinien abhängig von der Temperatur sind, muss auch die Temperaturschwankungen mitberücksichtigt werden. Das kann z. B. mit einer Referenzmessung mit einem zweiten Transistor ohne sensitive Schicht erfolgen [Sch02].

Stellt man die vorher diskutierten Möglichkeiten mit Transistoren Gaskonzentrationen zu messen gegenüber, zeigt sich, dass der ISFET leichter zu stabilisieren ist. Aufgrund der definierten Potenzialverhältnisse auf der Oberfläche des Sensors kann man den Ladungstransport im Feuchtigkeitsfilm unterdrücken, der an Oberfläche adsorbiert wird. Dafür zeichnen sich die Sensoren mit Luftspalt durch schnelle Ansprechzeiten aus, da keine Volumenreaktion stattfinden muss. Außerdem können durch die frei zugängliche Oberfläche theoretisch alle Gase und Materialien, auch solche die keine Leitfähigkeit besitzen, als sensitive Schicht verwendet werden. Das meistversprechende Konzept unter den Sensoren mit Luftspalt ist wohl der FGFET. Er besitzt nicht nur eine höhere Empfindlichkeit, zudem kann der Transistor sehr frei ausgelegt werden. So kann v. a. die Einsatzspannung in den Bereich von einem Volt reduziert werden. Die hohen Einsatzspannungen des HSGFET dagegen eignen sich kaum für mobile Applikationen.

4 Entwurf eines Gassensors auf Feldeffektbasis

Die in den vorangegangenen Kapiteln erarbeiteten Spezifikationen sollen in einer Architektur und einem Layout umgesetzt werden. Das Prinzip der Messung einer Potenzialänderung einer sensitiven Schicht unter Gasbeaufschlagung mit einem GASFET mit Luftspalt stellte sich als am günstigsten heraus. Im Vergleich erwies sich der FGFET als aussichtsreicher Aufbau.

4.1 Architektur des Sensors

Der FGFET ergibt sich aus einer einfachen Verknüpfung der einzelnen Elemente. In einem Spannungsteiler wird eine Potenzialänderung an einer sensitiven Schicht an einem Mittelabgriff heraus geführt. Diese wird mit einem MOSFET in einen Strom als Ausgangssignal gewandelt. Mit Hilfe eines Guard-Rings werden durch Feuchte induzierte Leckströme innerhalb des Spannungsteilers verhindert. Eine Regelung zur weiteren Stabilisierung der Schaltung wird von außen beschaltet.

Für künftige Untersuchungen und Versuche wird zusätzlich ein baugleicher Referenztransistor mit in die Schaltung aufgenommen. Anhand des Referenztransistors kann man den Auslesetransistor charakterisieren, weil er über das Gate ansteuerbar ist. Obendrein ist man dadurch in der Lage, das genaue Potenzial, welches am Floating Gate herrscht, zu bestimmen. Dazu wird der Referenztransistor so geregelt, dass er und der Auslesetransistor den gleichen Ausgangsstrom liefern. Die an das Gate des Referenztransistors angelegte Spannung entspricht dann derjenigen des Floating Gates.

Das funktionale Modell ist in Abbildung 4.1 synthetisiert. Es ist auf das Minimum der benötigten Elemente reduziert. Aus Gründen der Redundanz werden jedoch mehrere Transistoren und Spannungsteiler in einem Sensor integriert. Für eine Umsetzung müssen die Transistoren noch genauer spezifiziert werden.

4.2 Vertikaler Auslesetransistor

GASFET basieren, wie schon in Kapitel 3 erläutert, im Wesentlichen auf MOS-Transistoren. Dieser Typ unterscheidet sich von den stromgesteuerten Bipolartransistoren und wird in die Gruppe der spannungsgesteuerten Transistoren eingeteilt. Als weitere Klassifizierung kann man ihn, im Gegensatz zu den Sperrschichttransistoren, als Transistor mit isolierter Steuerelektrode bezeichnen. Lilienfeld postulierte

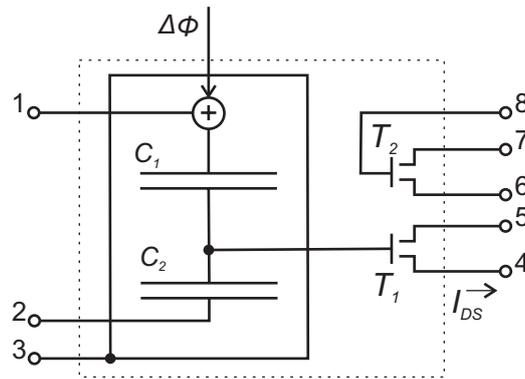


Abbildung 4.1: Funktionales Modell des Sensors

dieses Bauelement schon 1926 und reichte mehrere Patente dazu ein. 1930 erhielt er ein Patent, in dem der Grundaufbau des typischen planaren Feldeffekttransistors skizziert wird [Lil30]. Erst 1960 wurde dieser von Atalla und Kahng realisiert [KA60], weil vorher noch keine hoch qualitativen Halbleitermaterialien verfügbar waren. Kern dieses Bauelements ist eine Metall-Isolator-Halbleiter-Kapazität. Wie Abbildung 4.2 zeigt, schließen seitlich neben dieser hoch dotierte Gebiete an, deren Dotierungen komplementär zu der des Substrates sind. Liegt keine Spannung U_G an der Elektrode an, befindet sich immer einer der beiden pn-Übergänge der npn-Struktur zwischen Source und Drain im Sperrzustand. Es kann kein Strom fließen. Überschreitet hingegen U_G den Wert der Einsatzspannung, so bildet sich, wie in 3.1 schon beschrieben, eine Inversionsschicht aus. Die Ladungsträger in dieser sind parallel zur Gate-Elektrode beweglich und bilden so einen leitenden Kanal aus der Source und Drain elektrisch miteinander verbindet. Wie schon in Abschnitt 3.1 geschildert, kann über die Gate-Spannung die Anzahl der Ladungsträger in dieser Schicht gesteuert werden und somit die Leitfähigkeit des Kanals.

Durch dieses Prinzip wirkt dieses Bauelement wie eine spannungsgesteuerte Stromquelle mit extrem hoher Eingangsimpedanz. Nicht zuletzt deshalb besitzt dieser Transistortyp eine überragende Bedeutung in der Herstellung integrierter Schaltungen. Aber auch für die Nutzung als Ausleselement in einem Gassensor eignet sich dieser hervorragend. Aufgrund des hohen Eingangswiderstandes ist es möglich, das Potenzial der sensitiven Schicht stromlos zu messen, ohne es zu verändern. Die Verwendung der Stromstärke als Signalhöhe macht das Auslesen unabhängig von der Länge der Leitungen zur Auswerteelektronik.

Infolge des Einsatzes dieses Bauelementes in den integrierten Schaltungen, wurde es kontinuierlich weiterentwickelt und verbessert. Die eingeführten Techniken und Konzepte dienen hauptsächlich, um die lateralen Dimensionen immer weiter zu verkleinern und die dabei auftretenden Effekte, wie z. B. Kurzkanaleffekte, zu beherrschen. Die Motivation zum stetigen Schrumpfen der Transistoren wird durch die sinkenden Kosten getrieben. Kleinere Strukturen bedeuten mehr Transistoren pro Fläche und damit mehr Chips pro Wafer und letztendlich noch mehr Profit. In Gang gehalten wurde und wird dieser Fortschritt durch die Fotolithografie mit ihrer

immer feineren Auflösung. Diese bestimmt mit der kleinsten Strukturgröße im klassischen Fall des lateralen Transistors (Abbildung 4.2) die minimale Kanallänge. Letztere ist deshalb eine Kenngröße für die Technologie.

In dieser Arbeit soll ein FET als das verstärkende Element des Sensors eingesetzt werden. Interessant ist hier weniger die Menge der Transistoren, sondern deren Signalverstärkung. Betrachtet man die Gleichung (3.4), sieht man, dass die Beweglichkeit μ die Kapazität C'_i und das Verhältnis W/L aus Kanalweite zur Kanallänge des Transistors als Faktoren in die Kennlinien eingehen. Die Beweglichkeit ist jedoch als Materialkonstante vorgegeben. Auch die Gate-Kapazität, die sich aus dem Material und der Dicke des Dielektrikums ergibt, lässt sich nicht einfach erhöhen. Aber es kann durch das Design die Kanalweite W sehr viel größer gewählt werden als die Kanallänge L und somit über das W/L -Verhältnis die Verstärkung des Transistors erhöht werden.

Um das W/L -Verhältnis noch weiter zu erhöhen, kann auch eine möglichst kurze Kanallänge gewählt werden. Allerdings ist diese, wie erwähnt, in einem lateralen Konzept durch die kleinste Strukturgröße, die durch die Lithografie vorgegeben ist, nach unten begrenzt. Deshalb bietet es sich hier an, ein vertikales Konzept für den verstärkenden Transistor zu verwenden.

Bei dem vertikalen MOS-Transistor fließt der Strom durch den Kanal nicht lateral oder parallel zu Oberfläche des Wafers, sondern vertikal, d. h. senkrecht zu dieser. In Abbildung 4.3 ist der Aufbau eines solchen Transistors schematisch dargestellt. Die Source- und Drain-Gebiete befinden sich nicht mehr wie herkömmlich lateral neben dem Kanal, sondern unter und über dem Kanalgebiet. Dadurch ist die Kanallänge nicht mehr durch die Strukturauflösung der Fotolithografie definiert.

Wegen des vertikalen Stromflusses bestimmt hier die Schichtdicke des Materials, das den Kanal bildet (im Falle des Beispiels aus Abbildung 4.3 ist dies die Dicke der p-Schicht). Technologisch hat das den Vorteil, dass die verschiedenen Dotiergebiete mittels verschiedener Methoden der Schichtabscheidung übereinander gestapelt werden. Dabei lassen sich die Schichtdicken sehr klein, bis hinunter auf die atomare

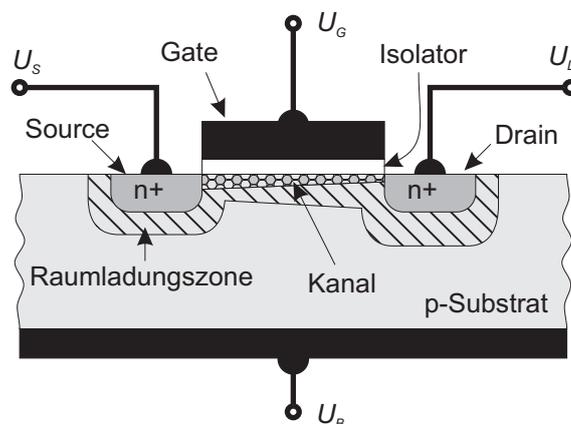


Abbildung 4.2: Schematischer Aufbau eines n-Kanal-Transistors

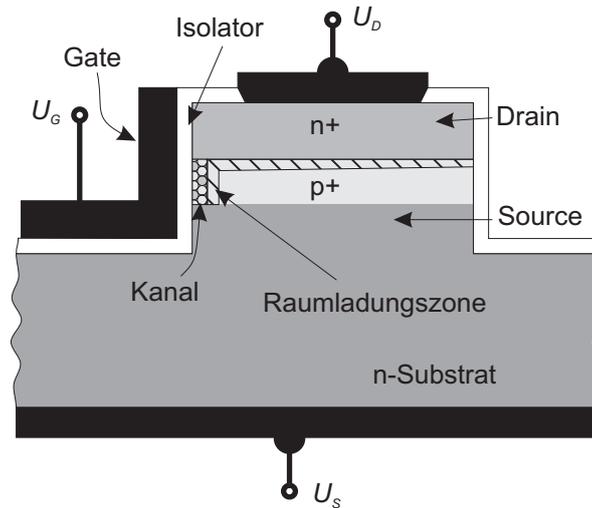


Abbildung 4.3: Schematischer Aufbau eines vertikalen n-Kanal-Transistors

Ebene einstellen. Kanallängen von wenigen 100 nm lassen sich vertikal sehr einfach herstellen. Lateral ist das nur mit großem Aufwand erreichbar [AMR03]. Aber nicht bloß die Schichtdicke, auch die Zusammensetzung oder das Dotierprofil ist mit modernen Methoden der Schichtabscheidung, wie MBE (Molecular Beam Epitaxie) oder CVD (Chemical Vapor Deposition), genau kontrollierbar. Erst nach der Abscheidung wird eine Mesa in den Schichtstapel geätzt, an deren Flanke, nach einer ganzflächigen Abscheidung des Gate-Dielektrikums, das Gate konstruiert wird. Die elektrische Funktion und die physikalischen Effekte sind die gleichen, wie die eines herkömmlichen lateralen MOS-Transistors.

Angesichts des oben genannten günstigen W/L -Verhältnisses und der daraus resultierenden hervorragenden Eignung in einem FGFET-Gassensor (siehe 3.4), wird in dieser Arbeit eben solch ein vertikaler Transistor als Ausselement dienen. Die Beweglichkeit von Elektronen in Silizium ist um den Faktor 2-3 größer als die der Löcher [Sze69]. Dieser Parameter geht proportional in die Ausgangskennlinie Gleichung (3.16) des Sensors ein. Deshalb ist ein n-leitender Transistor für den Sensor besser geeignet. Folglich wird auf einem hoch dotierten n-Substrat ein Schichtstapel in der Reihenfolge p- und dann n-leitendes Silizium epitaktisch aufgewachsen.

Um den Arbeitspunkt eines MOS-Transistors in einen gewünschten Bereich zu legen, wird oft das Potenzial der Gate-Spannung auf einen entsprechenden Wert eingestellt. In dem Floating-Gate-FET ist dieses aber nicht direkt zugänglich. Man vergleiche dazu die Abbildungen 3.4 und 3.5. Ein Potenzial der Gate-Elektrode kann lediglich kapazitiv über die Referenzelektroden eingekoppelt werden. Jedoch ist das ungünstig. Bei Feuchtigkeitsbeaufschlagung legt sich ein Feuchtigkeitsfilm auf der Oberfläche des Sensors nieder. Wegen unterschiedlicher Potenziale im Bereich des kapazitiven Teilers werden Leckströme hervorgerufen, die wiederum die Potenziale empfindlich beeinflussen. Besser ist es deshalb, alle Elektroden auf das gleiche Potenzial zu legen. Damit der Ladungstransfer mit anderen Bereichen,

wie dem Gehäuse oder Ähnlichem ausgeschlossen wird, wählt man hierfür am besten Masse. Um dennoch den gewünschten Arbeitspunkt zu erreichen, kann man bei herkömmlichen MOSFETs das entsprechende Potenzial mit umgekehrten Vorzeichen an die Transistorwanne legen [KBF⁺03].

Ein vertikaler Transistor besitzt selbst keine Wanne. Man kann ihn nicht über das Substrat steuern, weil das Bulk nicht kontaktiert werden kann (siehe Abbildung 4.3). Hier werden deshalb Source- und Drain-Spannung gleichzeitig um den entsprechenden Betrag verringert oder erhöht. Wichtig ist hier, dass Referenzelektroden und Source- oder Drain-Gebiete elektrisch isoliert sind, um an allen Punkten das Potenzial unabhängig von einander einstellen zu können.

4.3 SOI-Substrat

Am praktikabelsten lässt sich ein Spannungsteiler in der Siliziumtechnologie lateral aus Plattenkondensatoren realisieren. Dessen Spannungen müssen unabhängig von denen des Transistors eingestellt werden können. Leckströme bedingt durch die unterschiedlichen Potenziale dürfen nicht auftreten.

Der Spannungsteiler kann von dem Transistor entweder durch einen pn-Übergang oder durch einen dielektrischen Isolator getrennt werden. Der pn-Übergang zeichnet sich durch eine einfache Herstellung aus. In die zu isolierenden Gebiete werden Wannen implantiert. Allerdings muss immer dafür gesorgt werden, dass die Diode auch in Sperrrichtung gepolt ist. Zusätzlich wird der Spannungsbereich eingeschränkt, weil beim Erreichen der Durchbruchspannung U_{BD} wieder Strom fließt. Diese Spannung ist in erster Näherung umgekehrt proportional zu Dotierkonzentration N_D bzw. N_A . Je geringer also die Dotierkonzentration, desto größer ist die Durchbruchspannung [Gro67]. (Dieser Umstand wird im Abschnitt 5.2 noch von Belang sein.) Durch jede Diode fließt in Sperrrichtung ein Sperrstrom. Somit können Leckströme mit dem pn-Übergang nicht komplett unterbunden werden.

Ein anderer Weg, die Elemente zu trennen, ist eine Isolierung mit Hilfe eines Dielektrikums. In der üblichen CMOS-Technologie geschieht das unter Anwendung von Dickoxiden. Die einzelnen Bauelemente sind damit jedoch nicht rundum dielektrisch voneinander getrennt. Durch die Metallisierung entstehen trotzdem parasitäre Transistoren. Es treten Leckströme auf. Um diese zu minimieren, erfordert diese Technik ebenfalls einen zusätzlichen Implantationsschritt [Hof90].

In dieser Arbeit soll ein anderer Ansatz untersucht werden. Der kapazitive Teiler und der Transistor werden gleich auf einem dielektrisch isolierenden Substrat gefertigt. Das ist geschickt unter der Verwendung von SOI-Substraten (Silicon On Insulator) zu bewerkstelligen. Wie in 5.1 noch gezeigt wird, ist es damit möglich, die Elemente vollständig zu isolieren und außerdem Prozessschritte einzusparen. Aufgrund dieser Vorzüge ist diesem Ansatz der Vorrang zu geben.

Einen zusätzlichen Nutzen erhält man aus dem Umstand, dass vertikale Transistoren auf SOI-Substraten keine Implantationswannen benötigen. Die Dotierungen

für die Schichten diese Transistoren können sehr hoch gewählt werden. Dadurch ist es möglich den Sensor mit einem großen Temperatureinsatzbereich zu entwerfen. Das soll im Abschnitt 5.2 gezeigt werden.

4.4 Prozessablauf der Herstellung des Gassensors

Vor einer Umsetzung der Architektur in ein Masken-Layout, muss der genaue Prozessablauf der Herstellung des Sensors bekannt sein. Infolge der oben angestellten Überlegungen wird für den Sensor ein vertikaler Transistor verwendet. Die dazu benötigten Schichten sollen mittels Molekularstrahlepitaxie (MBE) auf einem n-dotierten Substrat abgeschieden werden. Zudem wird ein Spannungsteiler benötigt. Dieser wird lateral auf dem Wafer als Plattenkondensator aufgebaut. Um den Transistor elektrisch von dem kapazitiven Teiler zu trennen, soll der Sensor mit einem SOI-Substrat realisiert werden. Dieser wird hybrid aus zwei Teilen, im Folgenden Chip und Deckel genannt, gefertigt. Auf dem Chip befinden sich hauptsächlich die Transistoren, die nötigen Leiterbahnen und der untere Teil des Spannungsteilers. Die Deckel tragen die sensitive Schicht. In diesem Abschnitt wird der Prozessablauf für die Herstellung des Sensors beschrieben.

Ausgehend von einem SOI-Substrat mit einem n-dotierten Device-Layer, wird dieses zunächst thermisch oxidiert. Das auf der Oberfläche des Device-Layers entstehende 300 nm dicke Siliziumdioxid wird später als dielektrische Isolierung der Kapazität C_2 aus Abbildung 4.1 bzw. zwischen Referenz-Elektrode 2 und dem floatenden Gate analog zu den Abbildungen 3.4 und 3.5 dienen. Anschließend wird Fotolack aufgebracht und strukturiert (Abbildung 4.4). Er dient als Maske beim Ätzen des thermischen Oxids.

Nun werden epitaktisch mit MBE die Schichten für den Kanal (eine 200 nm p-Schicht, Dotierung: 10^{18} cm^{-3}) und das Source-Gebiet (eine 300 nm n-Schicht, Dotierung: 10^{20} cm^{-3}) des Auslesetransistors, abgeschieden. Nur dort, wo das Siliziumdioxid abgetragen und das Silizium frei zugänglich ist, wächst kristallines Silizium. Auf dem vorher strukturierten Oxid entsteht Polysilizium. Wieder wird Fotolack aufgetragen, belichtet und entwickelt. Diese Lackmaske definiert die Mesa, die anschließend mit einem trockenchemischen Ätzprozess heraus geätzt wird. Die Tiefe reicht dabei bis kurz unter die p-Schicht, in der sich später den Kanal bildet, demzufolge etwa 600 bis 700 nm. Die zu ätzenden Bereiche sind in Abbildung 4.5 angedeutet.

Nach dem Mesaätzen werden die Bauelemente freigestellt. Das heißt, es wird um die Transistoren und Referenzelektrode herum bis hinunter auf das vergrabende Oxid (BOX) das Silizium abgetragen. Dadurch sind die Bauelemente völlig elektrisch voneinander isoliert. Dies geschieht wiederum durch einen trockenchemischen Ätzprozess mit Hilfe einer Fotolackmaske (Abbildung 4.6).

In einem Ofenprozess wird das Gate-Oxid in einer Dicke von 16 nm hergestellt. Es folgt ein CVD-Prozess, mit dem ein 300 nm starkes Polysilizium abgeschieden wird, das als Gate-Material Verwendung findet. Eine erneute Fotolithografie -

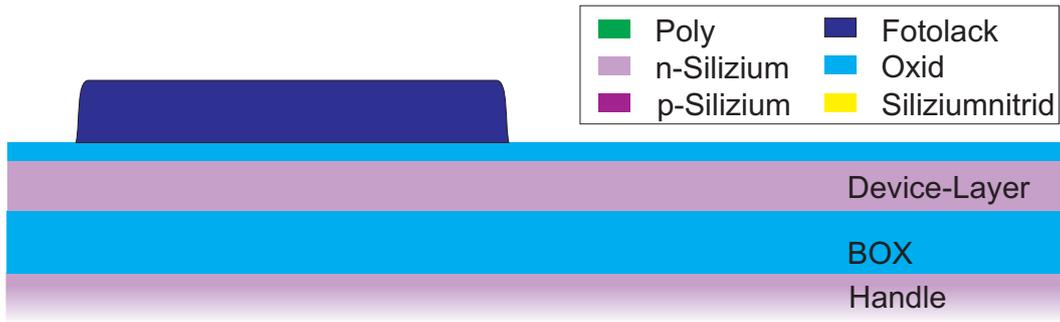


Abbildung 4.4: Oxidstrukturierung

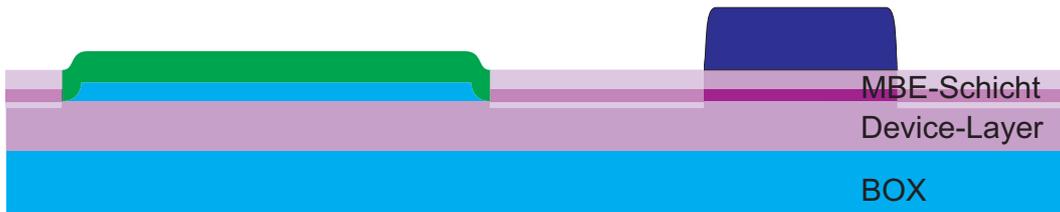


Abbildung 4.5: Mesastrukturierung

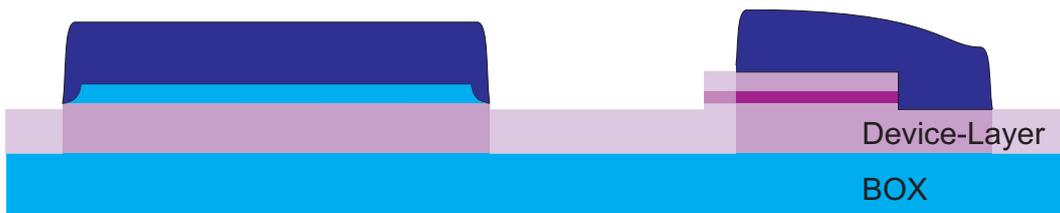


Abbildung 4.6: Freistellen der Bauelemente

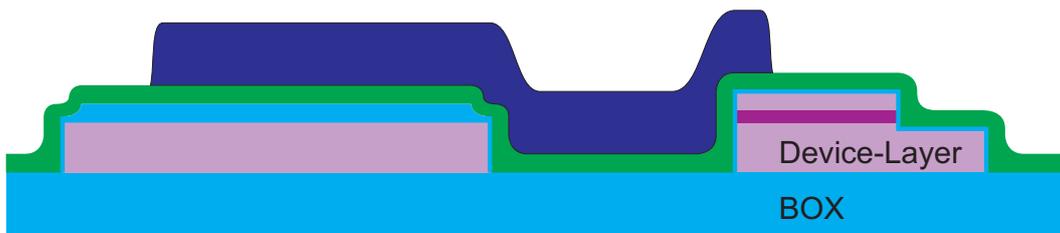


Abbildung 4.7: Strukturierung des floatenden Gates

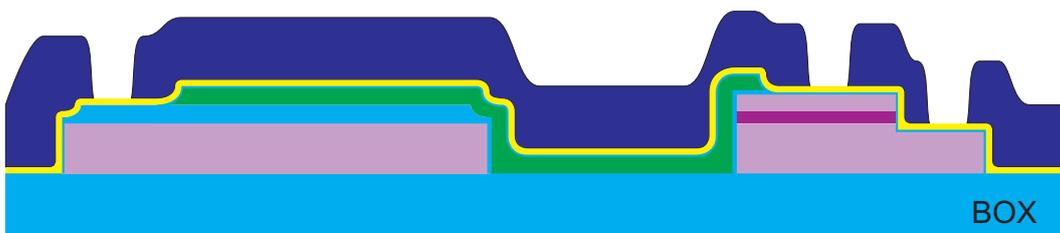


Abbildung 4.8: Öffnen der Kontaktlöcher

dieser Status ist in Abbildung 4.7 dargestellt - bestimmt dessen Geometrien. Die Strukturierung erfolgt nasschemisch.

Um das Floatende Gate und den Auslesetransistor zu verkapseln, wird nun eine Passivierung gefertigt. Dazu werden 10 nm eines thermischen Oxides gewachsen, welches mechanische Spannungen der nächstfolgenden Schicht abschwächen soll. Anschließend wird eine 130 nm dicke Schicht LPCVD-Siliziumnitrid (Low Pressure Chemical Vapor Deposition) abgeschieden. Durch eine erneute Lithografie werden die Kontaktlöcher definiert (Abbildung 4.8). Das Öffnen der Kontaktlöcher geschieht mit einer Kombination aus trocken- und nasschemischen Ätzen, weil am Transistor nur die Passivierung geöffnet werden muss, während für ein Kontaktieren der Referenz-Elektrode 2 zusätzlich die elektrische Isolierung (das 300 nm Oxid der ersten Oxidation) zwischen dieser und dem floatenden Gate geöffnet werden muss.

Zum Schluss wird die Metallisierung hergestellt, die das elektrische Anschließen der Bauelemente ermöglicht. Sie wird mittels Sputtern ganzflächig abgeschieden. Ihre Gesamtdicke beträgt ungefähr 450 nm. Darauf folgt nochmal ein Lithografie-durchlauf zur Herstellung einer Ätzmaske mit den Strukturen der Zuleitungen und Kontakte (Abbildung 4.9). Diese werden nasschemisch herausgeätzt.

Parallel dazu wird der Träger der sensitiven Schicht, der Deckel, gefertigt. Diese haben im Wesentlichen zwei Funktionen. Sie tragen die sensitive Schicht und die Abstandhalter. Sie werden mit Epoxidharzkleber auf dem Chip befestigt. Deshalb werden außerdem Vertiefungen eingearbeitet, um den Kleber aufzunehmen (siehe Abbildung 4.10 links). Diese werden mit Hilfe einer Hardmask strukturiert. Dazu wird auf ein Siliziumsubstrat Aluminium gesputtert und fotolithografisch und nasschemisch strukturiert. Anschließend werden die Vertiefungen mit einem RIE-Prozess (Reactive Ion Etching) hineingeätzt. Aus Gründen der Ausbeute werden diese in den Deckel integriert, um im schlimmsten Fall nur die Deckel zu zerstören, nicht aber die Chips.

Dieselbe Überlegung ist auch der Anlass, die Abstandhalter in den Deckel zu integrieren. Prinzipiell können diese subtraktiv oder additiv hergestellt werden. Verfahren, die wie das Ätzen die Abstandhalter subtraktiv aus dem Substrat formen, rauhen aber die Oberfläche auf. Daher wird Aluminium in einer Dicke abgeschieden, die der Höhe der Abstandhalter entspricht. Sie hängt im Wesentlichen von den benutzten SOI-Substraten ab, denn sie beeinflussen die Höhen der Schichtstapel auf dem Chip. Damit wird später die Größe des Luftspaltes definiert. Das Metall wird mit Fotolithografie und nasschemisch strukturiert (siehe Abbildung 4.10 rechts). Bei Silizium oder Glas als Substratmaterial wird so die Oberfläche nicht aufgeraut. Gerade bei der Messung von Dipolen (an Grenzflächen adsorbierte Gasmoleküle sind solche) ist das wichtig, weil diese sich immer senkrecht zur Oberfläche ausrichten. Ist diese rau, sind nicht alle Dipole parallel orientiert. Entgegengesetzte Anteile heben sich gegenseitig auf. So resultiert nicht die volle Potenzialänderung, wie sie mit der gleichen Anzahl an Dipolen auf einer absolut planen Oberfläche messbar wäre.

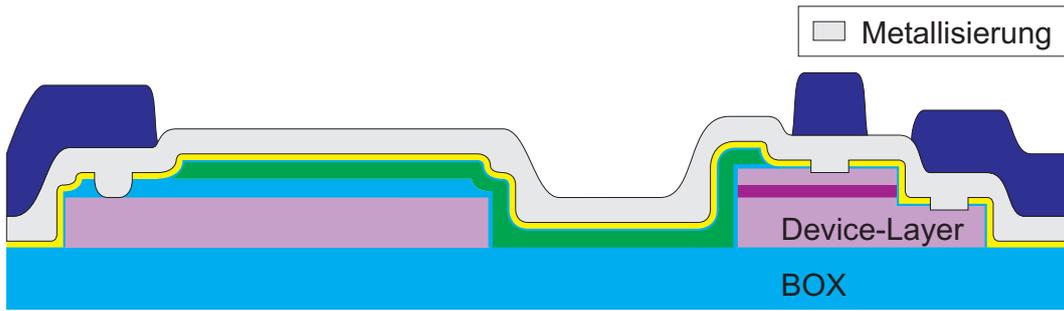


Abbildung 4.9: Metallisierung

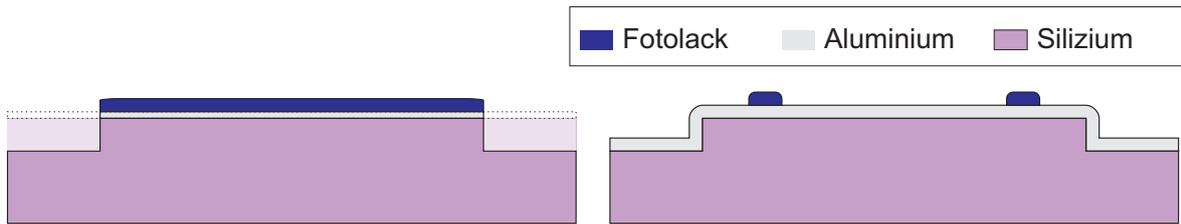


Abbildung 4.10: links: Siliziumätzen mit Hilfe einer Hardmask. rechts: Strukturierung der Abstandhalter.

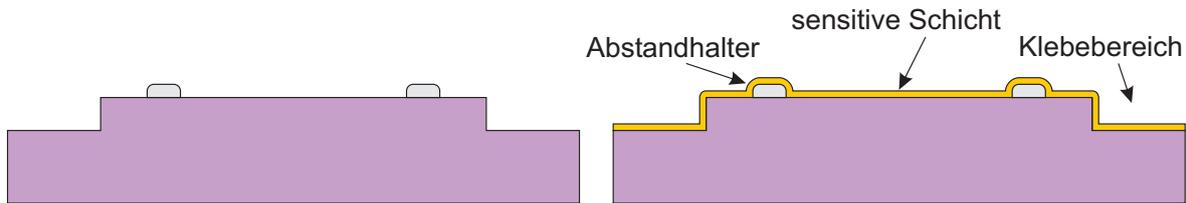


Abbildung 4.11: links: Deckel nach der Strukturierung der Abstandhalter. rechts: Aufbringen der sensitiven Schicht.

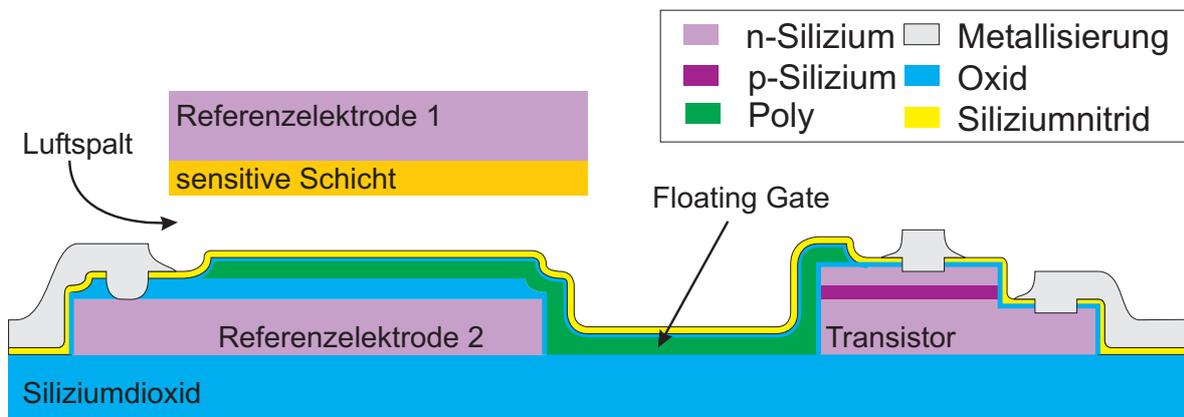


Abbildung 4.12: FGFET auf SOI mit vertikalem Auslesetransistor

Nun wird die sensitive Schicht aufgebracht (siehe Abbildung 4.11). Das Verfahren dafür kann relativ frei gewählt werden, weil dieses die Fertigung der Chips nicht beeinflusst oder stört. Diese zusätzlich aufgebrachte Schicht muss lediglich den Bedingungen genügen, auf dem Material des Deckels zu haften und man muss auf ihr kleben können. Ansonsten muss diese zusätzlich strukturiert werden, sodass Substrat wieder frei zugänglich ist.

Beide Teile des Sensors, der Chip und der Deckel werden mit Hilfe der Flip-Chip-Technik miteinander verklebt, sodass sich die sensitive Schicht frei hängend über dem Floatenden Gate befindet und ein Luftspalt bildet, in den das Gas bis an die sensitive Schicht gelangt. Schematisch wird der Aufbau des fertigen Sensors in Abbildung 4.12 illustriert. Die Abstandhalter sind aus Gründen der Übersichtlichkeit darin nicht enthalten. Referenzelektrode 1 und 2 sowie das floatende Gate dazwischen bilden den kapazitiven Teiler. Die einzelnen Elemente des Sensors, der Transistor, die Referenzelektroden sowie die Gate-Elektrode, sind elektrisch voneinander völlig isoliert.

4.5 Layout des Sensors

Nachdem die Transistoren ausgewählt und der Prozessablauf des Sensors bestimmt sind, kann die Architektur für diesen Sensor in ein Layout für den Sensor umgesetzt werden. Später werden die im Folgenden angeführten Geometrien unter Berücksichtigung der „Designrules“¹ in einen Maskensatz übertragen.

4.5.1 Der Chip

Das Layout des Chips wird in Abbildung 4.13 präsentiert. Jeder einzelne Chip trägt vier sensitive Bereiche mit jeweils einem Auslesetransistor. Diese Redundanz stellt sicher, dass der Sensor auch dann noch gebraucht werden kann, wenn ein Transistor infolge eines Prozessfehlers oder eines Anschlussfehlers zerstört wurde. Zudem ergibt das die Möglichkeit die Gaskonzentrationen mit Hilfe einer Referenzmessung zu bestimmen, z. B. für eine Temperaturkompensation [Sch02].

Der Schichtstapel im sensitiven Bereich bildet den unteren Teil des Spannungsteilers. Die Ausmaße der Fläche beträgt $1000\ \mu\text{m} \times 500\ \mu\text{m}$. Damit wird gewährleistet, dass die Kapazität der Zuleitung vom sensitiven Bereich bis zu dem Transistor und die Gate-Kapazität gegenüber den Kapazitäten des Spannungsteilers klein sind und so dessen Potenzialverhältnisse nicht beeinflussen.

Um den sensitiven Bereich wird ein sog. Guard-Ring gelegt. Er reduziert die durch Feuchte induzierten Oberflächenleckströme [Bur03].

¹Das sind bestimmte Anforderungen, die an die Geometrie gestellt werden und welche sich aus der Technologieentwicklung ergeben, (wie z.B. die kleinste Strukturgröße oder der Vorhalt für das Ätzen).

Jeder Chip erhält zwei Referenztransistoren, wieder aus Gründen der Redundanz. Diese sind baugleich mit den Auslesetransistoren. Ihr Gate kann kontaktiert und somit der Referenztransistor angesteuert werden.

Um alle Anschlüsse zugänglich zu machen, werden die Bondpads am unteren Rand des Chips platziert, sodass sie außerhalb der Reichweite des Deckels liegen. Auf diese Weise ist auch ein größerer Deckel verwendbar.² Zur Orientierung sind sie beschriftet. Zwei große Flächen aus Aluminium bilden Kontaktflächen, auf denen die Abstandhalter des Deckels aufsetzen. Über diese kann der Deckel chipseitig kontaktiert werden.

Neben den für die Funktion des Sensors wichtigen Bestandteilen werden noch weitere Elemente auf den Sensor untergebracht. So bekommt jeder Chip zum späteren Unterscheiden eine individuelle Chip-Nummer. Da verschiedene Varianten des Sensors geplant sind, z. B. vergrabenes oder nicht vergrabenes Floating Gate, führt er eine Modellbezeichnung. Wie im Folgenden noch genauer beschrieben wird, werden drei unterschiedliche Layouts für die Auslesetransistoren untersucht. Auch der Typ des auf dem Sensor zu findenden Transistors ist angegeben. Damit später kontrolliert werden kann, welche Maskenebenen prozessiert wurden, hinterlässt jede Maskenebene entsprechend ihrer Reihenfolge eine Nummer. Für den Fall, dass nach dem Vereinzeln noch etwas prozessiert werden muss, um nachträgliche Ideen oder Experimente zu realisieren, befinden sich zwei Justiermarken auf den Chips.

Um den Chip herum wird ein Metallrand gelegt. Ähnlich wie der Goldrand, der das Obere eines Bierglases schmückt, stabilisiert er die Kante und verhindert er ein Abspringen kleinerer Stückchen. Alle diese Ränder sind auf dem Wafer miteinander verbunden. Über diesen Ring können alle Kontaktflächen für den Deckel elektrisch kontaktiert werden. Daher ist es für spätere Weiterentwicklungen denkbar auf diesen Flächen galvanisch etwas abzuscheiden.

4.5.2 Der Deckel

Das Layout wird in Abbildung 4.14 im Detail geschildert. Sie zeigt, wie der Deckel auf den Chip geklebt wird. Die Beschriftungen, die den Chip eindeutig identifizieren, bleiben trotzdem sichtbar. Dabei setzen die Abstandhalter auf den Kontaktflächen auf. Die sensitive Schicht kann so über ein Bondpad kontaktiert werden. Eine Klebersperre sorgt dafür, dass der Kleber nur innerhalb der Klebefläche bleibt. Die Notwendigkeit der dargestellten Geometrie dieser Vertiefungen für den Klebstoff wird in Abschnitt 6.7.1 ausgeführt.

²Außer der in dieser Arbeit vorgestellten Gasdetektion, ist es genauso von Interesse die Potenzialänderung des FGFETs für die Flüssigkeitsdetektion heranzuziehen. Die Anordnung in Abbildung 4.13 ermöglicht auch das Eintauchen der sensitiven Bereiche in eine Flüssigkeit ohne die Bondpads zu benetzen.

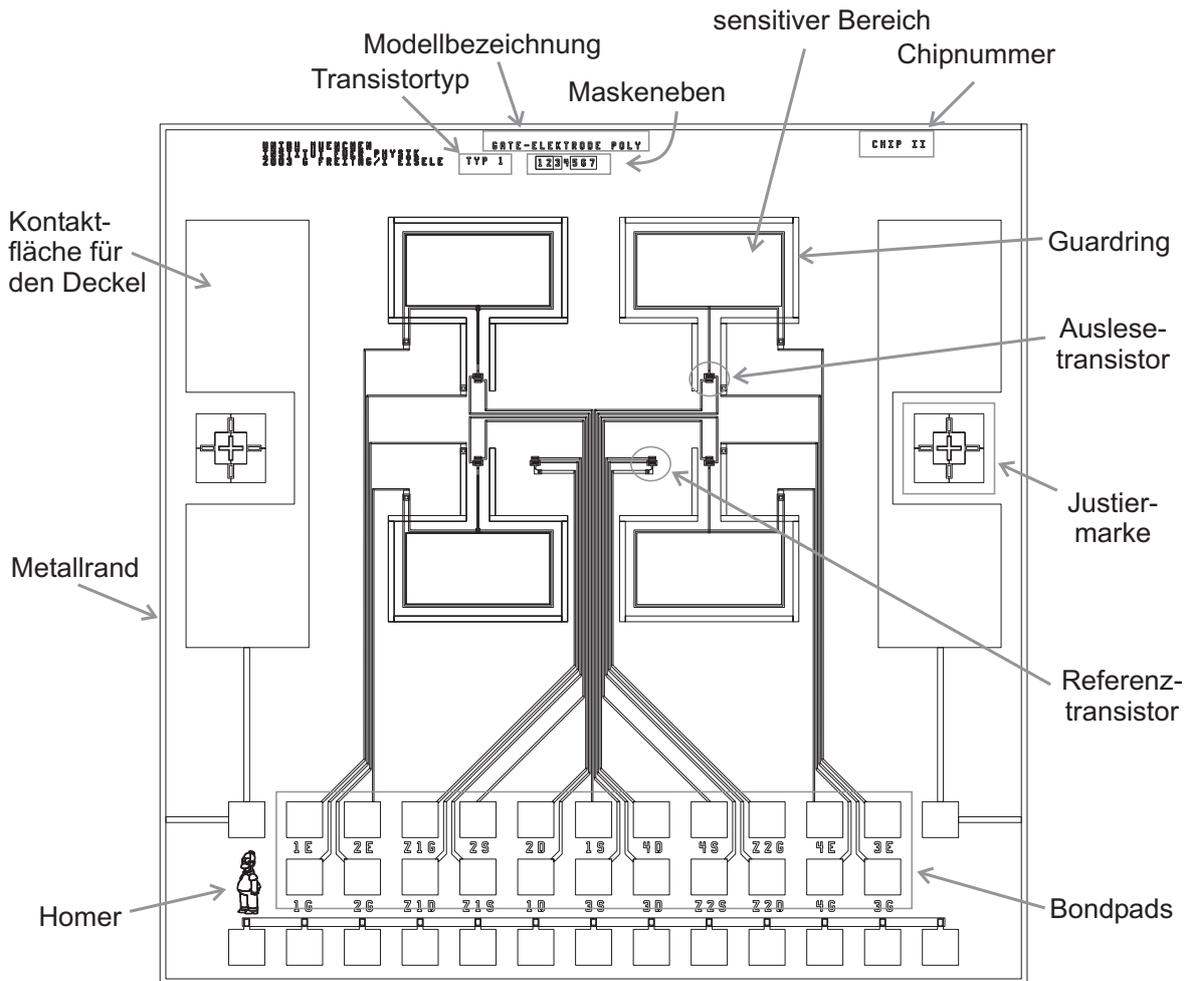


Abbildung 4.13: Layout des Chips

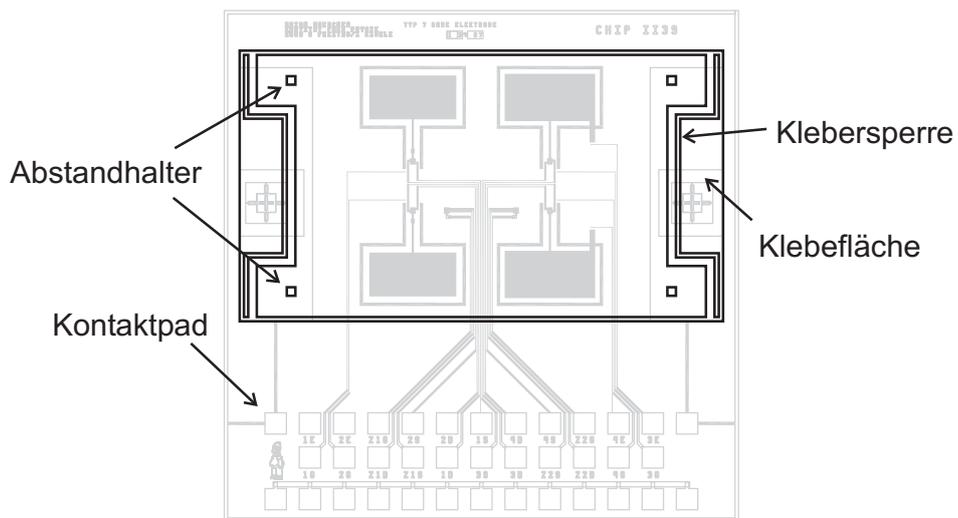


Abbildung 4.14: Design des Deckels und dessen Lage auf dem Chip

4.5.3 Die Transistoren

Es werden drei verschiedene Layouts auf die Eignung der Transistoren zum Auslesen der Potenzialänderung untersucht. Deren Layouts findet man in den Abbildungen 4.15 bis 4.17. Im Folgenden werden diese Transistoren Transistor Typ 1, Transistor Typ 2 und Transistor Typ 3 bezeichnet.

Die Mesen der Transistoren werden elektrisch vollständig isoliert, d. h. sie sind komplett, abgesehen von den Kontaktlöchern, mit Siliziumdioxid verkapselt. Eine seitliche Kontaktierung der Mesa oder gar eine Strukturierung auf der Seitenflanke ist technisch nicht machbar. Deswegen muss eine Stufe in die Mesa geätzt werden, damit Source und Drain kontaktiert werden können (siehe Abbildung 4.12 und Abbildung 4.18). Daher erfolgt hier das Fertigen der Mesen in zwei Schritten, dem Mesaätzen (Abbildung 4.5) und dem Freistellen (Abbildung 4.6).

Die zwei angesprochenen Prozessschritte werden mit einem RIE-Ätzprozess für Silizium durchgeführt. Sie erzeugen beide Flanken, die für den vertikalen Kanal des Transistors genutzt werden können. Sie unterscheiden sich darin, wie tief in das Silizium geätzt wird. Beim Mesaätzen liegt diese Tiefe zwischen 700 und 800 nm, beim Freistellen beträgt sie hingegen ca. 3000 nm. Die tiefere Flanke wird für Transistoren Typ 1 und 2 verwendet, wohingegen der Kanal des Transistors Typ 3 entlang der anderen Kante läuft. Abbildung 4.18 veranschaulicht das durch die Profilschnitte der Transistoren.

Die hoch dotierten Schichten werden mit MBE hergestellt, weil dieses Verfahren gute Kristallqualitäten liefert. Allerdings besitzen auch sie eine geringe Defektdichte. Je größer die Mesa des Transistors ist, desto höher ist die Wahrscheinlichkeit, dass ein Versetzungsfehler das Bauelement unbrauchbar macht. Kennt man die Defektdichte, also die Anzahl der Fehler pro Fläche, kann man abschätzen, wie groß die Mesafläche sein darf, mit der man trotzdem noch genügend funktionierende Bauteile erhält. Bei einer angenommenen Defektdichte von 10^4 cm^{-2} [Feh97, Fin00] errechnet sich die Ausfallwahrscheinlichkeit der Transistoren Typ 1 bis Typ 3 zu 33, 12 bzw. 35 %. Für die Herstellung der Sensoren im Rahmen dieser Arbeit ist $\frac{1}{3}$ völlig ausreichend. Der Transistor Typ 2 mit einer Fläche von nur $1200 \mu\text{m}^2$ wurde für noch höhere Zuverlässigkeiten entworfen. Dieser entspricht ansonsten Transistor Typ 1.

Um die Überlappkapazität der Transistoren so klein wie möglich zu halten, darf die Gate-Elektrode die Mesa nur so weit wie nötig überlappen. Eine Struktur, die exakt an die Mesaflanke heranreicht, diese aber nicht überlappt, ist nicht realistisch. Die Justiergenauigkeit der Fotolithografie und eventuelle Unterätzungen müssen mitberücksichtigt werden. Für den Transistor Typ 1 und Typ 2 beträgt die Überlappung $2 \mu\text{m}$ bzw. $1 \mu\text{m}$ (Abbildung 4.15 und 4.16). Die Überlappkapazität des Transistors Typ 3 ist am größten, weil der Kanal sich an der Kante befindet, die beim Mesaätzen entsteht, und somit die Gate-Elektrode Source und Drain überdeckt. Man erkennt das anhand der Profilschnitte in Abbildung 4.18. Ein schmales Gate ausschließlich auf diese Kante zu legen, ist wieder aus den genannten Gründen ausgeschlossen.

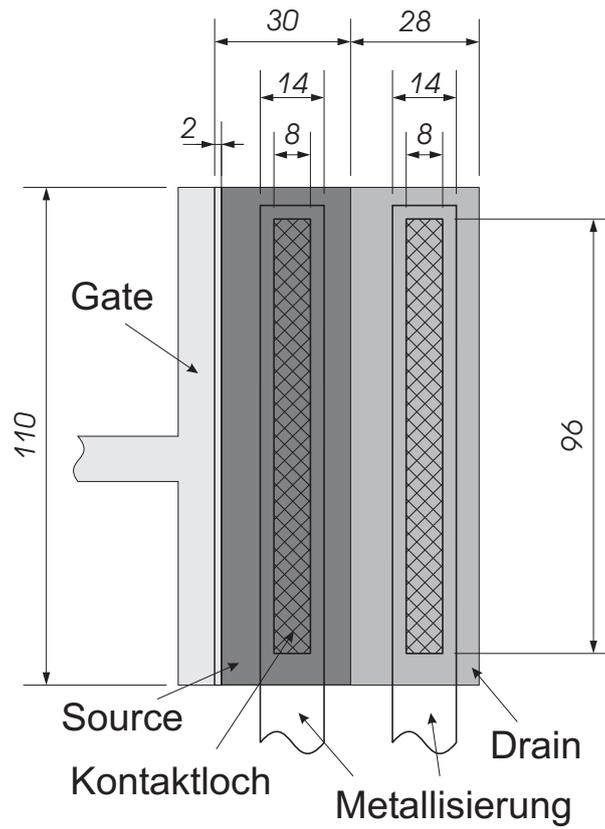


Abbildung 4.15: Transistorlayout Transistor Typ 1 (Maße in μm)

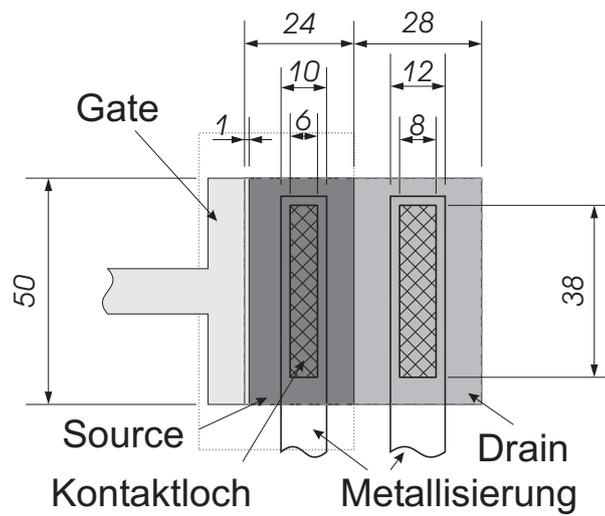


Abbildung 4.16: Transistorlayout Transistor Typ 2 (Maße in μm)

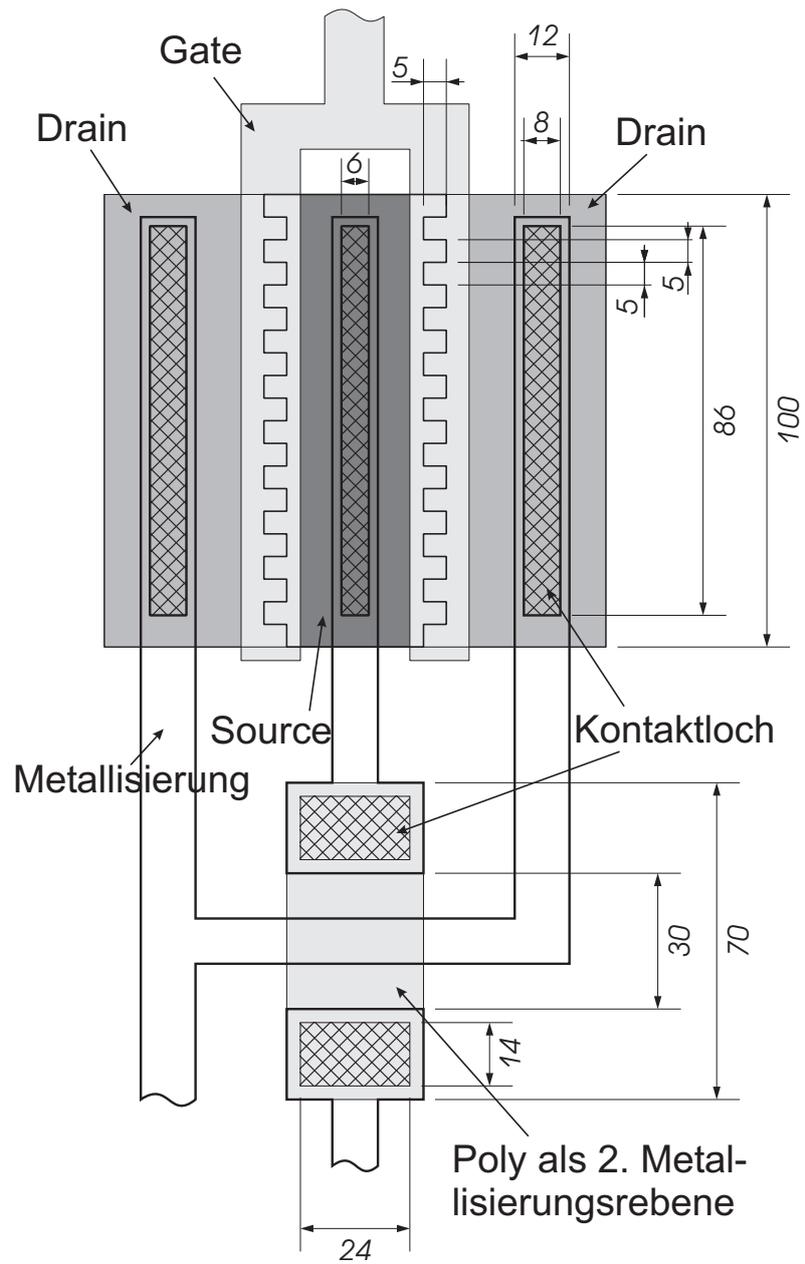


Abbildung 4.17: Transistorlayout Transistor Typ 3 (Maße in μm)

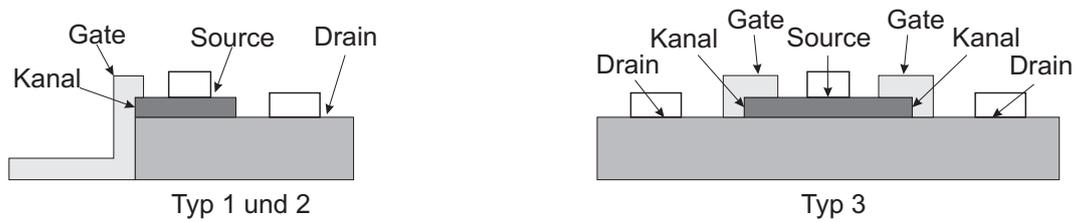


Abbildung 4.18: Profilschnitte der Transistoren

Transistor Typ	1	2	3
Ausfallwahrscheinlichkeit	33 %	12 %	35 %
Kanallänge	110 μm	50 μm	400 μm
W/L -Verhältnis	550	250	2000

Tabelle 4.1: Zusammenstellung der unterschiedlich dimensionierten Kenngrößen für die verschiedenen Transistortypen

Da für den Transistor Typ 3 eine hohe Kapazität in Kauf genommen werden muss, bietet es sich an, das Gate in Form eines Mäanders zu konstruieren (Abbildung 4.17). Dadurch vergrößert sich das W/L -Verhältnis in diesem Fall noch einmal um den Faktor 2. Um dieses Verhältnis weiter zu steigern, wird der Transistor Typ 3 symmetrisch aufgebaut und erhält zwei gegenüberliegende Kanäle. Dazu ist eine zweite Metallisierungsebene notwendig, weil sich so entweder die Zuleitungen für Source und Drain oder für Source und Gate einmal kreuzen müssen. Günstiger ist die erste Variante, weil dann die floatende Gate-Elektrode nicht beeinflusst wird. Um diese zweite Metallisierungsebene einfach herzustellen, wird dazu Polysilizium verwendet, das ebenfalls für die Gate-Elektrode abgeschieden wird. Als Dielektrikum fungiert die Passivierung aus Siliziumdioxid und -nitrid.

Alle Typen sind n-Transistoren. Ihre Kanallängen sind alle identisch 200 nm, weil die Schichten epitaktisch aufgewachsen werden. Daraus resultieren W/L -Verhältnisse im Wert bis zu 2000. Alle Kenngrößen sind in Tabelle 4.1 zusammengefasst.

5 FGFET auf SOI

SOI(Silicon On Insulator)-Wafer sind eine besondere Form der Siliziumsubstrate. In diesen Siliziumscheiben befindet sich eine elektrisch isolierende Schicht unterhalb einer einkristallinen Siliziumschicht. Solche Substrate werden bereits mehr als drei Dekaden verwendet. Als Erstes wurden dafür sog. SOS (Silicon On Sapphire)-Substrate verwendet. Heute kommen fast ausschließlich Substrate zum Zuge, bei denen eine vergrabene Schicht aus Siliziumdioxid als isolierende Schicht dient.

Handelsübliche SOI-Wafer ähneln den normalen Siliziumscheiben. Sie werden in den üblichen Abmessungen hergestellt und verkauft. 100-mm-Scheiben besitzen ebenso wie herkömmliche Wafer eine Dicke von ungefähr 350-500 μm . Diese Substrate sind wie folgt aufgebaut:

- Auf dem Substrat befindet sich ein einkristalliner Siliziumfilm. Üblicherweise bezeichnet man diesen, auf dem die Bauelemente gefertigt werden, als aktive Schicht, Device-Layer oder kurz Device. Dieser hat eine Dicke von wenigen Nanometern (z. B. 25 nm) bis zu 50 μm , kann aber auch mehrere hundert μm stark sein.
- Darunter liegt eine Schicht Siliziumdioxid. Diese wird Burried Oxid oder kurz BOX genannt. Die möglichen Dicken gleichen denen der Device-Layer, müssen aber nicht notwendigerweise gleich sein.
- Die mechanische Stabilität liefert ein einkristalliner Siliziumwafer, der die beiden oberen Schichten trägt. Man spricht vom sog. Handle. Letzterer hat die Dicke eines herkömmlichen Siliziumwafers.

Device-Layer und Handle können dabei beliebige Dotierungen und Kristallorientierung besitzen. Auch alle Kombinationen sind denkbar. Es werden auch Substrate verkauft, die aus einem Schichtstapel mehrerer alternierender Lagen Si und SiO₂ aufgebaut sind. Diese werden aber im Rahmen dieser Arbeit nicht betrachtet. In diesem Kapitel werden die Vorzüge dieser Wafer für den Sensor, der in dieser Arbeit gefertigt werden soll, diskutiert. Die Herstellungsverfahren und deren Eignung werden Anhang B erörtert.

5.1 Allgemeine Vorteile der SOI-Substrate für Sensoren

SOI-Wafer fanden von Anfang an Verwendung für extreme Umgebungsbedingungen, wie zum Beispiel Applikationen, die für den Einsatz in hohen Temperaturen oder der Raumfahrt ausgelegt waren. Unabhängig davon bieten sie Vorteile, die

es sehr wahrscheinlich machen, dass diese Substrate bald auch in Halbleiter-Applikationen des Massenmarktes die Standard-Substrate verdrängen.

Leckströme

Durch die isolierende Schicht ist es möglich Bauelemente vollständig elektrisch zu trennen. Zum einen werden vertikale Leckströme von Bauelement zum Bulk, d.h. zum Substrat durch diese Wafer selbst unterbunden, da zwischen diesen Gebieten eine vergrabene isolierende Schicht liegt. Zum anderen ist es aber auch leicht möglich Leckströme in lateraler Richtung zu unterbinden. Dazu müssen lediglich Gräben um die zu trennenden Gebiete bis herunter auf das Substrat geätzt werden (Mesaätzen). Durchschlagfestigkeit können bei diesen Substraten über die Dicke des Oxides und die Weite der Gräben eingestellt werden. Diese Grabenisolation (engl. Trenchinsulation) hat im Gegensatz zur Isolation durch in Sperrrichtung geschalteter Dioden (engl. Junction Insulation) den Vorteil, dass hier nicht auf das Anliegen der richtigen Potenziale geachtet werden muss. Die Gebiete sind unabhängig von der angelegten Spannung elektrisch von einander getrennt. Gegenüber der Grabenisolation bei der Verarbeitung konventioneller Wafer müssen die Gräben nur bis zum vergrabenen Oxid reichen und trotzdem sind die Bauelemente hier völlig von einem Dielektrikum umgeben.

Parasitäre Kapazitäten

Die Volumina der Sperrschichten an allen pn-Übergängen bestimmen im Wesentlichen die Gesamtkapazitäten. In integrierten Halbleiterschaltungen findet man diese nicht nur direkt unter den Gate-Elektroden, sondern auch, um die Source- und Drain-Gebiete und an allen, zur elektrischen Trennung in Sperrrichtung geschalteten Dioden (Junction Insulation). Wie oben erwähnt, lässt sich die Sperrschichtisolation beim Einsatz von SOI-Wafern vermeiden. Reichen die Implantierten Gebiete bis zur Grenzfläche des Siliziums und des Siliziumdioxides bilden sich dort keine Verarmungszonen aus. Aber auch wenn die pn-Übergänge nicht ganz bis zu dem vergrabenen Oxid heranreichen, verringern sich die Kapazitäten schon merklich [Sha02].

Ebenfalls verringern sich die parasitären Kapazitäten zwischen den vor allen unteren Verdrahtungsebenen und dem Substrat, da hier die vergrabene Siliziumdioxid-Schicht als zusätzliches Dielektrikum dient. Um diese Kapazitäten zu verringern, werden heutzutage Kupfermetallisierungen und Low-K-Materialien eingesetzt. Für diesen Fall gewinnt man noch zusätzlich eine Reduzierung der Leitungskapazitäten.

Parasitäre Bipolartransistoren

In jedem MOS-Transistor steckt immer auch ein parasitärer Bipolartransistor. Wobei das Substrat unterhalb des Kanalgebietes als Basis und Source und Drain

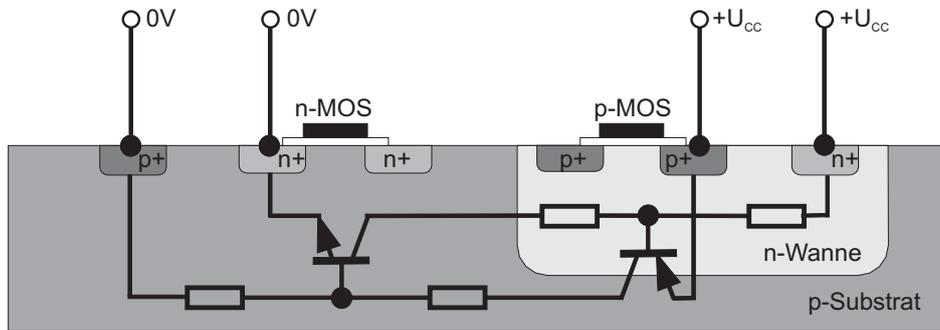


Abbildung 5.1: Schematische Darstellung eines CMOS-Inverters mit den Strompfaden während des Latchup-Effektes

als Emitter bzw. Kollektor wirken. Normalerweise sind alle pn-Übergänge in Sperrrichtung geschaltet, sodass kein bipolarer Verstärkungseffekt auftritt.

In CMOS-Schaltungen können zwei komplementäre Bipolartransistoren zusammen eine Vierschichtdiode bilden (siehe Abbildung 5.1). Unter ungünstigen Umständen kann diese gezündet werden, wobei ein hochleitender Strompfad zwischen Versorgungsspannung und Masse entsteht. Das kann durch die großen dabei auftretenden Ströme zur Zerstörung der Bauelemente führen. Dieser Effekt wird in der Literatur genauer beschrieben [Hof90].

Im Gegensatz dazu lassen sich diese parasitären Bipolartransistoren in SOI-CMOS-Invertern allein durch das Layout vermeiden.

Softerrors

Beim Eintritt hoch energetischer Strahlung, wie α - oder γ -Strahlung, in Silizium werden Elektronen-Loch-Paare erzeugt. Trifft nun ein Strahlungsteilchen eine Raumladungszone, wird diese entlang dessen Weges erweitert. Diese Ausdehnung liegt in der Größenordnung von $10\ \mu\text{m}$. Die Ladungsträger werden getrennt und während die eine Sorte über das Substrat abfließt, erhöht die andere den Driftstrom. Dieser Vorgang dauert Bruchteile einer Nanosekunde. Danach diffundieren die verbliebenen Ladungsträger außerhalb der Raumladungszone in diese hinein und erzeugen einen zusätzlichen Driftstrom. Dieser zweite Strom ist viel kleiner als der Erste, hält aber einige hundert Nanosekunden an. In dieser Zeitspanne werden integrierte Schaltungen gestört und es treten Softerrors auf. Durch auftretende Potenzialverschiebungen kann an Knoten ein falsches Potenzial anliegen, was zu Bit-Verdrehern führt. Die freien Ladungsträger konnten aber auch den oben diskutierten Latchup-Effekt erzeugen. Schlimmstenfalls wird die Schaltung zerstört.

Durch die vergrabene SiO_2 -Schicht kann sich die Raumladungszone nur noch bis zu dieser ausdehnen. Ladungsträger, die unterhalb dieser Schicht erzeugt werden, können nicht mehr von den Raumladungszonen der aktiven Schicht eingefangen werden. Die auftretenden Störströme werden verringert, wie auch ihre Wirkung.

Einsparung von Prozessschritten

Bei herkömmlichen Substraten wird die elektrische Isolation durch einen in Sperrrichtung geschalteten pn-Übergang realisiert (Junction Insulation). Deshalb muss die n-Wanne immer auf höherem Potenzial liegen als die p-Wanne. In modernen integrierten Schaltungen müssen deshalb mittlerweile kaskadierte Wannen prozessiert werden. So sind zwei und drei Wannen mittlerweile üblich (Twin Well / Tripple Well). Das macht zusätzliche Masken- und Implantationsschritte notwendig. Im Gegensatz dazu werden beim Einsatz von SOI-Technologien die Schaltungen auf einem Isolator aufgebaut. Durch die vollständige Isolation (Dielectric Insulation) können kleine Inselchen oder auch einzelne Bauelemente eigene Substratspannungen besitzen. Wie oben erwähnt, lässt sich das hier durch einen Maskenschritt in Verbindung mit einem Ätzschritt bewerkstelligen.

Aber nicht nur bei der elektrischen Trennung zweier Gebiete vereinfacht das vergrabene SiO_2 die Prozessierung. Oft wird auch erwähnt, dass sich durch die Verwendung von SOI-Substraten das sog. „Spiking“ vermeiden lässt. Das ist ein pyramidenartiges Hineinwachsen des Aluminiums aus der Metallisierung in das darunter liegende kristalline Silizium [Eis01]. Dabei wird der sich unter dem Kontakt befindende pn-Übergang (der Wanne, Shallow Junction) kurzgeschlossen [WMF88]. Um das zu verhindern, wurden zum Teil aufwendige Verfahren eingeführt, die die Anzahl der Prozessschritte sowie die Produktionskosten erhöhen [Fra83]. Wird aber ein SOI-Substrat mit einem dünnen einkristallinen Film verwendet, so endet das Implantationsgebiet nicht am pn-Übergang sondern direkt auf der Grenzfläche zwischen SiO_2 und dem einkristallinen Si, die als Diffusionssperre das Hineinwandern des Aluminiums automatisch stoppt. Da in der vorliegenden Arbeit aber vertikale Transistoren verwendet werden, wirkt sich dieser Vorteil nicht auf die Prozessierung des Sensors aus, weil bei solchen Bauelementen immer ein Kontakt über einem pn-Übergang liegt.

Steigerung der Packungsdichte

Aus fertigungstechnischer Sicht ergeben sich aus den vorher diskutierten Aspekten einige interessante Möglichkeiten. So bieten SOI-Substrate die Möglichkeit die Integrationsdichte zu erhöhen. Das reduziert die Kosten, da pro Fläche mehr Bauteile und so mehr Chips pro Wafer produziert werden können.

Der größte Flächengewinn resultiert aus der oben erwähnten Einsparung von Implantationswannen. Zusätzlich ist es möglich mache n- und p-Gebiete der komplementären Bauelemente gemeinsam zu kontaktieren, was ebenso Fläche einspart. Bei einem SOI-CMOS-Inverter z. B. eignen sich dazu die Drain-Gebiete als Ausgang dieser Schaltung. Dadurch lässt sich ein gemeinsames Kontaktloch für deren Anschlüsse nutzen. Die Zahl der Kontaktlöcher pro Gate sinkt, genauso wie noch einmal der Platzverbrauch.

Erhöhung der Taktfrequenz

Hohe Taktfrequenzen interessieren nicht nur, weil dadurch die nominelle Rechenleistung steigt, sondern lassen sich vor allen als Verkaufs und Werbeargumente gegen die Konkurrenz ins Feld führen. IBM gibt an, dass der Einsatz von SOI-Substraten eine Erhöhung der Taktfrequenz bei sonst gleicher Leistungsaufnahme um bis zu 35 % gegenüber herkömmlichen Substraten möglich macht [Kuh98]. In modernen CMOS-Schaltungen beeinflussen die parasitären Kapazitäten mittlerweile im starken Maße die Verzögerungszeit der Gatter. Da die Kapazitäten durch den Einsatz von SOI-Wafern wesentlich verringert werden können, steigt die relative Treiberfähigkeit der Transistoren. Ebenso erzielt man steilere Ausschaltflanken. Beides kann man für höhere Schaltfrequenzen nutzen.

Reduzierung des Energieverbrauchs

Bei mobilen Anwendungen steht der niedrige Energieverbrauch im Vordergrund, da die zur Verfügung stehende Energie, die in der Regel aus Batterien entnommen wird, begrenzt ist. Für diesen Fall kann die erhöhte Treiberfähigkeit der Transistoren dazu genutzt werden, bei ansonsten gleicher Taktfrequenz die Versorgungsspannung zu verringern. Zusätzlich verringert sich die Leistungsaufnahme bei CMOS-Schaltungen, da in diese die Gesamtkapazität proportional eingeht. IBM spezifiziert für herkömmliche Technologien einen 1,7- bis 3fach höheren Energieverbrauch als für Schaltungen auf SOI-Substraten bei ansonsten gleicher Rechenleistung [Kuh98].

Individuelle Einstellung der Einsatzspannung

Die Einsatzspannung eines MOS-Transistors kann nicht nur über seine Substratdotierung oder die Dicke des Gate-Oxides beeinflusst werden, sondern auch über die Spannung U_{SB} zwischen Source und Substrat (Bulk). Die Einstellung der ersten beiden Parameter ist allerdings für den gesamten Wafer gleich. Deswegen eignet sich zur individuellen Bestimmung der Schwellwertspannung besonders die Potenzialdifferenz U_{SB} . Durch die hervorragende dielektrische Isolation auf SOI-Substraten, kann diese Spannung für unterschiedliche Bereiche oder auch die einzelnen Bauelemente unterschiedliche Werte annehmen. Diese können auch innerhalb der gleichen Schaltung gegenüber Masse positiv und negativ sein.

Freistellen von kristallinem Silizium

Nicht nur für die Fertigung von integrierten Schaltungen zeigen SOI-Substrate ihre Vorzüge. Sie werden auch zur Herstellung in der Mikrosystemtechnik genutzt. Die vergrabene Oxid-Schicht lässt sich sehr einfach als Ätzstopp verwenden. In der Vergangenheit wurden viele nass- und trockenchemische Ätzprozesse entwickelt, die Si selektiv gegenüber SiO_2 ätzen [WMF88].¹ So können z. B. Membranen im

¹z. B. Plasmaätzen mit den Gasen CF_4/O_2 oder SF_6

selbstjustierenden Prozess von der Rückseite des Wafers herausgeätzt werden, ohne dass die Ätzrate sehr genau kontrolliert werden muss. Mit einer SiO_2 -Ätzlösung, die selektiv gegenüber Si ist,² lässt sich das einkristalline Silizium unterätzen und frei hängende Zungen oder Brücken herstellen.

5.2 Vorteile der SOI-Substrate für den FGFET

5.2.1 Vollständige Isolierung

Die Potenziale des Transistors müssen unabhängig von denen des Spannungsteilers eingestellt werden können. Es dürfen keine Leckströme zwischen den unterschiedlichen Potenzialen fließen. Leckströme überlagern den Ausgangsstrom des Transistors additiv und verfälschen so das Messergebnis.

Es wurde schon in den Abschnitten 4.3 und 5.1 darauf hingewiesen: Durch die Verwendung von SOI-Substraten können Leckströme vollständig unterbunden werden. Und das bei gleichzeitigem Einsparen von Prozessschritten.

5.2.2 Hochtemperatur-GasFET

Durch den Einsatz von SOI-Wafern, v. a. in Verbindung mit der Verwendung des vertikalen Transistors gewinnt man neben einer sehr guten Isolation zusätzlich einen größeren Temperatureinsatzbereich. Das soll im Folgenden gezeigt werden.

In einem reinen, nicht dotierten Halbleiter werden durch thermisch angeregte Gitterschwingungen ständig Elektronen-Loch-Paare generiert. Es gleicht also die Anzahl der Elektronen n der Anzahl der Löcher p . Gleichzeitig rekombinieren diese Ladungsträger, indem Elektronen über das verbotene Band zurückspringen. Eine Balance zwischen diesen beiden Vorgängen stellt sich im thermodynamischen Gleichgewicht ein. Die resultierende Konzentration der vorhanden Ladungsträger heißt Eigenleitungs-dichte oder intrinsische Dichte n_i . Sie hängt exponentiell von der Temperatur T und vom Bandabstand ($W_L - W_V$) ab, der wiederum selbst temperaturabhängig ist [Gro67]. Abbildung 5.3 zeigt ihren streng monotonen Anstieg mit der Temperatur.

Durch den Einbau von Dotieratomen bringt man zusätzlich Löcher oder Elektronen in den Halbleiterkristall ein. Damit erhöht man gezielt die Konzentration nur einer Sorte der Ladungsträger. In diesem Zustand ist der Halbleiter extrinsisch.

Physikalisch wird solch ein Überschuss durch das Fermi-Niveau W_F beschrieben [Gro67, Sze69]. Es gibt die Lage an, in der die Hälfte aller Energiezustände besetzt ist. Bei einer Überzahl an Elektronen nähert es sich dem Leitungsband und bei einem Mehr an Löchern dem Valenzband. Mit der Überlegung, dass ein Halbleiter im thermodynamischen Gleichgewicht in jedem Volumenelement elektrisch neutral ist, lässt sich die Lage dieses Fermi-Niveau bestimmen (siehe Anhang A).

²z. B. gepufferte Flusssäure BHF: 7 Teile NH_4 (40 %) : 1 Teil HF (49 %)

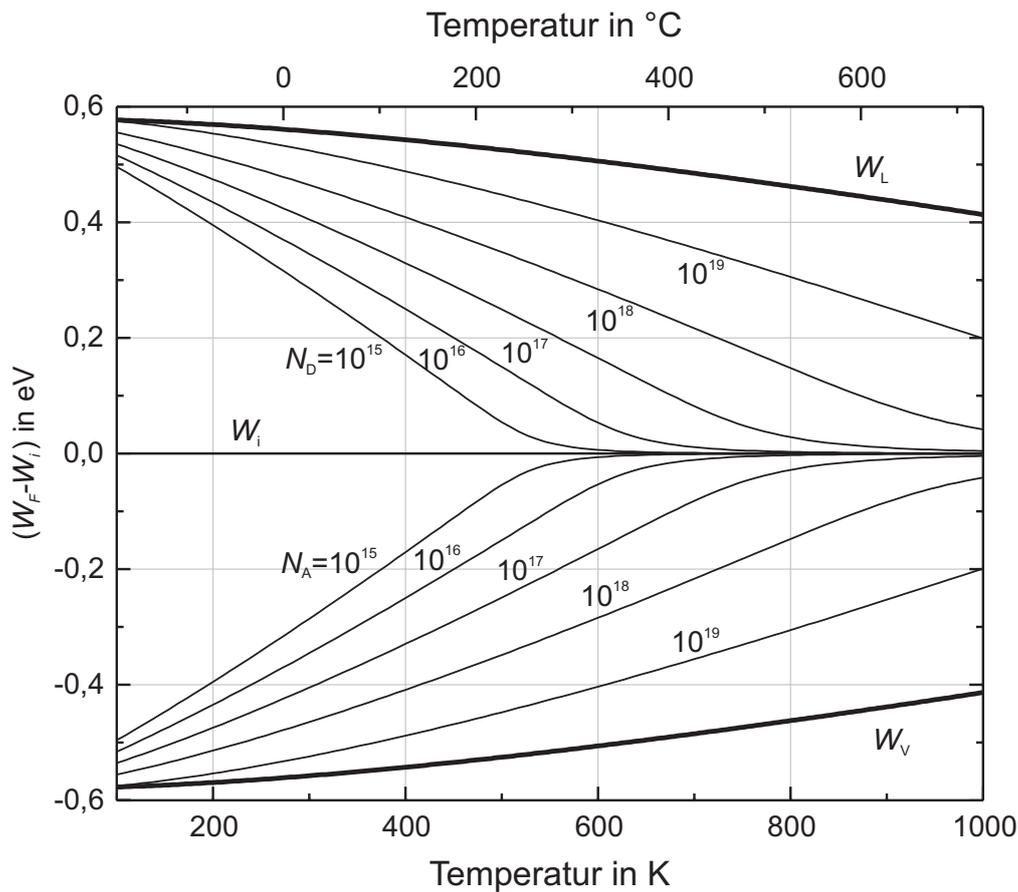


Abbildung 5.2: Lage des Fermi-Niveaus sowie des Valenz- und Leitungsbandes bezogen auf das intrinsische Energieniveau für verschiedene Dotierstoffkonzentrationen (in cm^{-3}) in Abhängigkeit der Temperatur. Aus diesem Graphen lässt sich die Potentialbarriere eines pn-Übergangs ablesen. Der Abstand zwischen zwei Linien passend zur jeweiligen Dotierung des p- und n-Gebietes entspricht der Barriere.

Diese Berechnungen sind für verschiedene Dotierkonzentrationen in Abbildung 5.2 angegeben. Dafür wurden die experimentell ermittelten Werte der intrinsischen Dichte aus [MM54] herangezogen. Auch die Energieniveaus des Leitungs- und Valenzbandes nach [Sze69] sind in Abbildung 5.2 mit eingezeichnet.

Im Bereich ab 100 K bis etwa 400 K überwiegen die extrinsischen Ladungsträger. Das Fermi-Niveau liegt entsprechend der Dotierung dicht am Leitungs- oder Valenzband. Mit zunehmender Temperatur steigt die Anzahl der generierten Elektron-Loch-Paare immer stärker an. Das Fermi-Niveau nähert sich dem intrinsischen Energieniveau, weil demzufolge der Anteil der extrinsischen Ladungsträger sinkt. Ab einer für die Dotierkonzentration typischen Temperatur dominiert die Generation so stark, dass die Differenz der beiden Energieniveaus verschwindet. Der Halbleiter ist nun intrinsisch. In diesem Falle zeigen pn-Übergänge kein charakteristisches Diodenverhalten mehr.

Das gleiche Ergebnis liefert die Betrachtung der Ladungsträgerkonzentrationen. Sie sind beispielhaft für Bor-dotierte Gebiete in Abbildung 5.3 logarithmisch über der Temperatur aufgetragen. Ergänzend ist die intrinsische Ladungsträgerdichte n_i schwarz gestrichelt dargestellt.

Zunächst steigt die Konzentration der Majoritätsladungsträger p sprunghaft an,³ da bei sehr niedrigen Temperaturen die Dotieratome nicht vollständig ionisiert sind. Ab etwa 100 bis 200 K sind näherungsweise alle Dotieratome ionisiert und die Konzentration der Majoritätsladungsträger verläuft entlang eines Plateaus. Das ist der Arbeitsbereich von integrierten Halbleiterschaltungen. Schließlich schwillt die Majoritätsladungsträgerdichte wieder an. Ab dieser Temperatur sind bereits so viele intrinsische Ladungsträger vorhanden, dass sie die Gesamtladungsträgerkonzentration beeinflussen. Ab einer von der Dotierung abhängigen Temperatur überwiegen die intrinsischen Ladungsträger die Konzentration der Dotierstoffatome und die Ladungsträger-Konzentrationen p und n werden durch die intrinsisch generierten Ladungsträger bestimmt. Durch das Dichteprodukt [Sze69]

$$p \cdot n = n_i^2 \quad (5.1)$$

kann man nun auf die Minoritätsladungsträgerkonzentration n schließen. Diese nähert sich natürlich von unten her an die intrinsische Dichte n_i . Es gibt eine von der Dotierung abhängige Temperatur, ab der sind die Dichten der Majoritätsladungsträger p und der Minoritätsladungsträger n annähernd gleich. Der Halbleiter ist nun intrinsisch und ist elektronisch nicht mehr zu verwenden. Je höher der Halbleiter dotiert ist, desto höher ist diese Temperatur. Für herkömmliche CMOS-Schaltungen liegt diese Temperatur bei etwa 300 °C, weil die niedrigsten Dotierungen bei etwa 10^{15} cm^{-3} liegen.

Erfahrungsgemäß fallen aber Silizium-Halleiterschaltungen schon weit früher aus, etwa bei 200 °C. Die Leckströme der in Sperrrichtung geschalteten pn-Übergänge nehmen bereits vorher so stark zu, dass sie die Funktion der Schaltung stören. Besonders sind das die Sperrströme von den Wannens ins Substrat und vom Bulk in die Source-Drain-Gebiete. Die Beziehung für den Sperrstrom I_S einer Diode [Hof90]

$$I_S = q \cdot A \left(\frac{D_p}{L_p} p_{n0} + \frac{D_n}{L_n} n_{p0} \right) \quad (5.2)$$

(mit den Diffusionslängen L_p und L_n , der Elementarladung q und der Querschnittsfläche der Diode A , die Indizes p und n stehen jeweils für das p- und n-Gebiet) macht eine starke Temperaturabhängigkeit deutlich, denn die Diffusionskonstanten D_p und D_n , und die Minoritätsladungsträger p_{n0} und n_{p0} steigen mit der Temperatur und somit nimmt auch der Leckstrom I_S zu. Gleichung (5.2) zeigt einen proportionalen Zusammenhang zwischen Sperrstrom I_S und den Minoritätsladungsträgerkonzentrationen p_{n0} und n_{p0} . Abbildung 5.3 zeigt, dass die Anzahl der Minoritätsladungsträger mit zunehmender Dotierstoffkonzentration

³Im Falle einer Bor-Dotierung sind das die Löcher.

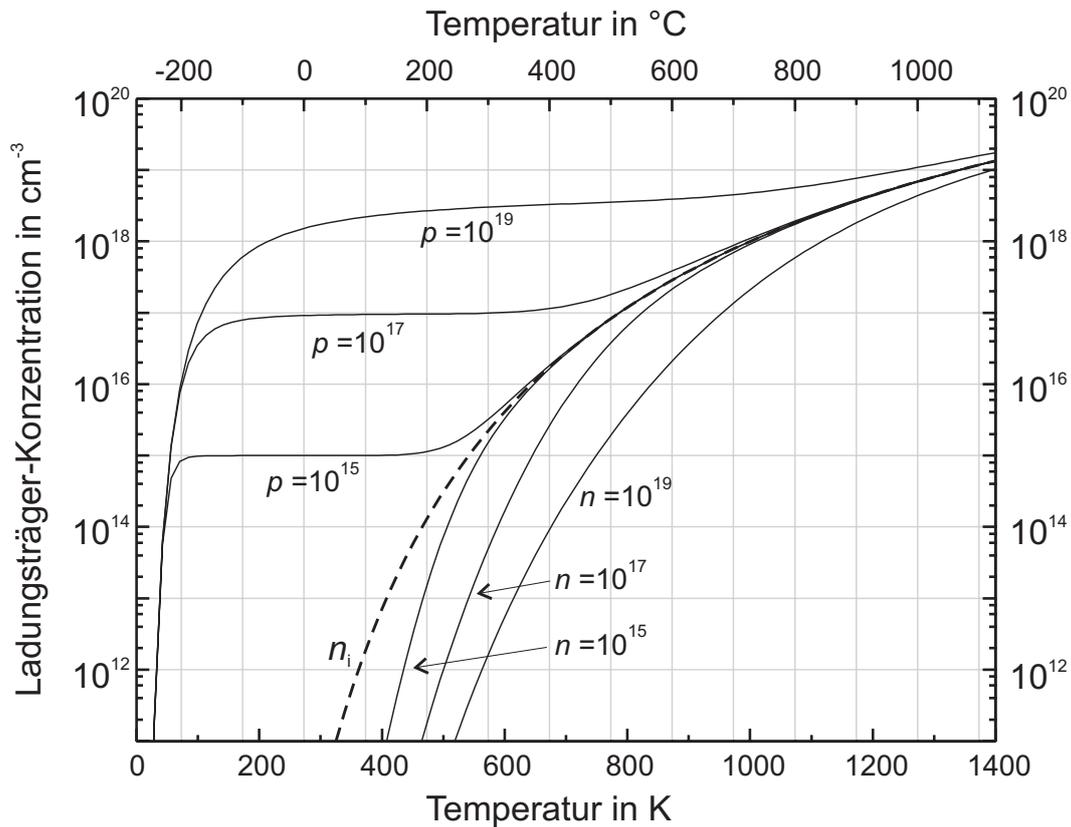


Abbildung 5.3: Majoritätsladungsträgerdichten p und Minoritätsladungsträgerdichten n in Silizium für verschiedene Bor-Dotierungen (in cm^{-3})

schrumpft. Folglich kann man bei konstanter Temperatur den Sperrstrom reduzieren, indem die Dotierung erhöht wird (siehe Abbildung 5.3). Das schwächer dotierte Seite des pn-Übergangs bestimmt maßgeblich den Leckstrom I_S , weil dort die Minoritätsladungsträgerkonzentration größer ist.

Die Folgerung aus diesen Betrachtungen heißt nun: Der Sensor muss für einen erweiterten Temperaturbereich so hoch wie möglich dotiert werden. Das steht aber im Widerspruch zu dem, was man für Isolationswannen fordert. Sie müssen nämlich für die Source- und Drain-Gebiete noch umdotiert werden. Und außerdem ist die Durchbruchspannung umgekehrt proportional zur Dotierkonzentration. Hohe Durchbruchspannungen fordern also niedrige Dotierungen. Üblich sind Konzentrationen von 10^{15} bis 10^{16} cm^{-3} .

Dieses Dilemma kann man lösen, indem die Elemente des Sensors direkt auf einem Dielektrikum aufgebaut werden. Das ist leicht unter der Verwendung von SOI-Substraten machbar. Wie schon im Abschnitt 4.2 angeführt, kann solchermaßen auf die niedrig dotierten Wannen verzichtet werden.

Mit der Forderung, den Transistor sehr hoch zu dotieren, stößt man auf ein weiteres Problem: Mit der Zunahme der Dotierung verringern sich die Raumladungszonen. Damit nimmt ebenfalls die Tunnelweite ab, die die Ladungsträger

überwinden müssen, um von Bulk ins Drain-Gebiet zu fließen. Folglich erhöht sich dieser Leckstrom mit steigender Dotierung.

Das Problem kann man umgehen, wenn keine pn-Struktur verwendet wird, sondern eine pin-Struktur. Mit dem Einbau einer intrinsischen Zone zwischen p- und n-Gebiet werden die Energiebänder auseinander gezogen [Eis01], und derart die Tunnelweite vergrößert. So etwas lässt sich lateral nicht leicht herstellen. Deswegen empfiehlt sich gerade ein vertikales Transistordesign. Bei dieser Bauart wird das Dotierprofil entlang des Kanals nicht durch Implantation oder kleinste Strukturgröße bestimmt, sondern kann durch Schichtabscheidung sehr genau eingestellt werden (siehe Abschnitt 4.2).

Die niedrigsten Dotierungen, die für den vertikalen Transistor eingesetzt werden sollen, liegen in der Größenordnung von 10^{18} cm^{-3} . Ein Vergleich der Konzentrationen mit den Berechnungen aus den Abbildungen 5.2 und 5.3 zeigt, dass der Halbleiter bei niedrigen Konzentrationen viel schneller intrinsisch wird. Für die angenommenen Werte lässt sich der Temperaturbereich um mehrere 100 K steigern. Es könnten Temperaturbereiche bis zu 400 oder 500 °C möglich sein.

5.3 Epitaxie auf SOI-Substraten

Für die Herstellung des Sensors standen SOI-Substrate mit den Parametern wie in der Tabelle 5.1 angegeben zur Verfügung. Das sind gebondete SOI-Wafer, deren Device-Layer auf die gewünschte Schichtdicke geschliffen wird. Eine detailliertere Abhandlung der Herstellungsverfahren für SOI-Substrate findet sich in Anhang B.

Um die Frage zu beantworten, ob die gebondeten SOI-Wafer für die Herstellung des Gassensors verwendet werden können, muss deren Eignung zur Schichterzeugung mit hoher Kristallqualität mittels Molekularstrahlepitaxie geklärt werden.

Entscheiden für Qualität dieser Schichten und der mit ihnen gefertigter Bauelemente, ist eine geringe Defektdichte. Diese setzt sich v. a. aus Punktdefekten, die Rekombinationszentren darstellen, sowie Versetzungen, die wie Kurzschlüsse zwischen den aktiven Bereichen wirken, zusammen. Zum Ermitteln der Kristallqualität haben sich zwei Verfahren etabliert. Zum einen das oben schon genannte Defektätzen und zum anderen das Auswerten von PIN-Dioden. Letzteres Verfahren wird in [Fin00] beschrieben. Das Defektätzen hat den Nachteil, dass damit kei-

	Device	BOX	Handle
Dicke:	$2 \mu\text{m} \pm 0,5 \mu\text{m}$	$2 \mu\text{m} \pm 5 \%$	$500 \mu\text{m} \pm 0,5 \mu\text{m}$
Dotierung:	n / Sb		n / Sb
Leitfähigkeit:	0,008 - 0,02 Ωcm		0,008 - 0,02 Ωcm
Orientierung:	$\langle 1-0-0 \rangle$		$\langle 1-0-0 \rangle$

Tabelle 5.1: Parameter der verwendeten SOI-Substrate

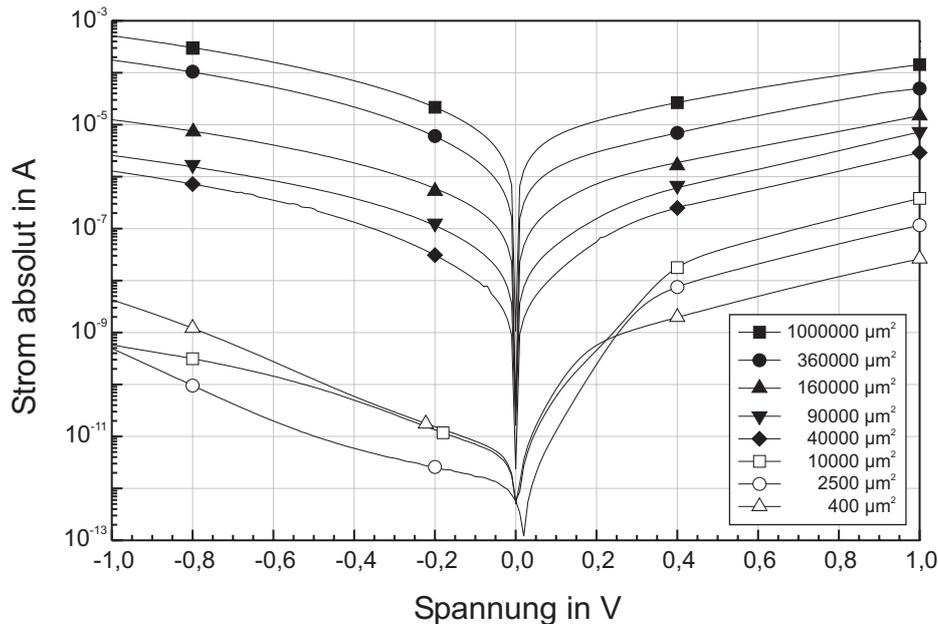


Abbildung 5.4: I-U-Kennlinien von PIN-Dioden, die auf SOI hergestellt wurden

ne Punktdefekte beobachtet werden können. PIN-Dioden geben außerdem direkt Aufschluss, ob sich die hergestellten Schichten für Bauelemente nutzen lassen.

Zur Evaluierung der Eignung gebundener SOI-Substrate wurden PIN-Dioden auf diesen hergestellt. Das Substrat besitzt eine Dotierung mit Arsen und hat einen spezifischen Widerstand von $0,004 \Omega\text{cm}$. Das entspricht einer Konzentration von 10^{19}cm^{-3} . Die i-Zone hat eine Dicke von 100 nm und die darüber liegende p-Schicht eine von 300 nm mit einer Dotierkonzentration von 10^{20}cm^{-3} Boratomen. Beide Schichten wurden mittels Molekularstrahlepitaxie an der Universität der Bundeswehr hergestellt. Durch das anschließende Mesaätzen in die so aufgewachsenen Schichten, wurden vertikale PIN-Dioden erzeugt.

Eine PIN-Diode ist eine spezielle Diode, die zwischen dem n- und p-dotierten Gebiet eine kaum dotierte, fast intrinsische Zone besitzt. Wegen dieser Bauweise fällt ein Großteil der Spannung über der undotierten Schicht ab. Versetzungslinien, die diese durchziehen, stellen Parallelwiderstände dar. Das äußert sich in einem hohen Sperrstrom der Dioden. Ab einer bestimmten Anzahl an Defekten zeigt das Bauteil keine Diodencharakteristik mehr. Um eine Aussage über die Defektdichte, also deren Anzahl pro Fläche zu treffen, haben die Dioden verschiedene Größen. Diese ergeben sich aus der Größe der dazugehörigen Mesafläche. Die zur elektrischen Charakterisierung aufgenommenen Strom-Spannungskennlinien dieser Bauelemente sind in Abbildung 5.4 zu sehen. Die jeweiligen Flächen dieser Dioden sind darin mit angegeben. An den Messungen ist zu sehen, dass es durchaus möglich ist brauchbare Bauelemente mit einer Fläche von $10^4 \mu\text{m}^2$ fertigen. Unter der Annahme, dass bereits eine Versetzungslinie, die von dem n-Gebiet über die i-Zone in das p-Gebiet reicht, das Bauteil zerstört, entspricht das einer

Defektdichte von 10^4 cm^{-2} . Das liegt durchaus im Rahmen der bisher gewachsen und untersuchten Schichten [Feh97, Fin00]. Allerdings kann hier aufgrund fehlender weiterer Messungen keine Statistik betreiben werden, um diese Aussage zu verifizieren. Die relativ hohen Sperrströme weisen trotzdem auf eine große Anzahl an Versetzungen hin, die nicht durch das intrinsische Gebiet hindurchreichen.

Das Ergebnis belegt dennoch, dass kristalline Schichten mittels MBE auch auf SOI-Substraten mit vergleichbarer Qualität abgeschieden werden können. Weiterhin bestätigt es die Richtigkeit der Annahmen aus Abschnitt 4.5.3, die zu den Abmessungen der Transistoren führten.

Um Punktdefekte zu betrachten, muss die Diode in Vorwärtsrichtung untersucht werden. Der Verlauf der Strom-Spannungskennlinie einer idealen Diode lässt sich nach Schokley als analytischer Ausdruck

$$I = I_S(e^{\frac{qV}{kT}} - 1) \quad (5.3)$$

ableiten. In Vorwärtsrichtung steigt der Strom I mit zunehmender Spannung U immer stärker an. In Sperrrichtung nähert sich dieser asymptotisch dem Sperrstrom $-I_S$ an. Das umgekehrte Verhalten zeigt die Diode in Abhängigkeit der Temperatur. Die Konstanten q und k stehen für die Elementarladung bzw. die Boltzmann'sche Konstante. In diesem idealisierten Fall geht man davon aus, dass keine Ladungsträger in der Raumladungszone generiert oder rekombiniert werden. Übertragen auf ideale PIN-Dioden findet auch keine Generation in der i -Zone statt. Diesem Anspruch genügen reale Dioden jedoch nicht. Gerade Punktdefekte stellen Störstellen dar, die Ladungsträger erzeugen oder vernichten. Zu dem Strom I der Diode addiert sich deshalb noch der Generations-Rekombinations-Strom:

$$I_{GR} = I_{GR0}(e^{\frac{qV}{2kT}} - 1). \quad (5.4)$$

Der Strom I_{GR} zeigt einen geringeren exponentiellen Anstieg, weil dieser nicht nur aus einer Sorte Ladungsträger hervorgeht, sondern die Störstellen Elektron-Loch-Paare erzeugen. Für die in Vorwärtsrichtung betrachtete Diode gilt für die Summe der beiden Ströme der Zusammenhang:

$$I + I_{GR} \sim e^{\frac{qV}{nkT}}. \quad (5.5)$$

Die Konstante n wird Idealitätsfaktor oder Emissionskoeffizient genannt. Sie gibt an, welcher der beiden Ströme überwiegt. Diese kann daher Werte zwischen 1 und 2 annehmen. Den Wert 1 hat n bei idealen Dioden. Liegt der Wert n jedoch nahe 2, überwiegt der Generations-Rekombinations-Strom. Man gewinnt den Idealitätsfaktor aus der gemessenen Strom-Spannungsmessung, indem die erste Steigung der Kennlinie in Durchlassrichtung entnimmt und aus dieser anschließend den Koeffizienten berechnet. Für die gefertigten Dioden ergeben sich so Werte zwischen 1,22 und 1,84. Es sind also zum Teil erheblich viele Punktdefekte in den Bauelementen vorhanden. Es muss versucht werden bei der künftigen Herstellung von epitaktischen Schichten diese während des Prozesses auszuheilen.

6 Technologieentwicklung

6.1 Fotolithografie

Die Fotolithografie spielt eine Schlüsselrolle in der Halbleitertechnologie. Sie wird genutzt, um die Vielzahl der abgeschiedenen Schichten in einzelne Bereiche, wie z. B. Leiterbahnen, zu unterteilen. Kern dieser Technik ist ein lichtempfindlicher Fotolack, der durch eine Chrommaske, nur in gewünschten Bereichen bestrahlt wird, wodurch unterschiedliche Ätzraten resultieren, sodass Lack-abhängig die belichteten oder nicht belichteten Bereiche selektiv entlackt werden können. Es entsteht ein positives oder negatives Abbild der Chromstrukturen der Fotolithografiemaske auf dem Wafer. Die stehen gebliebenen Lackstrukturen werden als Maskierung für einen nachfolgenden Prozessschritt genutzt. Nach diesem entfernt man sie wieder. Der Fotolack ist also kein Bestandteil der hergestellten Bauelemente.¹

Bevor der eigentliche Belackungsschritt erfolgt, wird auf die Substratoberfläche eine nur wenige Atomlagen dicke Haftvermittlungsschicht aufgebracht. Dazu verwendet man häufig HMDS (Hexamethyldisilazan). Anschließend wird mittels einer zentrischen Schleuder der Fotolack aufgeschleudert. Die gebräuchlichen Fotolacke bestehen hauptsächlich aus Novolack-Harz, eine fotoaktive Verbindung, und Lösungsmittel. Seine Schichtdicke ist von der Umdrehungszahl der Schleuder, sowie seiner genauen Zusammensetzung abhängig. Abbildung 6.1 zeigt resultierenden Lackdicken in Abhängigkeit der Umdrehungszahl. Die eingezeichnete Ungenauigkeit rührt nicht von den Messfehlern her, denn diese sind kleiner, sondern daher, dass die exakte Dicke zusätzlich von anderen Parametern wie Länge und Dauer des Ausbackens und der Entwicklungszeit beeinflusst werden. Auch die genaue Behandlung beim Aufschleudern selbst ist entscheidend. Für den Lack AR-P 3840 z. B. zeigt Abbildung 6.1 einen starken Abfall der Schichtdicken zwischen 1000 und 2000 Umdrehungen pro Minute. Dieses Verhalten zeigt dieser Lack, weil sich beim Aufschleudern mit kleineren Geschwindigkeiten als 3000 min^{-1} ein Wulst am Rand des Wafers bildet. Dieser hat eine Höhe von bis zu einigen $10 \mu\text{m}$ und verhindert den Kontakt zwischen Wafer und Maske während der Kontaktbelichtung. Deshalb wird nach dem eigentlichen Aufschleudern dieses Lackes mit niedrigen Geschwindigkeiten die Umdrehungszahl noch einmal für eine Sekunde auf 4000 min^{-1} erhöht, um den Lackwulst zu entfernen. Allerdings verringert sich dabei noch einmal die Schichtdicke.

¹Ausnahmen bilden hier Lötstoppmasken, wie z. B. Beispiel Peters ELPIMER 2467. Dabei fungiert der Lack selbst als Passivierung.

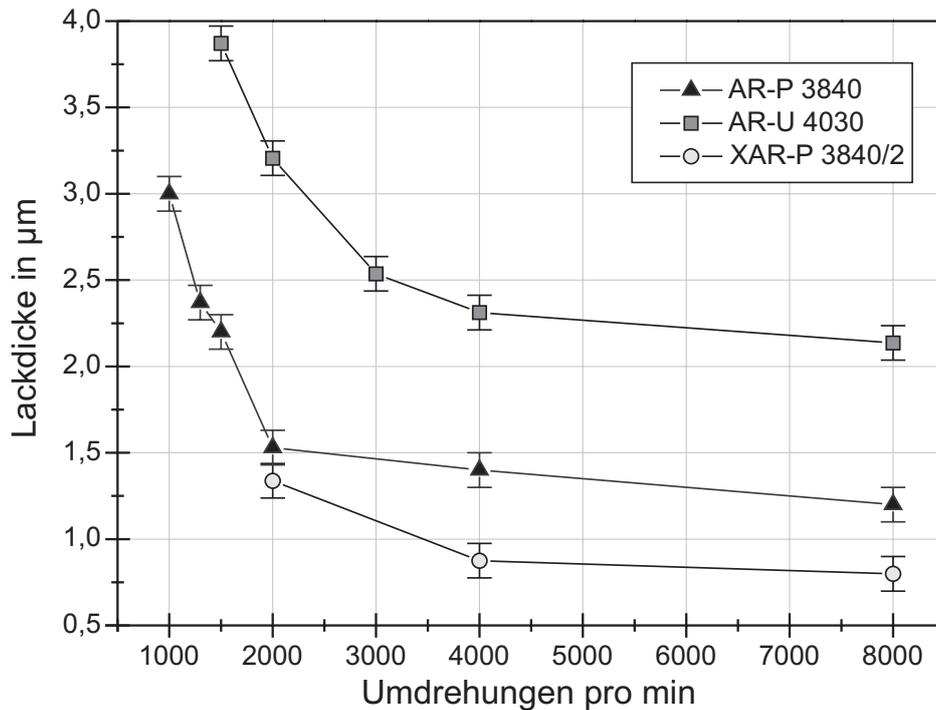


Abbildung 6.1: *Abhängigkeit der Dicke verschiedener Fotolacke von der Umdrehungszahl des Aufscheuderns*

Durch das folgende Vorbacken (engl. Prebake) bei ungefähr 100 °C wird das Lösungsmittel ausgetrieben und die Haftung zum Substrat verbessert. Nach einer kurzen Abkühlphase wird der Wafer in einem Belichter zur Maske justiert und belichtet. In der Regel wird hierauf der Fotolack entwickelt. Zum Schluss werden die Lackstrukturen noch einmal gebacken (engl. Postbake). Hiermit vernetzen die Polymere noch weiter und das Lösungsmittel wird vollständig ausgetrieben. Dadurch erhöht sich die chemische und mechanische Widerstandsfähigkeit des Lackes.

Das exakte Einhalten der Temperaturen und Zeiten bei den einzelnen Schritten ist Voraussetzung für ein Gelingen der Fotolithografie, denn die Eigenschaften des Lackes ändern sich in Abhängigkeit mehrerer Parameter. So nimmt z. B. die Belichtungsempfindlichkeit des Lackes mit steigender Temperatur des Vorbackens, niedrigerer Entwicklerkonzentration und Entwicklertemperatur ab. Andererseits steigt der Kontrast des Lackes mit geringerer Temperatur des Vorbackens und niedrigerer Entwicklungskonzentration und -temperatur. Die genauen Parameter können aber nicht pauschal angegeben werden, da diese sehr mit der Beschaffenheit des Substrates, wie z. B. Untergrund, Topologie und Strukturgrößen variieren. Es bedarf der Optimierung für jeden einzelnen Prozessschritt.

Bei dieser Anpassung muss immer wieder ein Kompromiss gefunden werden, um den Anforderungen eines speziellen Prozessschrittes gerecht zu werden. Durch die Wahl der Temperatur des Ausbackens vor dem Belichten (Prebake) wird der Kontrast beeinflusst. Geringere Temperaturen bedeuten dabei höherer Kontrast,

woraus glattere Seitenwände der Lackstrukturen resultieren. Auf der anderen Seite birgt das Ausbacken mit geringeren Temperaturen aber die Gefahr, dass der Fotolack nicht genügend aushärtet und somit während der Kontaktbelichtung stellenweise an der Maske anhaftet und abgerissen wird oder sogar der ganze Wafer an der Maske kleben bleibt.

An der Universität der Bundeswehr werden für die Fotolithografie die zwei Fotolacksysteme XAR-P 3840/2 und AR-U 4030 der Fa. Allresist verwendet. In den Tabellen C.1 bzw. C.2 im Anhang C werden die standardisierten Abläufe der Strukturierung dieser Fotolacke angegeben. Beim Vergleich der Tabelle C.1 mit dem oben beschriebenen allgemeinen Ablauf fällt auf, dass das Ausbacken zur chemischen Stabilisierung nicht erst zum Schluss (Postbake) sondern direkt nach dem Belichten (Postexposure Bake) erfolgt. Das hat zwei Vorteile. Es ist erforderlich, den Fotolack auszubacken, damit dieser eine genügende chemische sowie mechanische Festigkeit bekommt. Unterbleibt dieses Ausbacken, hält er z. B. einer Ätzung mit Buffered HF nicht stand oder die Ätzraten beim Plasma-Ätzen sind zu groß um tiefere Gräben herzustellen. Allerdings fängt der Lack beim Erwärmen ab einer Temperatur von etwa 110 °C an zu fließen. Deshalb zerlaufen die Flanken der Lackstrukturen beim Erhitzen nach dem Entwickeln und verrunden. Durch ein Ausbacken vor dem Entwickeln bleiben die Flanken stabil, denn der Lack ist noch homogen auf dem Wafer verteilt und kann so nicht weglaufen. Zum anderen werden Rauigkeiten der Fotolackstruktur ausgeglichen, weil im Bereich der Lackflanken vor dem Entwickeln die fotoaktivierten und nicht aktivierten Gebiete durch Diffusion ineinander fließen. Daraus resultieren geglättete Fotolackflanken. Versuche zeigten, dass der so behandelte XAR-P 3840/2 genügend stabil gegen verschiedene Ätzlösungen und Trockenätzen ist.

Der Nachteil dieser Methode ist die sich ergebende geringere Auflösung bzw. eine größere minimale Strukturgröße, die mit diesem Ablauf erreicht wird. Durch das Verlaufen verschieden belichteter Bereiche ineinander geht der Kontrast verloren. Die Übergänge sind nicht mehr so scharf. Trotzdem sind mit dieser Technik und dem verwendeten Lacksystem Strukturen in der Größe von einem Mikrometer herstellbar. Für die Fertigung des Sensors im Rahmen dieser Arbeit ist das ausreichend.

Ein Vergleich des Temporns jeweils vor und nach dem Entwickeln zeigt Abbildung 6.2. Die erste Probe (Abbildung 6.2a) wurde nach dem Entwickeln bei 110 °C für 2 min ausgebacken. Deutlich sind die abgerundeten Kanten, die aufgrund der Erhöhung der Temperatur verlaufen sind, bei 6000facher Vergrößerung zu sehen. Eine zweite Probe wurde noch vor dem Entwickeln nach dem Belichten bei 110 °C für 4 min ausgebacken. Bei 5000facher Vergrößerung sieht man die nahezu senkrechten Fotolackflanken (Abbildung 6.2b). Sie bilden mit der planen Oberfläche des strukturierten Fotolacks eine Kante.

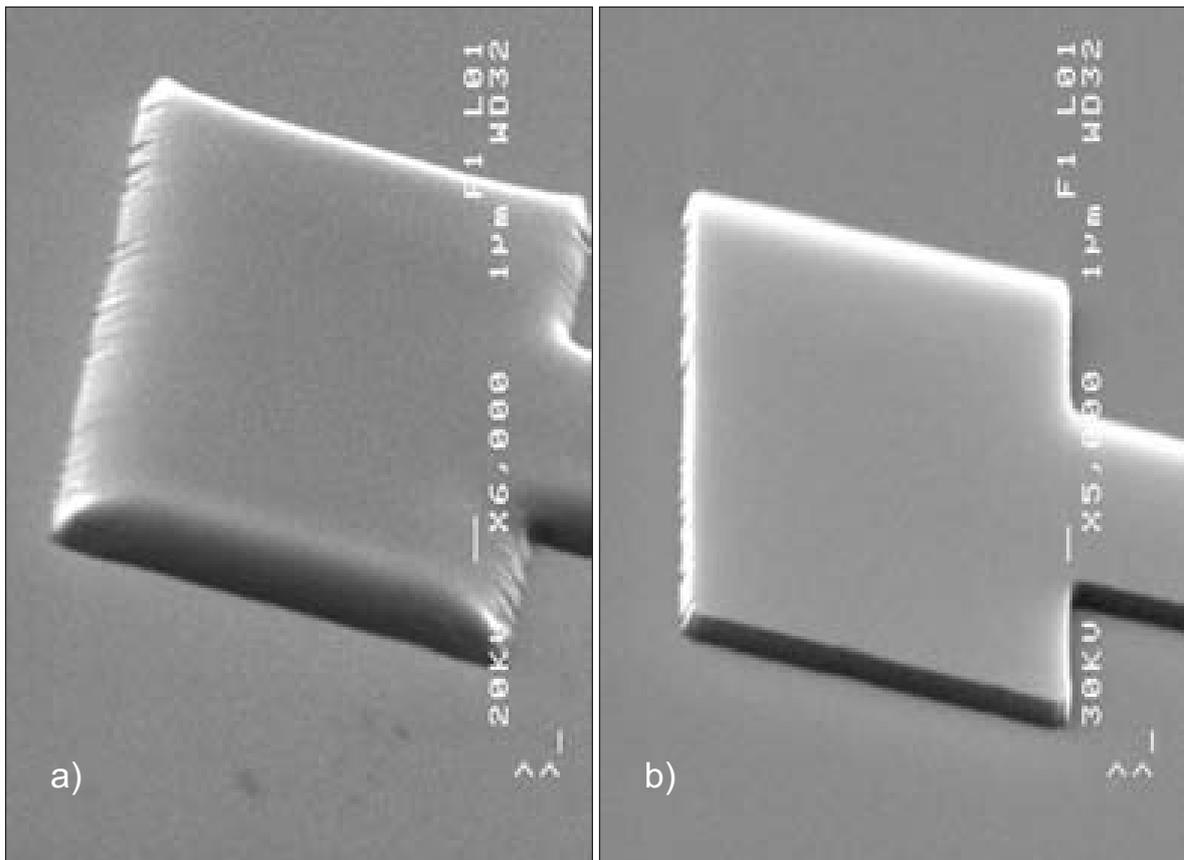


Abbildung 6.2: *Fertig entwickelter und unterschiedlich thermisch stabilisierter Fotolack XAR-P 3840/2: a) Postbake 2 min bei 110 °C, 6000fache Vergrößerung; b) Postexposure Bake 4 min bei 110 °C, 5000fache Vergrößerung.*

6.2 Substratreinigung

Um eine hohe Reproduzierbarkeit und Ausbeute zu garantieren, muss das Substrat genau definiert sein. Das betrifft gleichermaßen seine Oberfläche. Optimal wäre, wenn diese nur aus einem bestimmten Material besteht und keinerlei Kontaminationen enthielte. Das ist bei realen Substraten nicht der Fall. Typischerweise findet man auf diesen:

- Organische Anlagerungen,
- Metallische und ionische Verunreinigungen und
- Partikel.

Erstgenannte Verschmutzungen rühren z. B. von Lösungsmitteln oder Lacken her. Auch Metalle und Ionen stammen oft aus der Prozessführung selbst (z. B. K^+ -Ionen aus der KOH^-). Beide Arten findet man gleichmäßig über den gesamten Wafer verteilt. Im Gegensatz dazu trifft man Partikel nur unregelmäßig auf dem

Wafer verteilt an. Sie setzen sich aus der Luft und aus Flüssigkeiten auf der Substratoberfläche ab. Sie können sowohl organisch als auch metallisch sein. Ein Beispiel dafür sind Kunststoff oder Metallabrieb von Anlagen. Von Personen stammen ebenfalls v. a. organische Partikel.

Um Partikel von vornherein so weit wie möglich zu vermeiden, ist es unabdingbar, die Herstellung von Halbleiterbauelementen in einer Reinraumumgebung durchzuführen. Genauso reduzieren partikel-, metall- und organikarme Chemikalien eine Kontamination. Ansonsten trägt das richtige Verhalten des Personals im Reinraum zu einer hohen Ausbeute bei.

Zwischen den einzelnen Prozessschritten ist es immer wieder notwendig, das Substrat von Partikeln und Rückständen zu reinigen. Das hat im Wesentlichen drei Gründe:

- Für die Qualität einer nachfolgenden Epitaxie ist eine saubere Oberfläche wichtig.
- Eine Verschmutzung der Anlagen durch eine Verschleppung von Verunreinigungen muss vermieden werden.
- Nachfolgende Prozessschritte dürfen nicht durch verbliebenen Chemikalien gestört werden.

Für ein Aufwachsen kristalliner Schichten mittels MBE spielt eine saubere Oberfläche eine entscheidende Rolle [GOT86]. Besonders Kohlenstoffreste sind für die Epitaxie schädlich. Sie können z. B. Cluster oder Siliziumcarbid bilden. Als Folge entstehen von diesen Einschlüssen aus Versetzungslinien, die die Kristallqualität mindern.

Die Kontaminationen auf den Substraten können Anlagen verschmutzen. Sie desorbieren während der Prozesse und lagern sich auf den Kammerwänden ab. Später lösen sie sich wiederum und verunreinigen die Substrate. Auf diese Weise kann es sogar zu einer Verschleppung der Verunreinigungen kommen. Die gleiche Gefahr besteht ohne Unterschied bei nasschemischen Prozessen. Chemikalien eines vorherigen Prozessschrittes verbleiben auf dem Substrat und beeinflussen die weitere Prozessführung negativ. So ist es möglich, dass durch nicht vollständig entfernte Fotolackreste Maskierungseffekte hervorgerufen werden. Das kann z. B. beim Ätzen einer Metallisierungsschicht zu Kurzschlüssen zwischen zwei Leiterbahnen führen.

Zum Reinigen der Oberfläche von Substraten wurden verschiedene Methoden entwickelt. Einige davon findet man in [DeJ98] und [Feh97]. Das einfachste Verfahren dabei ist das Spülen im Rinser. Dabei können jedoch nur größere Partikel und wasserlösliche Substanzen entfernt werden.

RCA-Clean

Als Standard haben sich in der Halbleitertechnologie der so genannte „RCA-Clean“ [KP70] und Modifikationen davon etabliert. Diese nutzen alkalische und saure

	Lösung	Konz.	Temperatur	Dauer
1. „RCA1“	1 Teil NH ₃	29 %	70 °C	15 min
	1 Teil H ₂ O ₂	30 %		
	7 Teile H ₂ O			
2. Ätzen	1 Teil HF	49 %	RT	15 s
	100 Teile H ₂ O			
3. „RCA2“	1 Teil HCl	37 %	70 °C	15 min
	1 Teil H ₂ O ₂	30 %		
	6 Teile H ₂ O			

Tabelle 6.1: Prozessfolge des nasschemischen „RCA-Cleans“

Reinigungsbäder. Die nasschemische Reinigung, wie sie an der Universität der Bundeswehr durchgeführt wird, basiert auch auf diesen. Der genaue Ablauf ist in Tabelle 6.1 angegeben.

Im ersten Schritt wird in einer alkalischen Lösung das natürliche Oxid der Siliziumoberfläche weggeätzt. Dabei lösen sich Partikel. Die Ätzrate für Silizium ist dabei sehr gering. Die meisten organischen Verunreinigungen werden hierbei entfernt, da diese sich im Allgemeinen sehr gut in Ammoniumhydroxid² lösen. Gleichzeitig bildet sich ein chemisches Oxid. Nachdem ersten Reinigungsschritt hat es eine ungefähre Dicke von 2 nm. Viele Metallverunreinigungen werden in diesem eingeschlossen.

Im folgenden Schritt wird das erzeugte Oxid in verdünnter Flusssäure wieder entfernt. Günstig ist eine nicht zu hohe Konzentration, da sonst zu große Ätzraten die Siliziumoberfläche aufrauen.

Als Nächstes werden im sauren Bad alkalische Ionen und nicht im alkalischen Milieu lösliche Hydroxide beseitigt. Dieses Bad zeigt sich des Weiteren sehr effektiv in der Beseitigung von Metallen, weil es Salzsäure enthält, die alle Metalle mit negativem Normalpotenzial löst. Spuren anderer Metalle wie z. B. Kupfer werden durch das Wasserstoffperoxid³ aufoxidiert. Die dabei entstehenden Komplexe sind im sauren Milieu löslich und können so ebenfalls gelöst werden. Im dritten Schritt wird nochmals ein chemisches Oxid gebildet. Es wird genutzt, um den Wafer bis zum Beginn des folgenden Prozesses zu schützen.

Für die Durchführung dieser Reinigung ist die Einhaltung der Temperaturen besonders wichtig. Steigt diese über 70 °C zerfällt das Wasserstoffperoxid zu Wasser und Sauerstoff und das als Ammoniumhydroxid gelöste Ammoniak wird wieder abgegeben. Dies führt zu erhöhter Rauigkeit der Oberfläche und zu einer zu geringen Reinigungswirkung.

²NH₄OH³H₂O₂

Caro-Clean

Eine andere oft verwendete nasschemische Reinigung, die auch in der Universität der Bundeswehr verwendet wird, ist der sog. „Caro-Clean“⁴. Dieser eignet sich besonders zur Entfernung organischer Verunreinigungen. Hauptsächlich löst man damit Fotolacke ab. Diese Reinigungslösung setzt sich gewöhnlich aus drei bis fünf Teilen Schwefelsäure und einen Teil Wasserstoffperoxid zusammen.

Beim Mischen dieser beiden Komponenten entsteht in der Gleichgewichtsreaktion



Caro'sche Säure⁵. Damit das Gleichgewicht zur rechten Seite hin verschoben wird, muss der Ansatz erwärmt werden. Allerdings verflüchtigt sich dabei, wie oben erwähnt, das Wasserstoffperoxid. Deshalb wird die Schwefelsäure auf etwa 80 °C erhitzt und Wasserstoffperoxid erst kurz vor der eigentlichen Reinigung zugegeben. Das entstehende Wasser verdünnt die Schwefelsäure in einer stark exothermen Reaktion. Die Lösung erhitzt sich dadurch auf 150 bis 180 °C. Die plötzlich entstehende Wärme unterstützt die chemische Reinigung zusätzlich. Der zu säubernde Wafer wird darin 10 bis 15 Minuten belassen. Anschließend spült man ihn gründlich mit DI-Wasser im Rinser.

An der Universität der Bundeswehr werden solche Spülprozesse in Bechergläsern durchgeführt. Das birgt die Gefahr, aufschwimmende Partikel beim Herausnehmen der Wafer wieder mit aufzunehmen. Ein kontinuierliches Spülen wäre dem vorzuziehen. Das ist jedoch in einem universitären Betrieb aus Kostengründen nicht machbar.

Nach den nasschemischen Prozessen werden die Wafer in einer Trockenschleuder getrocknet. Das vermeidet im Gegensatz zum Trocknen im Umluftofen, dass das verdunstende Wasser die in ihm gelöste Substanzen auf dem Substrat zurücklässt.

6.3 Oxidherstellung und Strukturierung

Die erste Oxidschicht soll hauptsächlich als Dielektrikum für einen Kondensator dienen. Dieser Kondensator bildet mit dem Luftspalt zusammen einen Spannungsteiler. Es ist aber gewünscht, dass Spannungsänderungen im Luftspalt mit einem möglichst hohen Verhältnis auf das Gate übertragen werden. Deshalb ist es hier wichtig eine kleine Kapazität zu erzeugen. Man muss also einen großen Abstand der Elektroden fordern und hier in diesem Fall zusätzlich noch eine niedrige Dielektrizitätskonstante. Außerdem muss das Dielektrikum thermisch und mechanisch stabil genug für die folgenden Prozesse sein. Aus diesen Gründen ist ein Siliziumoxid einem Siliziumnitrid vorzuziehen.

⁴Diesem ganz ähnlich ist der sog. „Piranha-Clean“. Zu dessen Erzeugung wird Ozon in Schwefelsäure eingeleitet, statt Wasserstoffperoxid zu zugeben.

⁵H₂SO₅

6.3.1 Oxidherstellung

Es gibt verschiedene Möglichkeiten Siliziumdioxidschichten zu herzustellen. Oxide klassifiziert man in Thermische und Abscheideoxide.

Abscheideverfahren, sind z. B. CVD- (Chemical Vapour Deposition), Plasma oder Sputter-Prozesse. Sie haben den Vorteil das die Prozesstemperaturen zum größten Teil unterhalb von 700 °C liegen. Abgeschiedene Siliziumdioxidschichten wachsen im Gegensatz zu thermischen Oxiden unabhängig vom Material der Unterlage auf. Die Topologie vorher erzeugter Strukturen wird nicht verändert, sonder nur überwachsen. Dafür sind die dielektrischen Eigenschaften der mit diesen Methoden hergestellten Oxide nicht so gut.

Das Herstellen thermischer Siliziumdioxide ist den CVD-Prozessen sehr ähnlich, nur, dass das Silizium, welches zur Reaktion benötigt wird, nicht aus der Gasphase, sondern aus dem Festkörper selbst geliefert wird. Für eine Trockenoxidation wird der Sauerstoff in Form von molekularem Gas zugeführt. Bei einer Feuchtoxidation stammt der Sauerstoff aus dem Wasserdampf, der dem Ofen zur Reaktion eingeblasen wird. Die Grenzfläche verschiebt sich während des Prozesses in das Silizium hinein. Deshalb sind die sauberen Oberflächen nicht ganz so zwingend notwendig, wie bei den anderen Verfahren.

RTP-Prozesse (Rapid Thermal Processing) zeichnen sich dadurch aus, dass sie das Temperaturbudget nicht so sehr belasten wie das Herstellen thermischer Oxide. Genauere Abhandlungen dieser Technologien finden sich in der Literatur [Ada83, Kat83, WMF88].

An der Universität der Bundeswehr in München stehen keine Abscheideoxide zur Verfügung, deshalb muss auf ein thermisches Verfahren zurückgegriffen werden. Ohnehin ist die Qualität thermischer Oxide besser. Es ist nicht notwendig, in dem ersten Prozessschritt der Herstellung des Sensors auf das Temperaturbudget zu achten. Die Feuchtoxidation eignet sich hier besonders, weil dieses Verfahren höhere Wachstumsraten hat als die Trockenoxidation und hier eher dicke Oxide gewünscht sind.

Das Feuchtoxid wird in einem gewöhnlichen Rohrofen der Firma Inotherm bei einer Temperatur von 1050 °C gewachsen. Die Dauer der Oxidation beträgt 30 min. Zur Erzeugung des Wasserdampfes wird molekularer Wasserstoff und Sauerstoff verbrannt.

Beim Oxidieren von Silizium wächst 45 % des Siliziumdioxidfilmes in das darunter liegende Silizium hinein. Das bedeutet beim Herstellen eines 300 nm starken thermischen Oxides, dass eine Schicht einkristallines Silizium mit Dicke von 135 nm verbraucht wird und der Schichtstapel insgesamt nur um 155 nm zunimmt. Dieser Volumenschwund im Silizium muss unter Umständen, wie z. B. bei der Verwendung von SOI-Substraten, mitberücksichtigt werden. Für die Herstellung des in dieser Arbeit beschriebenen Sensors hat die Herstellung des Dielektrikums durch eine Oxidation den Vorteil, dass das Siliziumdioxid gleichzeitig als Opferoxid für die Reinigung verwendet werden kann. Durch die hohen Temperaturen werden

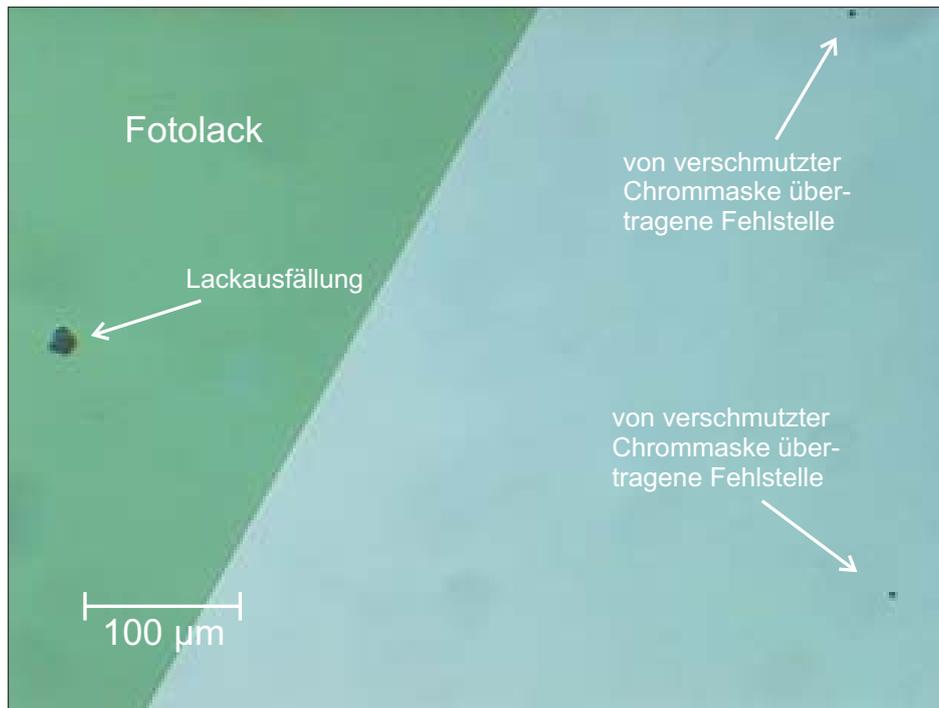


Abbildung 6.3: Fotolackablagerungen und -ausfällungen

auf der Waferoberfläche anhaftende Kohlenwasserstoffe verbrannt und sonstige Verunreinigungen entweder in das Substrat oder in das Oxid getrieben [Feh97].

6.3.2 Herstellen der Lackmasken

Die thermisch gewachsene Siliziumdioxidschicht wird mittels Fotolithografie und Ätzen strukturiert. Bei fehlender Sorgfalt während der Lithografie findet man in der entwickelten Lackmaske eventuell Fehlstellen, wie sie in Abbildung 6.3 markiert sind. Diese wirken während des Ätzens dann wie Mikromasken und verhindern das vollständige Freilegen des Siliziums. Ablagerungen, die in der Ätzmaske stecken, stören den Prozess nicht (linke Seite auf Abbildung 6.3). Dort jedoch, wo der Fotolack weggenommen werden soll, bleiben Reste als kleine Klümpchen zurück (rechte Seite auf Abbildung 6.3). Die daraus resultierenden „Inselchen“ aus Siliziumdioxid stören danach wiederum das kristalline Aufwachsen bei der anschließenden Epitaxie und vermindern die Kristallqualität. Abhilfe kann durch folgende Vorkehrungen getroffen werden.

Vorkehrung bei Hinzufügen des Fotolackes

Wichtig bei der Fotolithografie in diesem Prozessschritt ist die Verwendung von frischem Fotolack und die Durchführung der Lithografie unter Reinraumbedingungen. Bei langen Stehzeiten des Fotolackes bilden sich Ausfällungen und Ablagerungen. Gelangen diese auf den Wafer und werden aufgeschleudert, setzen sie

sich auf diesem fest. Beim Entwickeln des Lackes werden diese dann nicht mehr abgelöst und verbleiben dort.

Mit Hilfe neuer sauberer Pipetten werden diese Fehler auch vermieden, weil der Lack gerade in den Pipetten sehr schnell eintrocknet.

Vorkehrungen während der Entwicklung

Ähnliche Rückstände finden sich ebenfalls nach einem ungünstigen Entwickeln im Becherglas auf dem Wafer. Diese Ablagerungen führen genauso zu den oben genannten Problemen bei der nachfolgenden Prozessführung.

Sofort nach dem Eintauchen in den Entwickler löst sich der belichtete Fotolack. Dabei bilden sich deutlich sichtbar rotviolette Schlieren in der Lösung. Beim einfachen Herausnehmen des Substrats aus der Lösung, schlagen sich diese auf dem Wafer nieder und verbleiben als kreisrunde tropfenförmige Lackreste zurück. Für eine industrielle Fertigung ließe sich dies durch eine Sprühentwicklung, die den gelösten Fotolack wegschwemmt, verhindern. Aus Kostengründen wird die Entwicklung des Fotolackes an der Universität der Bundeswehr aber als Tauchentwicklung in Bechergläsern durchgeführt. Um in diesem Verfahren eine möglichst fehlerfreie Fotolackmaske herzustellen, muss die Entwicklungslösung durch leichtes Schütteln ausreichend bewegt werden. So können sich die Schlieren rasch auflösen und der Entwickler färbt sich homogen transparent rosa bis violett. Ein bloßes Rühren, z. B. mit einem Magnetrührer, reicht nicht aus, denn dadurch bildet sich ein Strudel im Becherglas, in dessen Zentrum sich die Schlieren sammeln und sich in der Mitte des Wafers ablagern.

Als Entwicklungsstopp wird das Substrat in ein mit DI-Wasser gefülltes Becherglas getaucht und umspült, indem es reichlich bewegt und anschließend abgebraust wird. Dies verhindert, dass sich Fotolackreste wieder anlagern.

Vorkehrungen vor Benutzung der Chrommasken

Die Hauptquelle für Fehlstellen in der Lackmaske sind die Chrommasken selbst. Im Verlauf der Kontaktbelichtung wird der belackte Wafer an die Chrommaske gepresst, was unweigerlich dazu führt, dass Lackreste an ihr fest haften. Während der nächsten Belichtung wirken diese Reste wie von Chrom maskierte Stellen und werden so in die Lackmaske übertragen (siehe Abbildung 6.3). Da diese ungewollten Strukturen einen Durchmesser von bis zu 50 μm haben, lassen sie sich nicht durch ein langes Überentwickeln entfernen.

Das alleinige Spülen der Chrommasken mit Aceton zum Entfernen jener Verunreinigungen hat außer einer neuen Verteilung keinen Effekt. Gänzlich abzuraten ist von einem Ultraschallbad, weil das die Chrommasken mechanisch beschädigt. Aus demselben Grund sollte ein Abwischen mit (selbstverständlich fusselfreien) Tüchern nur im Ausnahmefall stattfinden. Aber auch solche Maßnahmen verringern die Verschmutzungen nicht vollständig. Nur eine kombinierte Reinigung der Masken aus dem in 6.2 beschriebenen Caro-Clean (in diesem Fall reichen 60-80 °C)

und Kochen in einem Gemisch aus Aceton und Isopropanol bei 80 °C löst diese vollständig.

Weitere Vorkehrungen

Empfehlenswert ist statt einer Kontaktbelichtung eine Justierung mit Hilfe der Proximity-Fähnchen und eine Proximity-Belichtung, denn so kommt die Maske gar nicht erst mit Fotolack in Berührung. Des Weiteren muss nach einer fehlgeschlagenen Belichtung oder Entwicklung der Fotolack wieder vollständig vom Wafer abgelöst werden. Wie soeben erwähnt, kann der Fotolack nicht rückstandsfrei durch bloßes Kochen in Aceton entfernt werden. Das gilt gleichermaßen für Masken und Substrate. Geeignet sind die in Abschnitt 6.2 geschilderten RCA-Clean und Caro-Clean sowie ein Kochen in einem Gemisch aus Isopropanol und Aceton oder Remover.

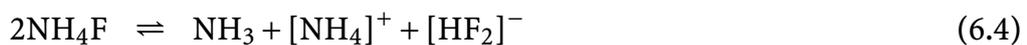
6.3.3 Oxidätzen

Das Ätzen des Siliziumdioxids kann physikalisch, trocken- oder nasschemisch erfolgen. Trockenchemische und physikalische Ätzverfahren zeigen nicht so hohe Selektivitäten wie Nasschemische, und greifen von daher das Siliziumsubstrat wesentlich stärker an. Daraus resultiert eine größere Rauigkeit. Da sich bei dem später folgenden Epitaxieschritt glatte Oberflächen als günstig erwiesen haben, soll jedoch das Silizium des Substrates möglichst wenig angeraut werden. Demzufolge wird hier ein nasschemischer Schritt angewendet.

Eine oft verwendete Lösung zum Ätzen von Siliziumdioxid ist gepufferte Flusssäure oder kurz BHF. Sie zeigt Selektivitäten zu Silizium größer als 100. Außerdem greift diese Lösung Fotolacke nicht so sehr an wie die ungepufferte Lösung. Siliziumdioxid wird in der Reaktion



zu dem flüchtigen Produkt Siliziumtetrafluorid⁶ umgesetzt. Eine gepufferte Säure besteht aus einer schwachen Säure und ihrem Salz einer starken Lauge. Für BHF wird dazu Ammoniumfluorid⁷ genommen. In der Lösung stellen sich die Gleichgewichtsreaktionen



ein. Das Salz erhöht zunächst die Anzahl der Fluoridionen und liefert sie nach, wenn diese während der Reaktion der Gleichung (6.2) verbraucht werden. Gelangt Säure in die Lösung oder entsteht sie in ihr, reagieren die Produkte der rechten Seite der Gleichung (6.3) und (6.4) zu Flusssäuremolekülen. Die zugegebene

⁶SiF₄
⁷NH₄F

Lösung	Konz.	Menge
NH ₄ F	40 %	880 ml
HF	49 %	100 ml
H ₂ O		520 ml

Tabelle 6.2: Zusammensetzung der gepufferten Flusssäure (BHF)

Base reagiert mit dem Ammoniumfluorid und wird zu Wasser und weiteren Fluoridionen umgewandelt. Auf diese Weise wird durch die Zugabe des Salzes der pH-Wert gesenkt und anschließend konstant gehalten. Dadurch bleiben die Ätzraten gleich. Nebenbei schützt der niedrige pH-Wert den Fotolack.

Die gepufferte Flusssäure zeigt 2- bis 3-mal höhere Ätzraten als verdünnte Flusssäure mit gleicher Säurekonzentration, weil eine höhere Anzahl an Fluoridionen vorliegt. Das ist auch der Grund für die starke Abhängigkeit der Ätzrate von dem Anteil an Ammoniumfluorid in der Lösung.

Für diesen Prozessschritt wird ein Lösungsansatz entsprechend Tabelle 6.2 verwendet. Mit ihm beträgt die Ätzrate auf den hier verwendeten Substraten 51 nm^{nm}/min.

Um die Strukturen richtig auf den Wafer übertragen zu können, ist es wichtig, dass die Fotolackmaske dem Ätzmedium lange genug standhält. Für diesen Prozessschritt wurden die Fotolacke XAR-P 3840/2 und AR-P 3840 der Firma Allresist auf die Ätzbeständigkeit gegen gepufferte Flusssäure getestet. Die Abbildungen 6.4 zeigen Profilometermessungen während verschiedener Phasen des Prozessschrittes. Dabei ist das Höhenprofil, senkrecht zur Waferoberfläche gemessen, entlang einer gewissen Länge aufgetragen. Zunächst wird der Fotolack aufgeschleudert und strukturiert. Das sich ergebende Profil ist in der Abbildung 6.4 jeweils als Kurve (1) dargestellt. Um den Lack zur Maskierung gegen BHF zu verwenden, muss er zusätzlich nach dem Entwickeln ausgebacken werden. Das kann auf der Hotplate oder im Umluftofen geschehen. Währenddessen verliert der Fotolack Wasser und die Polymere verketteten sich noch fester. Dabei stabilisieren sich die Strukturen sowohl mechanisch als auch chemisch. Auch die Haftung der Maskierung auf dem Wafer wird dadurch erhöht. Der resultierende Volumenschwund ist an der Kurve (2) zu erkennen. Anschließend wird die Ätzung mit der gepufferten Flusssäure durchgeführt. Die Kurve (3) gibt das sich ergebende Profil nach dem Ätzen wieder. Schließlich ist das Profil nach dem Entfernen des Fotolacks in Kurve (4) aufgenommen. Aufgrund der Auflösung in x-Richtung ist die Unterätzung nicht zu erkennen.

Da es sich bei dem Ätzen mit gepufferter Flusssäure um einen isotropen Prozess handelt, liegt die Unterätzung in der Größe der geätzten Tiefe. Gewöhnlich verbleibt der zu ätzende Wafer noch länger als nötig in der Lösung, um sicher zu sein, dass alle Strukturen vollständig freigelegt sind. Deshalb ist sogar von einer größeren Unterätzung als der Äztiefe auszugehen. Für das hier untersuchte thermische Oxid bedeutet das eine Unterätzung größer als 300 nm. Das fand auch bei dem Layout

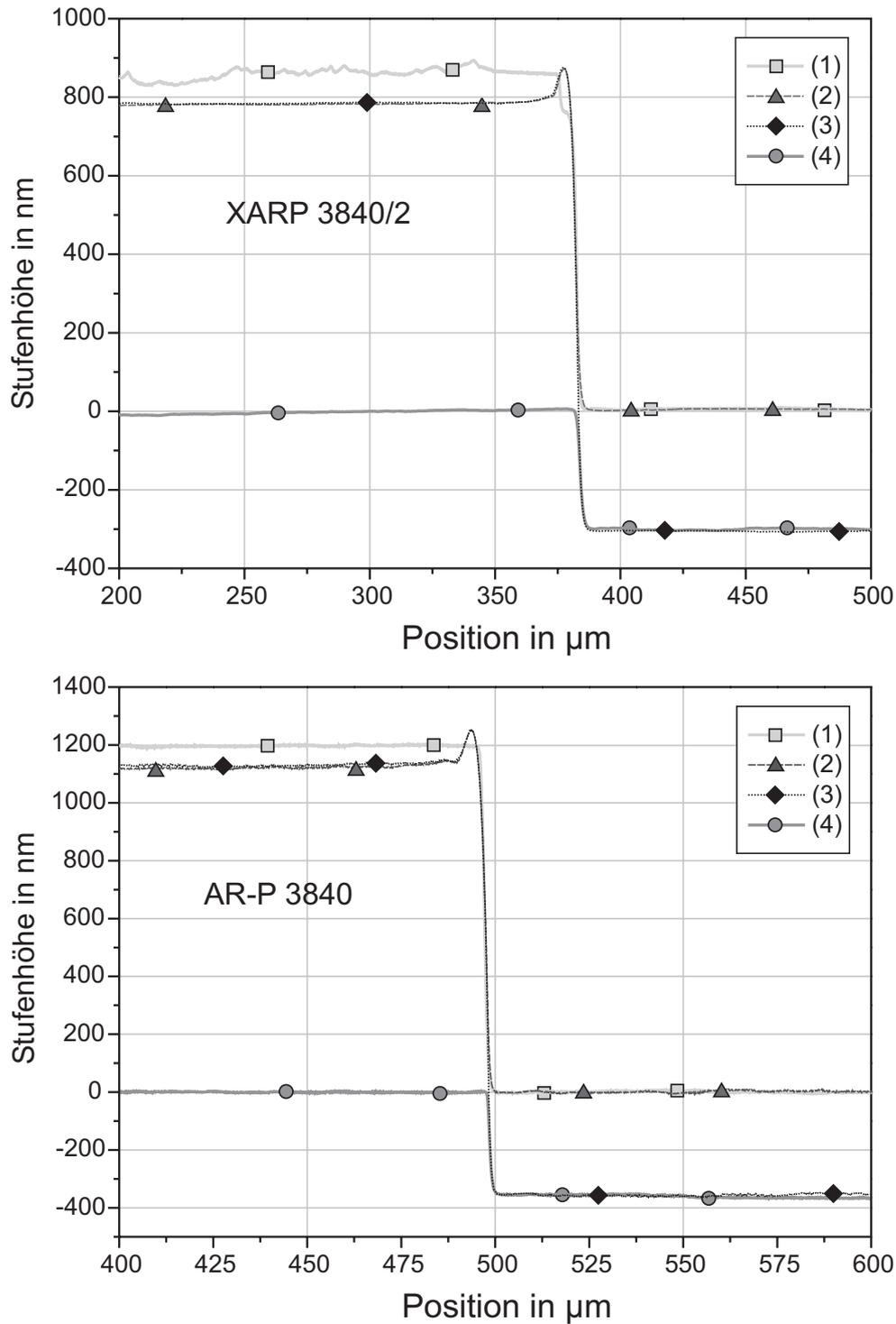


Abbildung 6.4: Profilometermessungen einer Stufe aus Siliziumdioxid während der verschiedenen Stadien der Strukturierung mit den Fotolacken XAR-P 3840/2 bzw. AR-P 3840: (1) nach der Strukturierung des Fotolacks; (2) nach dem Tempern zur chemischen Stabilisierung des Fotolacks, (3) nach der BHF-Ätzung; (4) nach dem Entfernen des Fotolacks.

der Masken Berücksichtigung, indem die Geometrien der Maske entsprechend größer ausgelegt wurden und so die Ätzmaskierung genügend Vorhalt hat.

Anhand der Deckungsgleichheit der Kurven (2) und (3) in den Abbildungen 6.4 zeigt sich, dass der Abtrag der so behandelten Fotolacke vernachlässigbar ist. Beide Lacksysteme eignen sich also für das Strukturieren der Oxidschicht mittels BHF. Aufgrund der geringeren Anschaffungskosten und der einfacheren Strukturierbarkeit des Fotolacks AR-P 3840 über schon topologisch vorgeformte Substrate (siehe dazu 6.4.3) ist dieser im Allgemeinen als Ätzmaske für BHF dem anderen System vorzuziehen.

Abschließend lässt sich aber sagen, dass beide Fotolacke XAR-P 3840/2 sowie AR-P 3840 der Fa. Allresist für diesen Prozessschritt geeignet sind.

6.4 Polystrukturierung

Zur Strukturierung des Polysiliziums wird ein Prozess benötigt, der es ausreichend selektiv gegen die Strukturierungsmaske und Siliziumdioxid ätzt. Zusätzlich ist hier die Strukturierung über eine Mesa, deren Höhe wesentlich größer als die Schichtdicke des Polysiliziums ist, zu berücksichtigen. Vor allem Letzteres bedeutet, dass hier ein isotroper Ätzprozess verwendet werden muss, was kurz erläutert werden soll.

In einem anisotropen Verfahren erfolgt der Materialabtrag definitionsgemäß vorwiegend in eine Raumrichtung, im Falle der üblicherweise verwendeten Plasmaätzer am schnellsten senkrecht zur Waferfläche. Das heißt für topologisch strukturierte Wafer, dass Schichten auf Flächen parallel zur Schnittfläche des Wafers schneller geätzt werden als auf dazu senkrechten Flächen.⁸ Abbildung 6.5a illustriert einen anisotropen Ätzprozess, wobei die gestrichelte Linie den Schnitt der abzutragenden Schicht vor dem Ätzen verkörpert. Bei weiterem Ätzen wird dann diese Schicht nur noch auf der vertikalen Flanke stehen bleiben. Noch weiteres Fortschreiten des Prozesses führt dann auch dort zum Abtrag der Schicht.

Bei einem isotropen Ätzprozess wird das Material gleichmäßig in alle Richtungen abgetragen. Deshalb wird die Maskierung oft stark unterätzt. Abbildung 6.5b zeigt ein charakteristisches Ätzprofil eines isotropen Verfahrens. Die Struktur der Maske kann nicht maßhaltig in die darunter liegende Schicht übertragen werden. Die Seitenflanken der geätzten Schicht sind nicht eben und senkrecht.

Die Polysiliziumschicht befindet sich auf einem 16 nm dünnen Gateoxid. Um dieses nicht durch zu ätzen, muss das Ätzverfahren eine ausreichende Selektivität haben. Damit auch an den Mesaflanken das Polysilizium zuverlässig entfernt wird, darf die Anisotropie des Ätzprozesses nicht zu hoch sein. Das soll kurz gezeigt werden.

⁸Das wird bei der sog. Spacer-Technik ausgenutzt, um Abstandsstücke oder Strukturen im sub- μm -Bereich herzustellen, ohne sie lithografisch erzeugen zu müssen [WMF88].

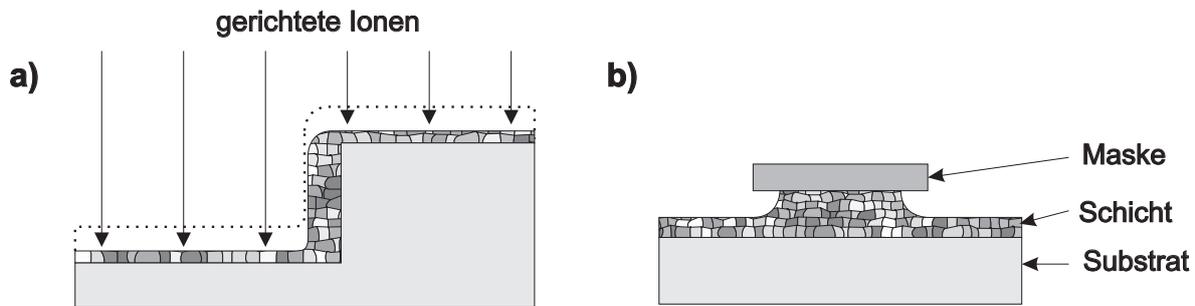


Abbildung 6.5: Typische Ätzprofile: a) anisotropes (physikalisches) Ätzen einer Stufe durch Ionenbeschuss, b) isotropes nasschemisches Ätzen.

Man definiert allgemein die Ätzrate

$$r = \frac{\Delta h}{\Delta t} = \frac{\text{Ätzabtrag}}{\text{Ätzzeit}} \quad (6.5)$$

und die Selektivität

$$S = \frac{r_1}{r_2} = \frac{\text{Ätzrate von Material 1}}{\text{Ätzrate von Material 2}} \quad (6.6)$$

Als Maß für die Anisotropie definiert man den Anisotropiefaktor

$$f = 1 - \frac{r_h}{r_v} = 1 - \frac{\text{laterale Ätzrate}}{\text{vertikale Ätzrate}} \quad (6.7)$$

Wird in laterale und vertikale Richtung gleich schnell geätzt, ergibt sich ein Anisotropiefaktor f gleich null. Für einen rein anisotropen Ätzprozess hat dieser den Maximalwert eins.

Abbildung 6.6a zeigt die Skizze eines Schichtstapels über einer Stufe. Die obere Schicht soll mit einem rein anisotropen Prozess, ohne horizontale Komponente vollständig entfernt werden. Der Anisotropiefaktor ist also 1. Die darunter liegende Schicht soll nicht vollständig durchgeätzt werden. Damit das erfolgen kann, darf das Wegnehmen des zu ätzenden Materials 1 in einer Dicke, die der Stufenhöhe h plus der Schichtdicke d_1 entspricht, nicht länger dauern als das Ätzen der Schichtdicke d_1 des Materials 1 und das der Schichtdicke d_2 des Materials 2. Das drückt nachstehende Beziehung aus:

$$\frac{d_1 + h}{r_1} < \frac{d_1}{r_1} + \frac{d_2}{r_2} \quad (6.8)$$

Durch Umformen erhält man:

$$\frac{r_1}{r_2} > \frac{h}{d_2} \quad (6.9)$$

Durch Einsetzen der Gleichung (6.6) ergibt sich das Minimum, das die Selektivität haben muss, um den Ätzprozess wie beschrieben durchführen zu können:

$$S > \frac{h}{d_2} \quad \text{mit} \quad f = 1 \quad (6.10)$$

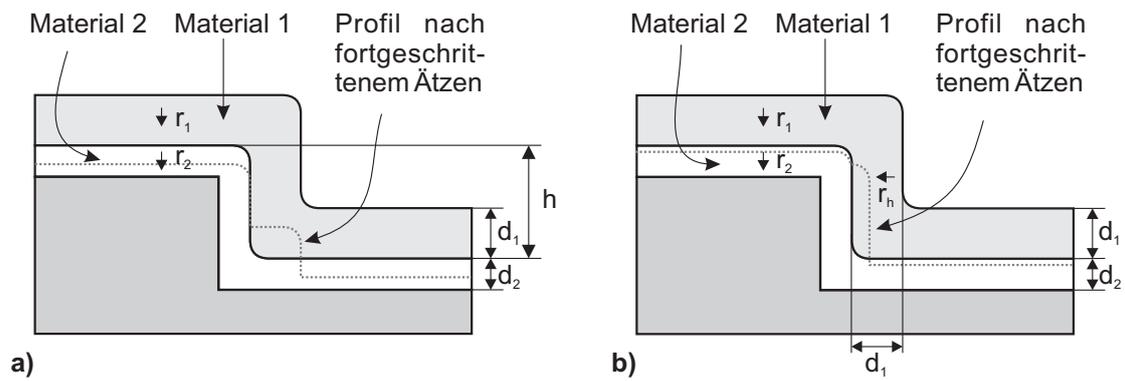


Abbildung 6.6: Profile beim Ätzen zweier Schichten über einer Stufe: a) rein anisotroper Prozess, b) mit horizontaler Komponente.

Ist Gleichung (6.10) erfüllt, wird das zu ätzende Material 1 auch an den Mesaflanken von oben her entfernt, noch bevor das Material 2 auf der Oberseite der Mesa durchgeätzt ist.

In dem konkreten Fall des zu fertigenden Sensors entspricht d_2 der Dicke des Gate-Oxids, welches sich unter dem zu ätzenden Polysilizium befindet. Setzt man die entsprechenden Werte des zu fertigenden Sensors ein (Stufenhöhe 2900 nm, Gate-Oxid 16 nm), erhält man die mindest erforderliche Selektivität von 181,25. Das ist auch mit guten anisotropen Plasmaprozessen nur schwer erreichbar. Es muss also eine gewisse laterale Ätzrate zugelassen werden.

In Abbildung 6.6b wird das Geschehen dargestellt, wenn eine horizontale Ätzkomponente existiert (f ist kleiner 1). Hier bleibt auf der vertikalen Flanke der Stufe ein Rest des Materials 1 übrig, nach dem auf den lateralen Flächen dieses bereits entfernt ist. Die Schichtdicke dieses Restes beträgt

$$d_1 - \frac{d_1}{r_1} \cdot r_h \quad (6.11)$$

und ist gegenüber der ursprünglichen d_1 um den Teil geringer, den die horizontale Ätzrate r_h in der Zeit ätzt, in der senkrecht mit der Ätzrate r_1 die Schichtdicke d_1 abgetragen wird. Dieser verbleibende Rest muss wiederum schneller entfernt werden als die Schicht des Materials 2. Deshalb kann man die Bedingung aufstellen:

$$d_1 \left(1 - \frac{r_h}{r_1}\right) \frac{1}{r_h} < \frac{d_2}{r_2} \quad (6.12)$$

Mit Einsetzen des Anisotropiefaktors 6.7 in Gleichung (6.12) erhält man

$$d_1 f \frac{1}{r_1(1-f)} < \frac{d_2}{r_2} \quad (6.13)$$

was man zu

$$\frac{d_1}{d_2} \cdot \frac{f}{1-f} < \frac{r_1}{r_2} \quad (6.14)$$

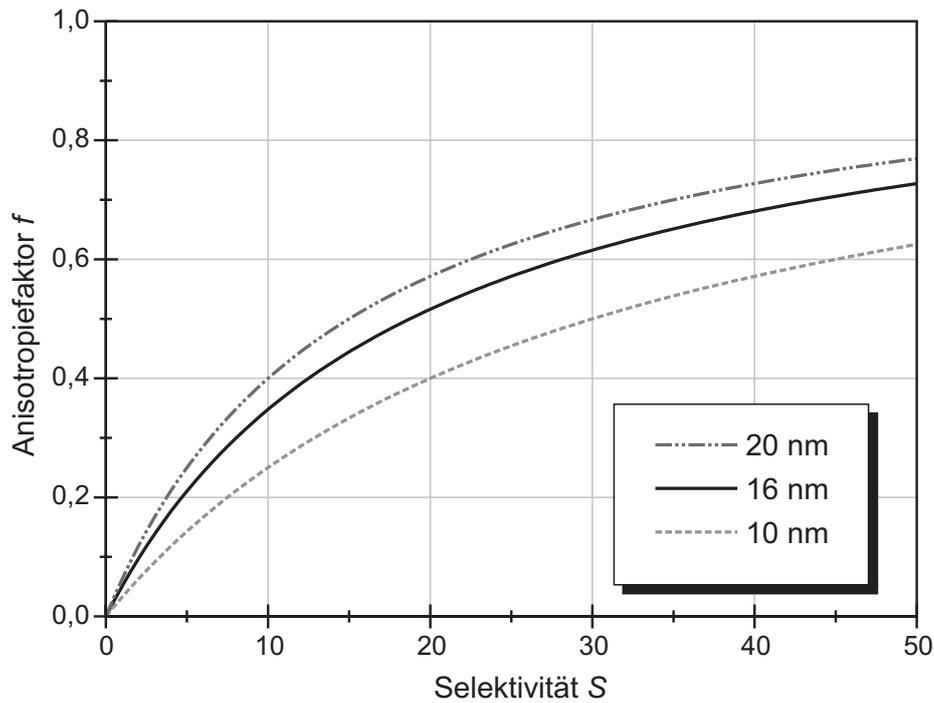


Abbildung 6.7: Maximaler Anisotropiefaktor in Abhängigkeit der Selektivität für das Ätzen eines Schichtstapels aus 300 nm Polysilizium und verschiedener Gate-Oxiddicken

umformen kann. Nach dem Einsetzen von Gleichung (6.6) in (6.14) und Auflösen erhält man die Bedingung:

$$f < \frac{S}{\frac{d_1}{d_2} + S} \quad \text{mit} \quad S < \frac{h}{d} \quad . \quad (6.15)$$

Ist die Bedingung (6.15) erfüllt, wird das Material 1 vollständig entfernt, noch bevor das Material 2 durchgeätzt wird. Vergrößert man den Anisotropiefaktor, verringert sich die horizontale Ättrate r_h im Verhältnis zur Vertikalen r_v . Steigert man f soweit, dass die Bedingung (6.15) nicht mehr gilt, wird das Material 1 auf der Flanke nicht mehr schnell genug geätzt.

Im Falle des Sensors entspricht d_1 der Dicke des Polysiliziumfilms und d_2 der des Gate-Oxids. Abbildung 6.7 zeigt den maximalen Anisotropiefaktor, der zum Strukturieren der Polysiliziumschicht des in dieser Arbeit behandelten Sensors, möglich ist. Dabei wurde vorausgesetzt, dass das Polysilizium 300 nm dick ist. Die angenommenen Gate-Oxiddicken sind in dem Grafen angegeben. Für den Fall, dass die Selektivität null ist, darf der Anisotropiefaktor nicht größer als null sein, damit das gesamte Polysilizium abgetragen werden kann. Das entspricht einem isotropen Prozess. Je höher die Selektivität S ist, desto größer kann der Anisotropiefaktor f sein. Ab einer Selektivität S von 181,25, gilt wieder die obige Bedingung (6.6). Aus dem Graphen 6.7 ist auch ablesbar, dass selbst bei einer für RIE-Prozesse sehr

hohen Selektivität von 20 der Anisotropiefaktor nicht größer als ungefähr 0,5 sein darf.

Die obigen Ausführungen zeigen, dass hier ein isotroper Ätzprozess notwendig ist. Da die Strukturen der Polysiliziumlage in der Anwendung des Sensors in lateraler Ausdehnung wesentlich breiter sind als die Schichtdicke, kann auch ein starkes Unterätzen in Kauf genommen werden. Deshalb werden im Folgenden Möglichkeiten isotroper Ätzung untersucht.

Die meisten nasschemischen Prozesse sind (bis auf wenige Ausnahmen) streng isotrop. Die Vorzüge dieser Technik liegen in der hohen Selektivität, der geringen Kontamination und Schädigung der Oberfläche, der einfachen Handhabung sowie der Gleichmäßigkeit und Reproduzierbarkeit der Prozessführung.

Die meisten in der Literatur vorgeschlagenen Mischungen für das isotrope Ätzen von Polysilizium sind im Prinzip auf zwei ähnliche Lösungen zurückzuführen. Zum einen ist dies eine Mischung aus Salpeter-⁹ und Fluorsäure¹⁰ mit Essigsäure¹¹ oder Wasser. Deren Ätzraten in Abhängigkeit der Konzentrationen und deren chemische Abläufe werden ausführlich in [RS60] diskutiert. Zum anderen findet auch eine aus Salpetersäure, Wasser und Ammoniumfluorid¹² bestehende Ätzlösung Verwendung. Außer diesen können prinzipiell auch Lösungen verwendet werden, um Polysilizium zu entfernen, die sonst eher als anisotrope Ätzen für kristallines Silizium bekannt sind. Die Anisotropie dieser Lösungen, wie z. B. TMAH, rührt von den unterschiedlichen Bindungsenergien her, die die Siliziumatome in unterschiedlichen Kristallrichtungen in einkristallinem Silizium haben. Da aber Polysilizium keine einheitliche Kristallrichtungen besitzt, tritt der anisotrope Effekt nicht auf. In der Tabelle 6.3 sind Ansätze, wie sie zum Ätzen von Silizium bzw. Polysilizium üblicherweise verwendet werden, aufgeführt.

Aus der Tabelle 6.3 kann man entnehmen, dass sich für einen Prozess, bei dem eine hohe Selektivität gegen Siliziumdioxid erforderlich ist, TMAH mit Selektivitäten größer als 1000 am besten eignet. Als Maske könnte dazu eine Hardmask aus Siliziumdioxid dienen. Für eine einfache Prozessführung ist es jedoch wünschenswert, Fotolack als Maske zu verwenden. Es wurden deshalb verschiedene Fotolacksysteme auf die Eignung zur Strukturierung getestet.

6.4.1 AR-U 4030

Der Fotolack AR-U 4030 der Firma Allresist hielt, als Positiv-Resist betrieben, keiner der in der Tabelle 6.3 aufgeführten Ätzlösungen stand. Auch ein Ausbacken nach dem Entwickeln brachten keine Verbesserung der Ätzbeständigkeit. Daher wurde dieser Fotolack nicht weiter für das Strukturieren des Polysiliziums betrachtet.

⁹HNO₃

¹⁰HF

¹¹CH₃COOH

¹²NH₄F

Lösungsansatz	Temperatur	Ätzzraten in nm ³ /min für		Selektivität	Bemerkung
		Si	Poly SiO ₂		
1 Teil HF 3 Teile HNO ₃ 8 Teile CH ₃ COOH	20 °C	3000	700-3000	≈100	»1-3-8-etch«; Ätzrate sinkt auf $\frac{1}{100}$ für $n_n < 3 \cdot 10^{17} \text{ cm}^{-3}$ oder $n_p < 3 \cdot 10^{15} \text{ cm}^{-3}$
1 Teil HF 2 Teile HNO ₃ 1 Teil CH ₃ COOH	20 °C	40000			»CP-4A«; Ätzrate $< \frac{1}{100}$ für $n_n < 3 \cdot 10^{17} \text{ cm}^{-3}$ oder $n_p < 3 \cdot 10^{15} \text{ cm}^{-3}$
TMAH	80 °C	800		>1000	Ätzrate sinkt auf $\frac{1}{10}$ für $n_{Bot} > 3 \cdot 10^{20} \text{ cm}^{-3}$
2 Teile HF 15 Teile HNO ₃ 5 Teile CH ₃ COOH	20 °C	5000			»Planar-Ätze« r_{Si} von der Dotierung abhängig
126 Teile HNO ₃ 60 Teile H ₂ O 5 Teile NH ₄ F	20 °C	300	100 - 340	30-100	
50 Teile HNO ₃ 20 Teile H ₂ O 1 Teil HF	20 °C	2500	1000-6000	8,7	

Tabelle 6.3: Ätzzraten verschiedener Poly-Silizium-Ätzzlösungen

Es ist trotzdem festzuhalten, dass sich dieser Fotolack sehr leicht über Mesen und Stufen strukturieren lässt. Außerdem ist dieser gegen BHF beständig, sodass die oben angesprochene Möglichkeit besteht, mit Hilfe dieses Lackes eine Hardmask aus Siliziumdioxid herzustellen, die darauf zur Strukturübertragung mit TMAH genutzt wird.

6.4.2 XAR-P 3840/2

Eine hohe Ätzbeständigkeit zeigt der Fotolack XAR-P 3840/2. Mit Ausnahme von TMAH kann dieser Lack für alle Lösungen, die in Tabelle 6.3 angegeben sind, als Fotolackmaske verwendet werden. Vorausgesetzt er wird nach der Entwicklung ausreichend ausgebacken. Dabei spielt weniger die Ausbackzeit als die -temperatur, die zwischen 110 °C und 180 °C liegen kann, eine Rolle. Je höher die Temperatur ist, je stärker verfestigen sich die Polymere des Fotolacks. Allerdings lässt sich der Lack nach einem Ausbacken mit Temperaturen größer als 140 °C nur sehr schwer wieder entfernen.

Um die Tauglichkeit des Fotolackes XAR-P 3840/2 als Maske für das Ätzen des Poly-Gates zu untersuchen, wurde dieser an Mesastufen getestet. Dazu wurden 500 nm und 3 µm hohe Mesen aus Silizium gefertigt. Anschließend wurde der Fotolack XAR-P 3840/2 aufgeschleudert und strukturiert. Trotz der Variation aller Parameter für das Aufschleudern, Ausbacken, Belichten sowie das Entwickeln wurde bei den 3 µm tiefen Stufen kein Erfolg erzielt. Auf Abbildung 6.8 sieht man ein Ergebnis dieser Versuche. Es zeigt entwickelte Lackstrukturen über einer Mesa aus Silizium. Der Fotolack (XAR-P 3840/2) ist auf der oberen Kante, der konvexen Ecke, schon soweit entwickelt oder überentwickelt, dass diese Kante freigelegt ist. Deutlich ist die Unterbrechung der Maske auf Abbildung 6.8 zu sehen. Da die Fotolackmaske aber Polysilizium als Gate-Material abdecken soll, griffe an genau dieser Stelle die Ätzlösung die Leiterbahn an und führte zu einer Leiterbahnunterbrechung. Auf der anderen Seite wiederum erkennt man auf Abbildung 6.8, dass der Fotolack aus der unteren konkaven Ecke noch nicht herausgelöst ist.

Diese Schwierigkeiten beim Strukturieren über der Mesa ergeben sich aus den unterschiedlichen Schichtdicken des Fotolackes über topologisch vorstrukturierte Substrate. Abbildung 6.9 skizziert ein resultierendes Profil der Lackoberfläche nach dem Aufschleudern und nach dem Ausbacken. Beim Aufschleudern wird der Lack vom Mittelpunkt des Wafers nach außen geschleudert. Dabei bleibt der Lack an den vertikalen Strukturen hängen und die Mesen sind an den Flanken zur Mitte des Wafers hin stärker belackt als an den ihnen gegenüberliegenden Flanken.

Weil der Lack bestrebt ist, seine Oberflächenenergie zu verringern, versucht er, eine Form anzunehmen, sodass die Krümmungsradien an der Oberfläche möglichst groß und gleich werden. Das ist nur möglich, solange sich noch genügend Lösungsmittel im Lack befindet. Durch die Erwärmung des Fotolackes beim anschließenden Ausbacken fängt dieser wieder an zu fließen. Da er weiter versucht, die Oberflächenenergie zu minimieren, werden wiederum die Krümmungsradien

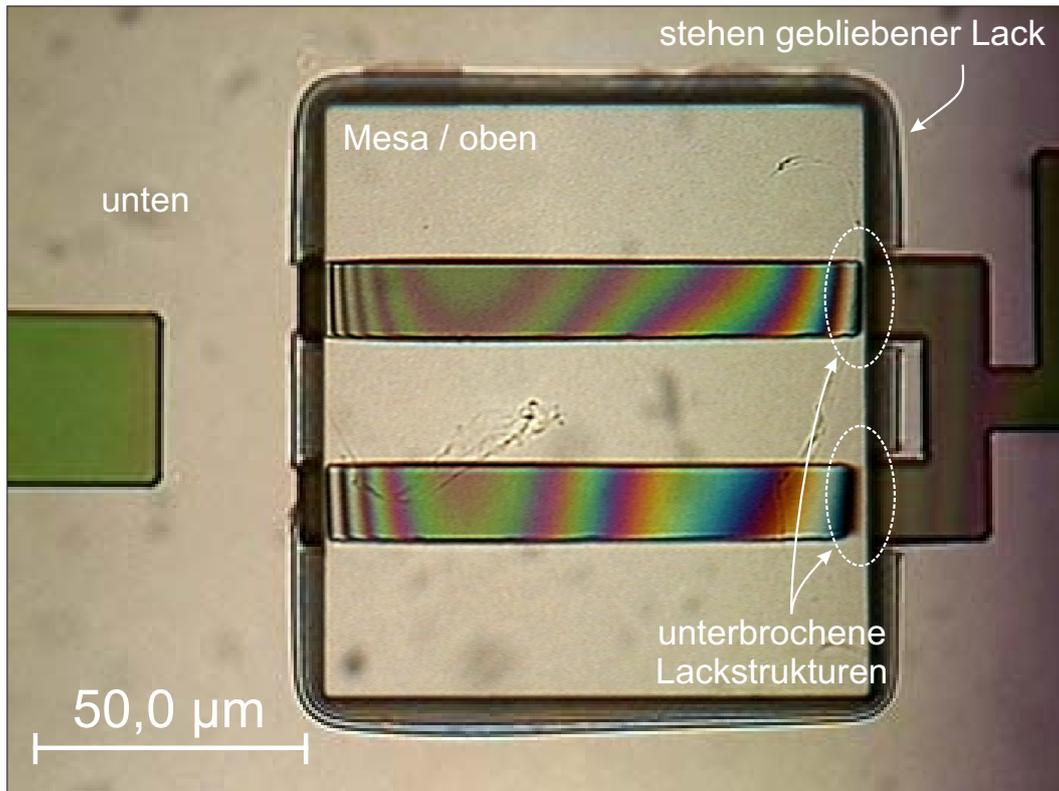


Abbildung 6.8: Fotolack XAR-P 3840/2 strukturiert über eine Mesa (Draufsicht)

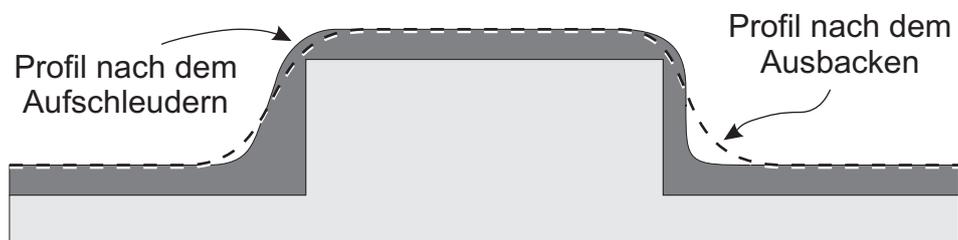


Abbildung 6.9: Querschnitt einer über einer Mesa aufgetragenen Fotolackschicht

immer größer. Deshalb fließt das Material in die konkaven Ecken hinein und zieht sich von den konvexen Ecken zurück.

Die Zeit des Entwickelns nach der Belichtung muss nun so hoch gewählt werden, dass der belichtete Fotolack in den konkaven Ecken vollständig gelöst wird. Dort ist der Lack relativ am dicksten und braucht am längsten, um fertig entwickelt zu werden. Während der Entwicklung wird auch unbelichteter Fotolack geätzt, nur nicht so schnell. Weil die Schichtdicke auf der Oberseite der Mesa nahe den konvexen Kanten nun aber am dünnsten ist, wird an diesen Stellen der Fotolack zuerst vollständig entfernt. Die Entwicklung darf den Zeitpunkt nicht überschreiten, an dem die relativ dünnere Lackschicht, ob belichtet oder nicht, auf der Oberseite der Mesa weggeätzt ist.

Um diesem Phänomen zu begegnen, gibt es mehrere Möglichkeiten. Um die Schichtdickenabnahme auf den Kanten zu verhindern, kann man auf das Ausbacken verzichten oder zumindest höhere Temperaturen vermeiden und den Fotolack z. B. nur bei 60 °C ausbacken. Für die Polystrukturierung scheidet diese Möglichkeit allerdings aus, weil der Fotolack nach dem Entwickeln, wie oben erwähnt, zwingend mit Temperaturen höher als 110 °C ausgebacken werden muss, um der Polyätzlösung standzuhalten.

Fotolack absorbiert zu einem gewissen Maße das einfallende Licht. Daher müssen stärkere Schichten auch länger belichtet werden. Durch eine Überbelichtung wird die Ätzrate des belichteten Fotolacks erhöht, und so wird die Entwicklungszeit gesenkt. Allerdings führt eine Überbelichtung von Positivlacken auch immer zu einer Strukturverkleinerung relativ zur Maske. Gerade für die Gate-Strukturierung ist es notwendig, die Toleranzen und auch die Überlappung des Gates über die Mesa so klein wie möglich zu halten. Durch einen Vorhalt in der Maske wäre das zu kompensieren. Das setzt aber die genaue Kenntnis des Prozessschrittes sowie der Mesahöhe und des verwendeten Fotolacks schon vor dem Masken-Design voraus.

Ein anderer Weg ist das doppelte Belichten und Entwickeln. Dabei wird der Fotolack direkt im Anschluss nach dem Entwickeln noch einmal belichtet. Wodurch bislang nicht vollständig belichtete Regionen, v. a. in konkaven Ecken, eine genügend hohe Lichtdosis erhalten, um vollständig vom Entwickler gelöst werden zu können. Bei dieser Technik ist die Strukturübertragung von Maske auf Fotolack durch die Justiergenauigkeit begrenzt. Die Genauigkeit der Linienbreite ist in diesem Fall abhängig von dem Versatz der zwei Belichtungen zueinander. Strukturgrößen des entwickelten Fotolacks verringern sich um den doppelten Versatz der zweiten Belichtung zur Ersten. Wie schon vorher steht ein Gewinn an besserer Kantenbedeckung einem Verlust an Genauigkeit der Strukturübertragung gegenüber.

Alle diese Maßnahmen führten zu keinem brauchbaren reproduzierbaren Ergebnis. Spätestens während des zur chemischen Stabilisierung notwendigen Ausbackens bei 110 °C verfließt der Lack und zieht sich von den Kanten zurück. Zum Strukturieren über 2 bis 3 µm hohe Mesen ist der Fotolack ARP-3840/2 nicht geeignet.

6.4.3 AR-P 3840

Der Fotolack AR-P 3840 ist beständig gegen die Polyätze aus 126 Teilen Salpetersäure, 60 Teilen Wasser und einem Teil Ammoniumfluorid (siehe Tabelle 6.3). Deshalb wurde untersucht, inwieweit sich dieser Lack über die Mesen und Stufen strukturieren lässt.

Auch dieser Lack zeigt ein typisches Profil nach dem Aufschleudern und Ausbacken, wie es in Abbildung 6.9 dargestellt ist. Es kommt also genauso zur Verdünnung der Lackschicht oberhalb von Stufen und zu Verdickung unterhalb. Auch hier darf der Wafer dem Entwickler nicht so lange ausgesetzt werden, bis der Lack an den dünnsten Stellen (in der Nähe der Kanten, auf der Oberseite der Mesen) komplett entfernt ist. Dieser Zeitpunkt ist wiederum von der Umdrehungszahl beim Aufschleudern und damit der Schichtdicke, sowie der Entwicklerkonzentration abhängig. Wie aus Abbildung 6.9 ersichtlich, ist der Weg, den das Licht durch den Lack zurücklegt, unterhalb der Stufe wesentlich länger als oberhalb dieser. Da die Intensität des Lichtes exponentiell mit der Eindringtiefe abnimmt, wird der Lack in den tiefer liegenden Bereichen weniger stark belichtet. So weisen eben diese Bereiche eine geringere Ätzrate beim Entwickeln auf. Ein Überbelichten wirkt dem entgegen und sorgt für eine vollständige Belichtung über die gesamte Tiefe und erhöht so die Ätzrate der belichteten Gebiete.

Wie oben schon angesprochen, ist auch hier ein Kompromiss zu finden, mit dem die Überbelichtung ausreicht, den Fotolack in den konkaven Ecken vollständig zu belichten, aber der die Strukturen trotzdem nicht zu stark verkleinert. Auf den konvexen Ecken muss so viel Fotolack stehen bleiben, dass dieser beim anschließenden Ausbacken nicht wegfließt oder reißt.

Für bis zu etwa $3,5\ \mu\text{m}$ hohen Mesen lässt sich ein solcher Kompromiss finden. Die Strukturen werden beim Aufschleudern mit 4000 Umdrehungen pro Minute und einer daraus resultierenden Schichtdicke von $1,4\ \mu\text{m}$ genügend bedeckt. Dabei darf die Entwicklungszeit nicht länger als 50 Sekunden gewählt werden. Die Belichtungsdosis sollte $105\ \text{mW}/\text{cm}^2$ nicht überschreiten. Für Mesen, die bis zu etwa $5\ \mu\text{m}$ hoch sind, muss mit der schon erwähnten doppelten Belichtung gearbeitet werden. Das Aufschleudern erfolgt mit 1500 Umdrehungen pro Minute. Das ergibt eine Schichtdicke von ungefähr $3\ \mu\text{m}$. Den dabei entstehenden Lackwulst am Rand des Wafers muss man in Kauf nehmen. Die Lichtdosis von $235\ \text{mW}/\text{cm}^2$ wird auf beide Belichtungen aufgeteilt. Nach jeder Belichtung erfolgt eine Entwicklung. Deren Gesamtzeit sollte 95 s nicht erreichen. Hier gilt wieder, dass die genauen Strukturmaße von dem Versatz der beiden Belichtungen zueinander bestimmt werden.

Damit die so erzeugte Fotolackstrukturen als Maske für das Polyätzen verwenden werden können, muss der Lack thermisch behandelt werden. Ohne diesen Schritt ist die Haftung nicht ausreichend und der Lack bekommt Risse. Durch höhere Temperaturen (ab etwa $130\ ^\circ\text{C}$) fängt der Fotolack jedoch an zu fließen und so wird die Maske v. a. an den Ecken der Mesastufen unbrauchbar.

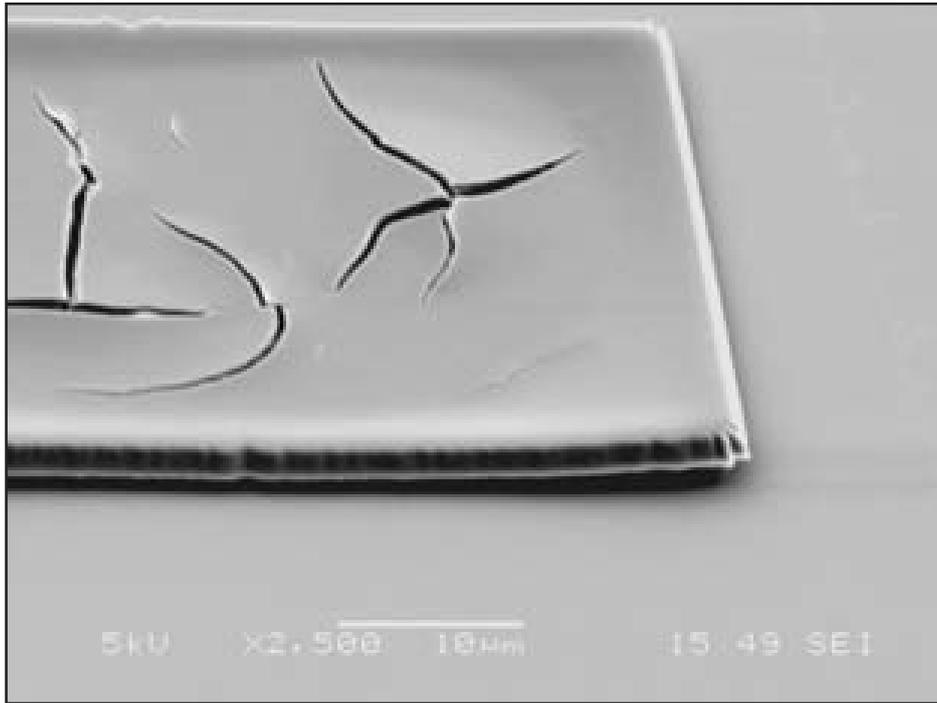


Abbildung 6.10: Durch den Postexposure-Anneal (120 °C) thermisch stabilisierter Fotolack AR-P 3840 nach dem Ätzen mit Poly-Ätze

Es wurde versucht, den Lack noch vor dem Entwickeln durch einen Postexposure-Anneal zu stabilisieren. Da zu diesem Zeitpunkt der Erwärmung der gesamte Wafer noch mit dem Lack bedeckt ist, wird ein schnelles Wegfließen von den Kanten verhindert. Mit einem Postexposure-Anneal bis zu 120 °C konnte die Haftung allerdings nicht ausreichend verbessert werden, sodass sich die Maske während des Ätzens löste. Durch diesen Postexposure-Anneal trocknet er nicht genügend aus. Während des Ätzens kommt es zu Verspannungen innerhalb der Lackschicht. Deswegen wölbt sie sich am Rand der Strukturen nach oben und löst sich vom Substrat, wie es in Abbildung 6.10 zu sehen ist. Es kommt zu Unterätzungen, die im Falle einer nur 300 nm starken Polysiliziumschicht ein bis zwei Mikrometer betragen können. Eine maßhaltige Strukturübertragung ist nicht mehr gegeben.

Brauchbare Ergebnisse konnten nur mit einem Ausbacken nach dem Entwickeln erzielt werden. Entscheiden ist dabei weniger die Dauer des Ausbackens als die Temperatur. Wird der Fotolack bei 110 °C ausgebacken, bekommt er beim Ätzen des Polysiliziums Risse. Das ist in Abbildung 6.11 zu sehen. Diese Risse bilden sich bevorzugt dort, wo Fotolack über Kanten hinweg strukturiert ist, oder in sehr großen Strukturen. Durch sie kann die Ätzlösung eindringen, was zu ungewolltem Abtrag des Polysiliziums führt. Erst bei einer Ausbacktemperatur von 120 °C vernetzten die Polymere des Fotolackes stark genug. Abbildung 6.12 zeigt den Fotolack AR-P 3840 nach dem Ätzen. Die Lackmaske zeigt noch keine Risse, genügt also den Ansprüchen der Strukturierung des Polysiliziums, wie sie in dieser

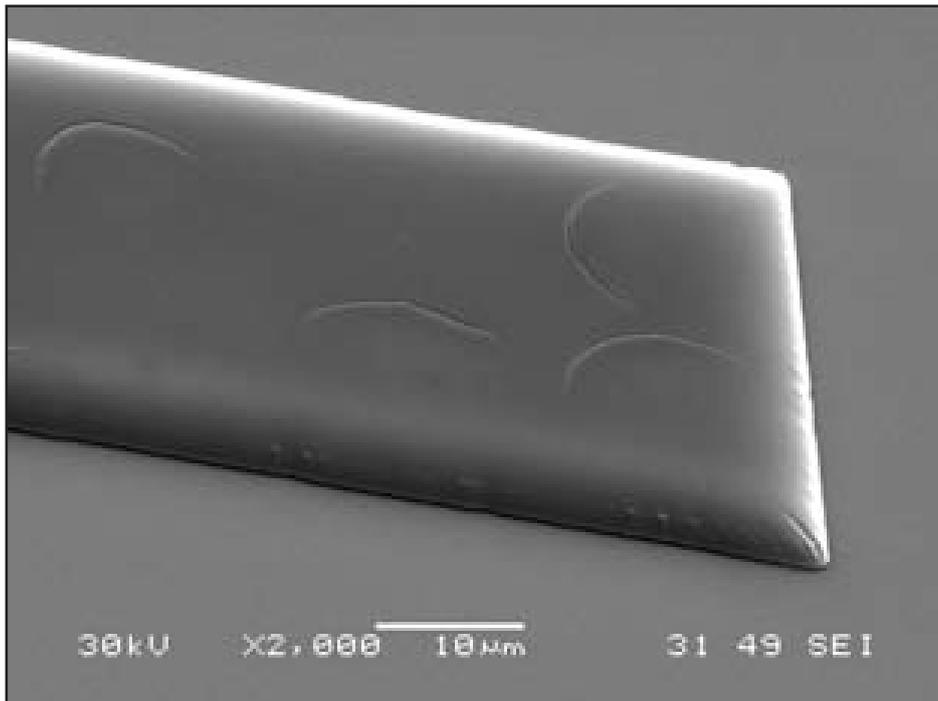


Abbildung 6.11:
AR-P 3840 nach dem Entwickeln mit 110 °C thermisch stabilisiert; nach dem nasschemischen Ätzen mit Poly-Ätze

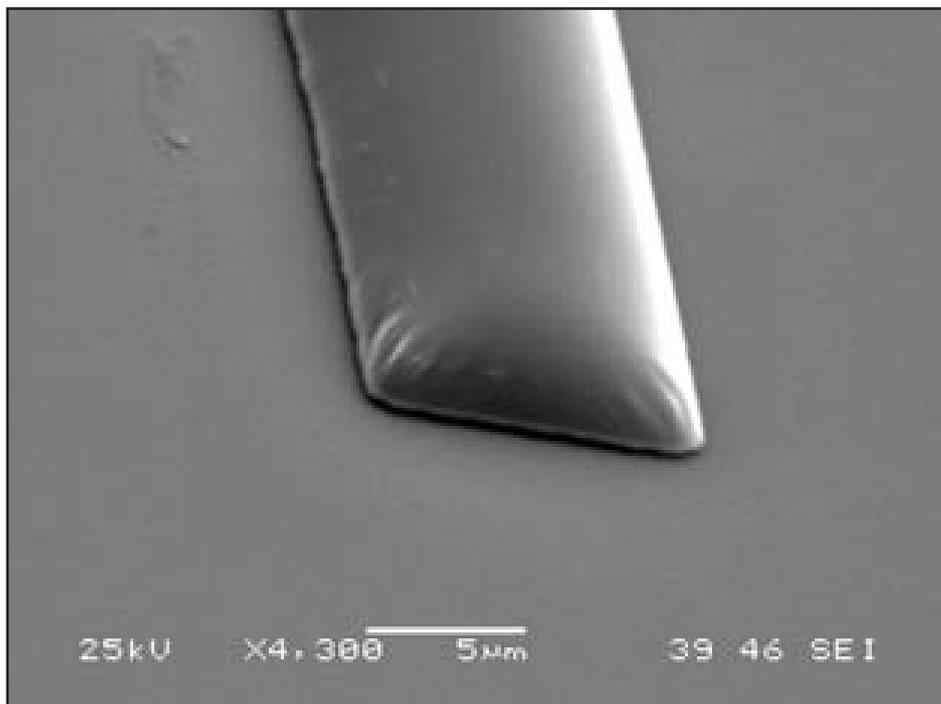


Abbildung 6.12:
AR-P 3840 nach dem Entwickeln mit 120 °C thermisch stabilisiert; nach dem nasschemischen Ätzen mit Poly-Ätze

Fotolack	Ätzbeständigkeit	Kantenbedeckung
AR-U 4030	-	++
XAR-P 3840/2	++	-
AR-P 3840	+	+
++ sehr geeignet	+ geeignet	- unbrauchbar

Tabelle 6.4: Vergleich der verschiedenen Fotolacke

Arbeit benötigt wird. Da die Ausbacktemperatur von 120 °C in der Nähe des Glaspunktes des Fotolackes liegt, wird der Wafer nicht sofort auf diese Temperatur gebracht, sondern davor für zwei Minuten auf der Hotplate bei 60 °C gehalten. So wird den Polymeren Zeit gegeben, sich schon vorher zu vernetzen und zu stabilisieren. Abbildung 6.14 zeigt eine Fotolackstruktur, die über eine 3 µm hohe Mesa strukturiert und anschließend, wie beschrieben bis zu 120 °C ausgebacken wurde. Es ist zu erkennen, dass der Lack, obwohl er sich zusammengezogen hat, die Kante trotzdem vollständig bedeckt.

In Tabelle 6.4 werden die Ergebnisse der Untersuchungen noch einmal zusammenfassend dargestellt. Der Fotolack AR-U 4030 eignet sich durchaus für das Strukturieren auf derart topologisch beschaffenen Substraten, wie sie hier behandelt wurden. Allerdings ist die Ätzbeständigkeit gegen die Polyätzlösungen völlig ungenügend. Andersherum verhält es sich beim XAR-P 3840/2, der chemisch äußerst resistent ist, aber sehr hohe Fließeigenschaften aufweist. Nur der Fotolack AR-P 3840 lieferte zufrieden stellende Ergebnisse. Mit ihm lassen sich die Polysiliziumstrukturen über 3 bis 3,5 µm hohe Mesen erzeugen.

Wie bereits erwähnt, eignet sich zusammen mit dem AR-P 3840 am besten der Ansatz: Salpetersäure, Wasser und Ammoniumfluorid im Verhältnis 126:60:1. Dieser Fotolack hält auch Ansätzen mit Zusätzen an Flusssäure stand, wenn man ihn mit genügend hohen Temperaturen behandelt (zwischen 140 °C und 150 °C). Der Fotolack wird beim Polyätzen am stärksten durch die Salpetersäure angegriffen. Das basische Ammoniumfluorid senkt den pH-Wert in der Lösung und schützt so den Fotolack. Trotzdem ist es bei allen dieser Lösungen eine Frage der Zeit, die der Lack der Ätze ausgesetzt wird. Irgendwann kommt es zu Verspannungen und schließlich reißt er. Große Strukturen (im Bereich mehrerer Millimeter) sind hier besonders gefährdet.

Das Verhältnis von 126:60:5 der Anteile an Salpetersäure, Wasser, und Ammoniumfluorid ist optimal. Abbildung 6.13 zeigt die Abhängigkeit der Ätzraten für Silizium und Siliziumdioxid vom Wassergehalt der Lösung. Ein Verdünnen des Ansatzes mit Wasser, um den Fotolack zu schützen oder die Ätzrate für Siliziumdioxid zu senken, um so das Überätzen, v. a. bei sehr dünnen Gateoxiden, weniger kritisch zu machen, bringen keinen Erfolg. Steigt der Anteil an Wasser über die

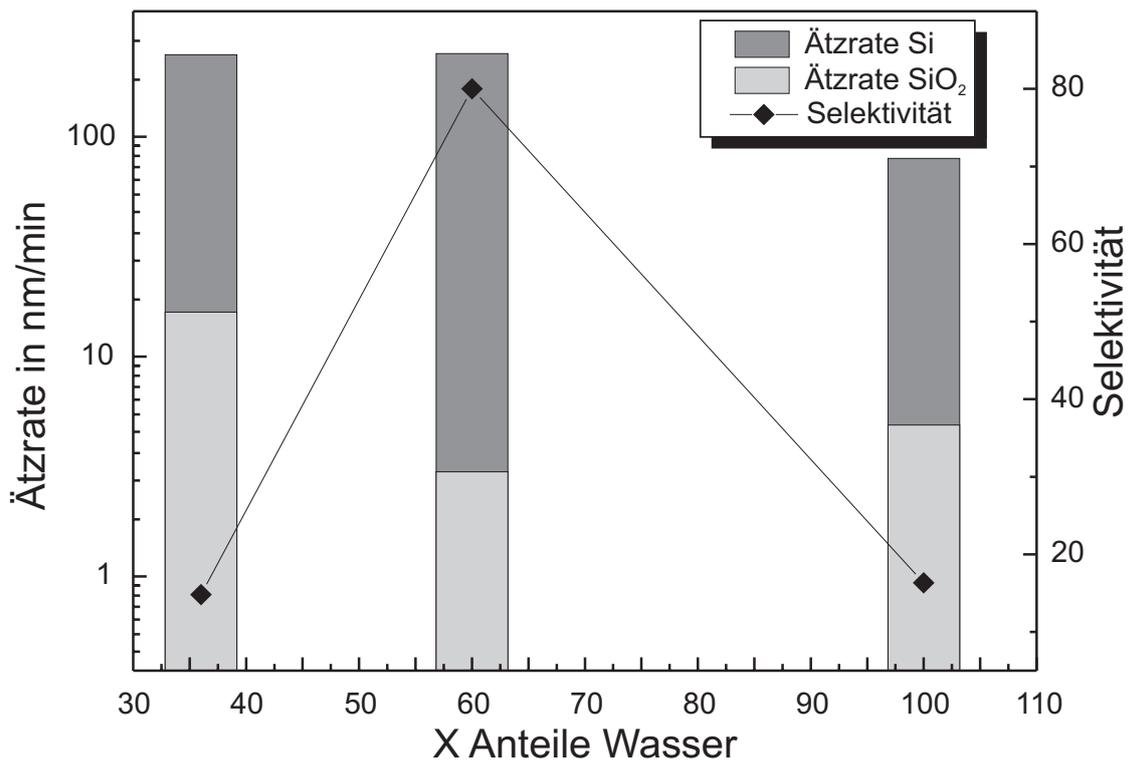


Abbildung 6.13: Ätzzraten von kristallinem Silizium und Siliziumdioxid in Poly-Ätze aus $\text{HNO}_3:\text{H}_2\text{O}:\text{NH}_4\text{F}$ in Abhängigkeit des Wasseranteils X in einem Mischungsverhältnis von $126:X:5$

oben genannte Konzentration, nimmt die Ätzzrate für Siliziumdioxid zu und die Selektivität sinkt.

Die Prozesskontrolle während des Ätzens im Becherglas gestaltet sich einfach. Das Polysilizium verändert seine Farbe mehrmals. Im Falle der hier verwendeten Schichten schlägt sie von rötlich Violett über Grün nach Rosa zu Blau um, und wird schließlich wieder rötlich Violett. Dieser Zyklus wird sogar wiederholt durchlaufen. Kurz vor Schluss verfärbt sie sich in ein gelbliches Weiß und am Ende sieht die Oberfläche des Wafers in der Lösung metallisch Grau, wie Silizium aus. Genau das ist der Zeitpunkt, an dem der Vorgang unterbrochen werden muss, weil ein weiteres Ätzen sonst das Gate-Oxid zerstören würde.

Die Ätzzraten des Polysiliziums hängen sehr stark von dessen Dotierung sowie seiner Korngröße ab. Auch das Einhalten der Temperatur ist wichtig. Prinzipiell steigt die Rate mit der Dotierung und der Temperatur. Beim Zusammenmischen der Lösung wird Energie freigesetzt und sie erwärmt sich. Um ein schnelleres Ätzen der zuerst eingetauchten Wafer zu verhindern, muss das Ätzbad wieder auf Raumtemperatur abgekühlt werden. Das geschieht z. B. im Wasserbad. Bei den Versuchen in dieser Arbeit lag die Ätzzrate bei Raumtemperatur im Bereich zwischen 240 bis $340 \text{ nm}/\text{min}$.

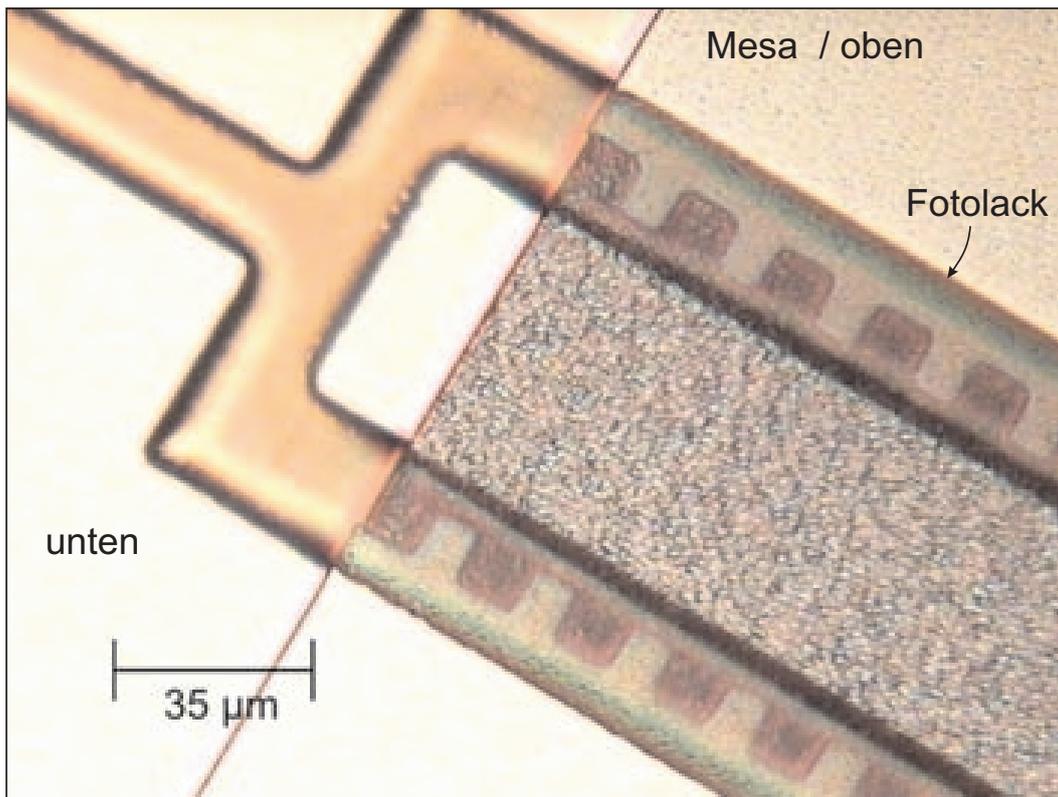


Abbildung 6.14: AR-P 3840 über eine 3 μm hohe Mesa strukturiert und bei 120 °C ausgebacken (Draufsicht)

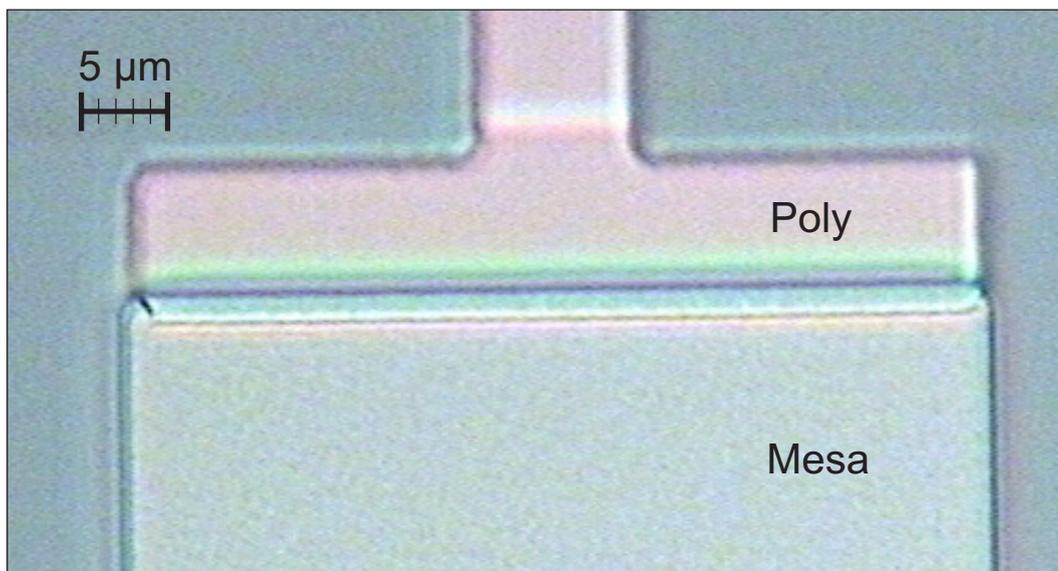


Abbildung 6.15: Geätzte Polysiliziumstruktur mit 1 μm Überlappung über der Mesa (Draufsicht)

Nun stellt sich die Frage, wie groß die Überlappung sein muss, damit reproduzierbar ein Polysilizium-Gate entlang einer Mesastufe konstruiert werden kann. Um die Überlappkapazitäten eines Transistors so klein wie möglich zu halten, sollte das Gate idealerweise nur über dem Kanal liegen. Das exakte Abschließen einer Struktur an einer Kante ist nicht zuverlässig herstellbar, weil die Justiergenauigkeit der Belichtung im Bereich von etwa einem Mikrometer liegt. Diese Positioniergenauigkeit gibt die kleinste Überlappung an, die reproduzierbar gefertigt werden kann. Im Masken-Layout muss zusätzlich noch ein Vorhalt berücksichtigt werden, der die Unterätzung des isotropen Ätzprozesses kompensiert, die so groß wie die Schichtdicke des Polysiliziums ist. Wie in Abbildung 6.15 zu sehen ist, kann mit der zuvor beschriebenen Technik eine Polysiliziumstruktur mit nur einem Mikrometer Überlappung über eine 3 μm hohe Stufe gefertigt werden.

6.5 Mesaätzen und Freistellen

Während der Herstellung der Sensoren muss immer wieder Silizium strukturiert werden. Dazu steht die RIE-Ätzanlage Plasmalab 80 der Firma Oxford zur Verfügung.

Besonders problematisch sind das Mesaätzen und das Freistellen der Bauelemente, bei der Fertigung der Chips. Dabei ist besonderes Augenmerk auf die vertikalen Flanken zu legen. Entlang dieser bildet sich später der Inversionskanal aus. Die elektrischen Eigenschaften, primär die Durchbruchspannung des Gate-Oxides und die Beweglichkeit der Ladungsträger, hängen sehr stark von der Qualität der Si – SiO₂ - Grenzfläche und diese hängt von der Flankenätzung ab.

Bisher wurde für diesen Arbeitsschritt ein Gasgemisch aus Schwefelhexafluorid, Argon und Sauerstoff¹³ verwendet [Fin00, Sed03]. Die REM-Aufnahmen in Abbildung 6.16 zeigen mit aus früheren Arbeiten bekannten Parametern geätzte Silizium-Mesen. In Abbildung 6.16a ist der Fotolack, der die nicht zu ätzenden Bereiche schützt, noch vorhanden. Die Lackmaske ist stark unterätzt. Die laterale Ätzrate ist etwa gleich groß wie die Vertikale. Abbildung 6.16b veranschaulicht den isotropen Charakter des Prozesses. In früheren Arbeiten wurde zugunsten einer glatten Siliziumoberfläche auf Anisotropie verzichtet [Sed03].

Auf diese Weise ist es allerdings nicht möglich, die Struktur der Lackmaske maßhaltig zu übertragen. Vor allem beim Freistellen der Bauelemente weichen die Strukturen mit diesem Prozess lateral um 2-3 μm von denen der Lackmaske ab. Das schließt eine Überlappung des Gates über der Mesa von 1 oder 2 μm aus (man vergleiche Abschnitt 6.4). Schwierig ist ein Vorhalt im Maskendesign zu berücksichtigen, weil die Dicke der Device-Layer der hier verwendeten SOI-Substrate, selbst um ca. 1 μm schwankt. Zudem wurden im Rahmen dieser Arbeit nur Restposten mit unterschiedlichen Schichtstärken des Device-Layers bezogen. Beide Punkte werden im Anhang B.6 erörtert. Zum Zeitpunkt des Masken-Designs ist also die exakte Dicke und folglich die genaue Ätztiefe noch nicht bekannt.

¹³SF₆ /O₂ /Ar

Zusätzlich führt diese isotrope Unterätzung stellenweise zur Ausbildung eines spitzen Winkels an der oberen Kante der Mesa. Das ist in Abbildung 6.16b gut zu erkennen und markiert. Genau an diesen Punkten treten später im Gate-Dielektrikum Feldüberhöhungen auf und der Transistor bricht schon bei niedrigen Gate-Spannungen durch.

Ziel muss also neben so glatten Flanken wie möglich auch eine gute Anisotropie sein. Um dieses zu erreichen, müssen zunächst die Mechanismen beim RIE-Ätzen betrachtet werden.

In einem RIE-Prozess vollzieht sich der Materialabtrag mit Hilfe zweier verschiedener Mechanismen.

- Aufgrund des Self-Bias oder der DC-Bias-Spannung werden die Ätzionen nahezu senkrecht auf das Substrat beschleunigt. Durch diesen Beschuss wird physikalisch geätzt.
- Zusätzlich laufen im Reaktor eine Vielzahl an chemischen Reaktionen ab. Im Plasma werden die vorhandenen Gase in Ionen und Radikale zerlegt. Diese greifen das Silizium chemisch an.

Um das physikalische Ätzen zu unterstützen, wird Argon in den Reaktor eingelassen. Als Edelgas ist es sehr inert und kaum an chemischen Reaktionen beteiligt.

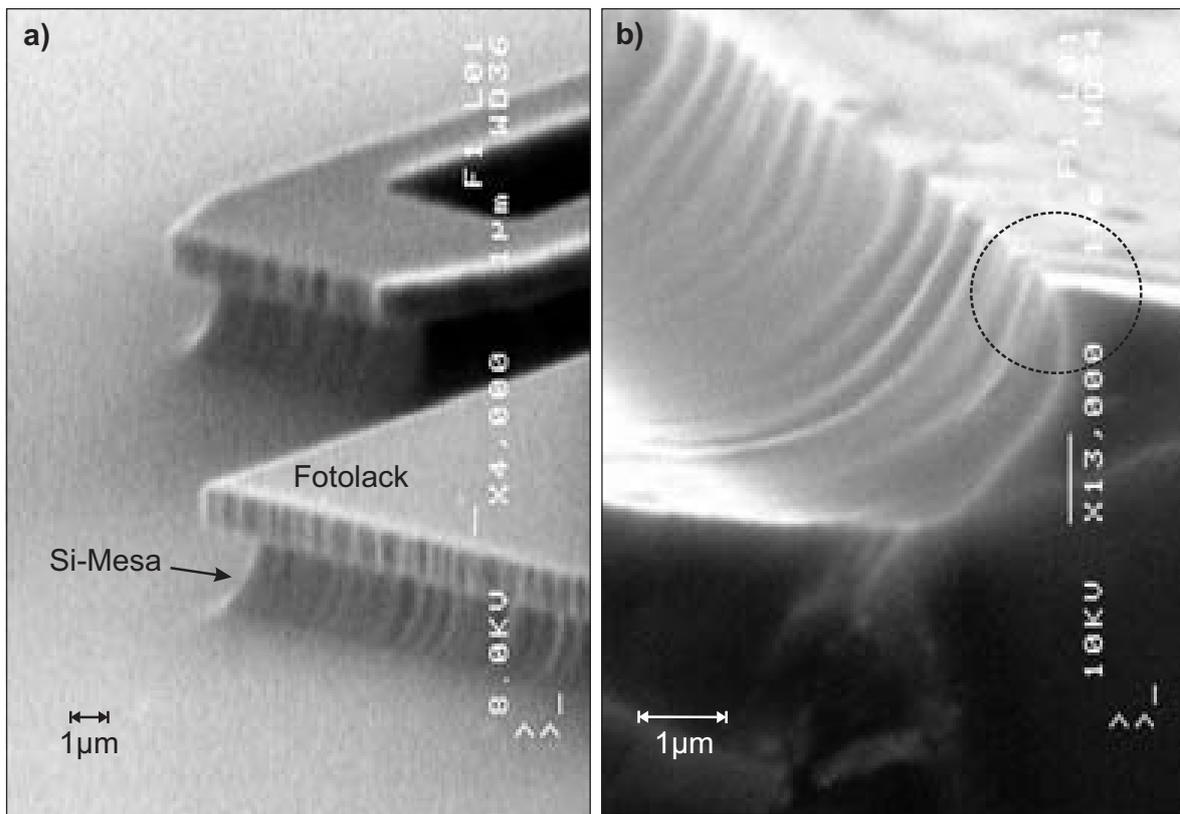


Abbildung 6.16: RIE-Ätzen ohne Anisotropie: a) vor und b) nach dem Strippen des Fotolackes

Die Ionen werden aufgrund des Potenzialunterschiedes zwischen Plasma und Substrat im Bereich der Randschicht beschleunigt. Durch gezieltes Manipulieren des DC-Bias kann deshalb die Ätzrate beeinflusst werden. Normalerweise werden die Ionen parallel zu den Feldlinien beschleunigt. Deshalb ist dieses Ätzen sehr anisotrop. Dieser Vorgang wird in Abbildung 6.5a geschildert. Kommt es allerdings im Bereich der Randschicht zu Stößen zwischen den Ionen, geht die gerichtete Wirkung der Ionen verloren. Um dem entgegen zu wirken, wird der Druck reduziert und so die freie Weglänge der Ionen vergrößert. Dieses Vorgehen hat aber nur einen Effekt, bis jene in etwa der Ausdehnung der Randschicht gleicht. Streuungen an dem Substrat selbst, die ebenfalls den Anisotropiefaktor mindern, können so nicht verhindert werden.

Das Beifügen von SF_6 zu dem Gasgemisch, trägt v. a. zum chemischen Ätzen bei. Im Plasma entstehen unter anderem Fluorradikale. Diese ätzen in der Reaktion



Silizium spontan. Das Reaktionsprodukt ist schon bei Raumtemperatur flüchtig und wird desorbiert. Dieser Vorgang ist streng isotrop. Das Ätzprofil sieht man in Abbildung 6.5b. Die Ätzrate dieses Ätzanteiles kann, wie bei den meisten chemischen Reaktionen, über die Temperatur beeinflusst werden.

Ein weiterer wichtiger Mechanismus ist die Bildung von Polymeren aus Fotolack und Sauerstoff während des Ätzprozesses. Diese kondensieren am Substrat und bilden eine Schicht, die das chemische Ätzen verhindert. Sie wird jedoch gleichzeitig durch die gerichteten Ionen physikalisch entfernt. Das ist insofern gewollt, als dass die Polymere auf den Flächen senkrecht zu den Feldlinien abgetragen werden und die Radikale das Silizium dort wieder chemisch angreifen. Auf den Seitenwänden werden die Polymere aber von den gerichteten Ionen nicht erreicht (siehe Abschnitt 6.4), verbleiben dort und bilden eine Seitenwandpassivierung.

Die Veränderung nur eines Parameters beim RIE-Ätzen hat sogleich Auswirkungen auf mehrere Eigenschaften des Ätzprozesses. Das ist v. a. problematisch beim Optimieren. Genauere Untersuchungen dieser Zusammenhänge und Beschreibungen der Mechanismen finden sich in [RBB91] und [TP91]. Es stellt sich heraus, dass es für ein anisotropes Ätzen von Flanken notwendig ist, die Temperatur auf -100°C abzusenken, um die chemische Reaktion und damit den isotropen Anteil so weit wie möglich zu unterdrücken. Gleichzeitig wurde der Druck bis zur technischen Grenze der Ätzanlage verringert, damit die Ionen möglichst parallel gerichtet auf das Substrat auftreffen.

Die drei beschriebenen Vorgänge, die während des RIE-Ätzens ablaufen, sind in Abbildung 6.17 illustriert.

Abbildung 6.18 präsentiert eine Mesa nach dem Ätzen mit den veränderten Parametern. Die Flanken sind nun senkrecht. Das Ergebnis beweist einen hohen Anisotropiefaktor. Allerdings belegt es eine weitere Schwierigkeit: Für äußerst glatte Kanten ist eine perfekte Fotolackmaske unabdingbar. Schon geringe Welligkeiten der Lackstrukturen übertragen sich auf die zu ätzenden Mesen. Des Weiteren müssen die Flanken des Fotolackes selbst bereits senkrecht sein. Bildet die Lackmaske

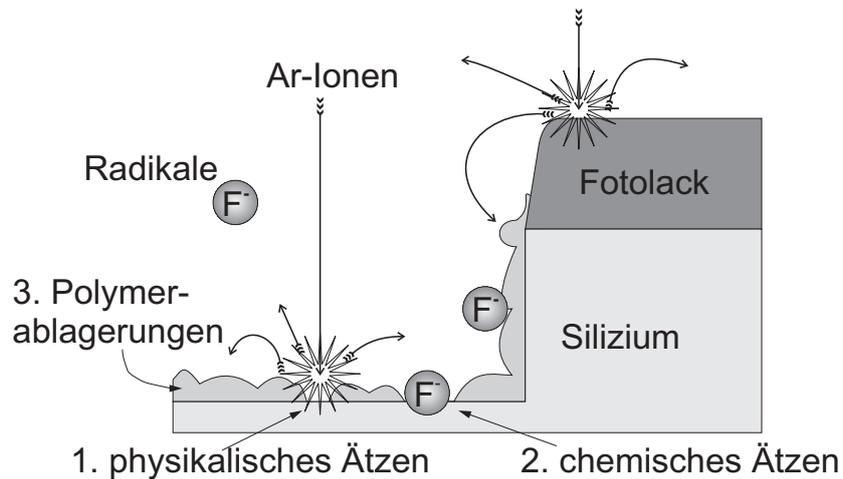


Abbildung 6.17: Vorgänge während eines RIE-Ätzprozesses

im Querschnitt ein Trapez, schwindet die Fotolackstruktur in lateraler Ausdehnung während des Prozesses und gibt mehr und mehr Silizium frei, weil sie selbst auch durch das physikalische Ätzen dünner wird [Mog83]. Daraus resultieren Flankenteilheiten kleiner als 90° . Auf schrägen Kanten kann jedoch der Ionenbeschuss wiederum die Passivierung aus den Polymeren leichter entfernen. Als Folge ätzen die Radikale die Kanten chemisch und rauhen sie auf. Die Herstellung der entsprechenden Fotolackmasken wurde bereits im Abschnitt 6.1 diskutiert.

Es zeigt sich, dass glatte Flanken hergestellt werden können. Allerdings nur, wenn die Mesahöhe $2\ \mu\text{m}$ nicht überschreitet. Versuche, eine $3\ \mu\text{m}$ tiefe Stufe zu ätzen, zeigten an manchen Stellen eine Aufrauung der Seitenflanken in der Nähe der oberen Kante. Einen Ausschnitt einer solchen Probe zeigt Abbildung 6.19. Die Seitenwandpassivierung hält nicht dauerhaft. Da die Ionen nicht perfekt parallel ausgerichtet sind und an der Oberfläche gestreut werden, wird sie nach und nach abgetragen. Deshalb beginnt die Seitenflanke von oben her rau zu werden, wo sie zeitlich dem Ionenbeschuss am längsten ausgesetzt ist.

Ungünstigerweise ist das genau der Ort, an dem sich später der Kanal des Transistors befindet. Die raue Oberfläche streut die Elektronen, die an ihr entlang fließen und vermindert damit ihre Beweglichkeit. Das mindert letztlich die Signalverstärkung des Transistors.

Ein weiteres Absenken des Druckes während des RIE-Ätzens könnte diesem Problem entgegenwirken. Ebenso könnten andere Fotolacke eine Verbesserung bringen. Aufgrund anderer Zusammensetzungen besteht die Chance, dass sie mehr Polymere produzieren, was zu stabileren Seitenwandpassivierungen führen könnte.

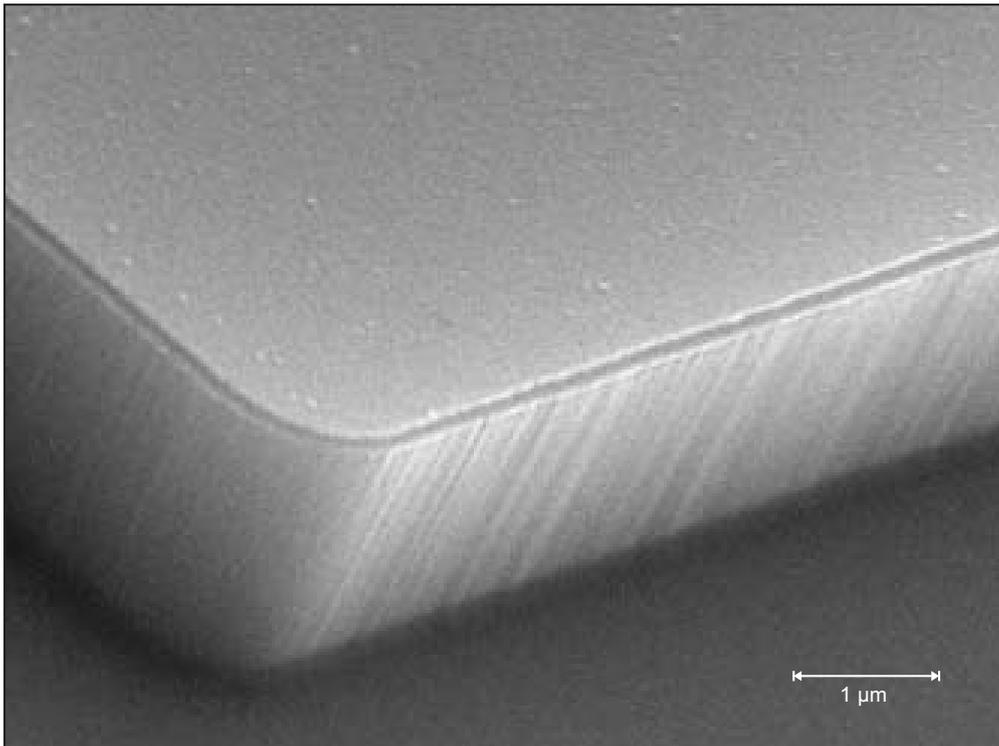


Abbildung 6.18: REM-Aufnahmen einer Silizium-Mesa-Flanke

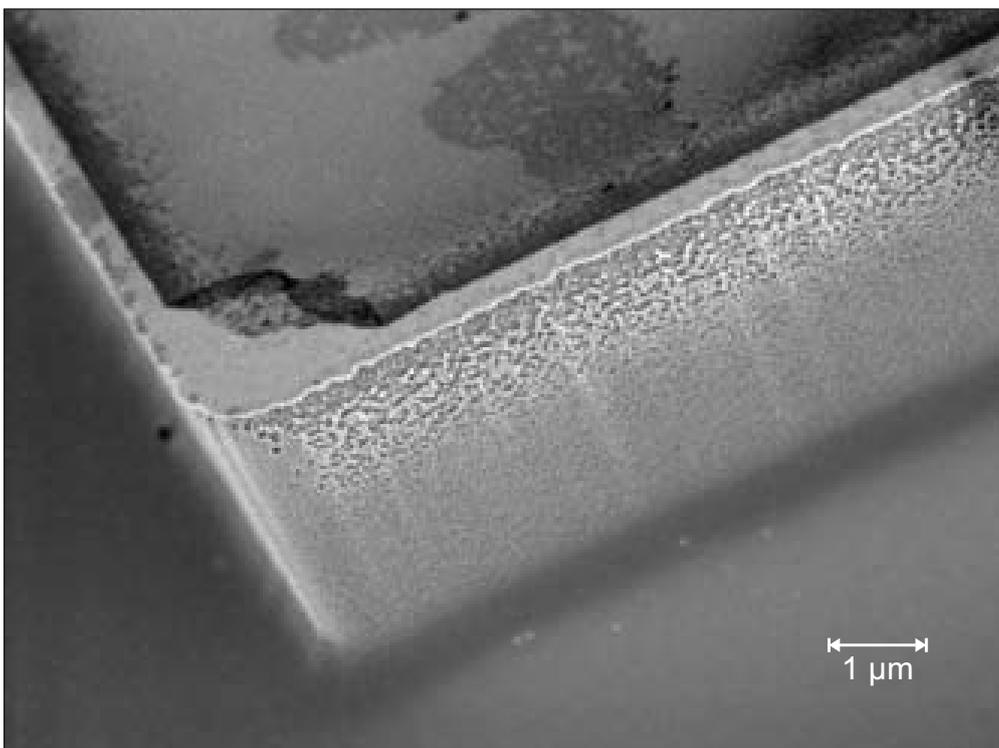


Abbildung 6.19: REM-Aufnahme einer 3 µm tief geätzten Siliziumstufe. Vom oberen Rand her beginnt die Oberfläche rauer zu werden.

6.6 Metallisierung

Auch in dieser Arbeit wurde die Metallisierung wiederum grundlegend überarbeitet und neu entwickelt. Sie muss für den in dieser Arbeit entwickelten Sensor folgenden Anforderungen genügen:

- hohe Leitfähigkeit,
- ohmscher Kontakt auf n-Silizium,
- gute Adhäsion auf Silizium, und auf Siliziumnitrid,
- Temperaturstabil bis 400 °C,
- leichte Strukturierbarkeit,
- hohe Reproduzierbarkeit,
- geringe Rauheit der abgeschiedenen Schichten,
- Erzielung homogener Schichtdicken über dem Wafer;
- das Metallisierungssystem sollte leicht gegen natürliche und thermische Oxidation zu Passivieren sein.

6.6.1 Entwicklung der Metall-Filme

Um diese Anforderungen zu erfüllen, soll eine dreilagige Schichtfolge aus Titan, Titannitrid¹⁴ und Aluminium eingesetzt werden.

Aluminium

Aluminium ist als guter Leiter mit einem spezifischen Widerstand von $2,7 \mu\Omega\text{cm}$ bekannt. Das Strukturieren mit Hilfe mit PNA-Ätzen¹⁵ mit einer hohen Selektivität zu Siliziumdioxid und -nitrid möglich. Eine Eigenschaft dieses Metalls ist die Selbstpassivierung. Bei vorhandenem Sauerstoff bildet sich eine Oxidschicht an der Oberfläche, die die weitere Oxidation stoppt.

Die Abscheidung erfolgt an der Universität der Bundeswehr in einer herkömmlichen Leybold-Heraeus Z550 Magnetron-Sputteranlage. Die Substrate liegen während der Abscheidung auf einem rotierenden Teller und bewegen sich ständig unter einem runden Sputterfleck hindurch. Zusätzlich ist die Sputterrate unterhalb des Targets nicht homogen. Diese drei Punkte führen zu inhomogenen Schichtdicken über den Wafer. Diese Homogenität kann in gewissen Grenzen durch die Sputterleistung und den Prozessdruck beeinflusst werden. Das ist aber ungünstig, weil dadurch auch andere Schichteigenschaften, wie z. B. Leitfähigkeit und Rauigkeit,

¹⁴TiN

¹⁵16 Teile H_2PO_4 : 2 Teile H_2O : 1 Teil HNO_3 : 1 Teil $\text{H}_3\text{C}-\text{COOH}$

verändert werden. Für die Abscheidung des Aluminiums wurde eine Schattenmaske berechnet und vor das Target eingebaut. Diese deckt bestimmte Teile des Targets ab und gleicht so die Inhomogenitäten aus. Eine Abbildung findet sich im Anhang D. Diese Schicht wird im DC-Modus gesputtert. Wie schon bemerkt, je höher die Leistung, desto größer die Abscheiderate. Jedoch steigt damit auch die Rauigkeit. Deshalb kann man diesen Parameter nicht zu weit steigern. Mit der Blende ergeben sich bei 400 Watt sehr gute Ergebnisse.

Titan

Direkt auf das Silizium wird Titan aufgebracht. Während des Temperns bei ungefähr 400 °C verbinden sich diese Feststoffe an ihrer gemeinsamen Grenzfläche zu Titansilizid¹⁶, wobei ein ohmscher Kontaktwiderstand zu n-Silizium entsteht. Zusätzlich hat der Einsatz von Titan den Vorteil, dass dieses das natürliche SiO₂ auflöst und so den Kontakt zusätzlich verbessert. Nebenbei ist Titan als guter Haftvermittler bekannt und erhöht die Adhäsion des Aluminiums auf Siliziumdioxid und -nitrid.

Dieses Material wird ebenso in der Sputteranlage Z550 abgeschieden. Da diese Schicht direkt auf dem Halbleiter aufgesputtert wird, besteht die Gefahr von Plasmaschäden. Deshalb wird hier eine nur geringe Leistung in das Plasma eingekoppelt. Die Titanabscheidung erfolgt im DC-Magnetron-Modus mit einem Target-Potenzial von 350 V.

Titannitrid

Ein weiteres Problem ergibt sich aus dem Gebrauch des Aluminiums. Da in diesem bei höheren Temperaturen (schon unterhalb der Schmelztemperatur $T_S = 660,2\text{ °C}$ von Aluminium) Silizium zu einem geringen Teil löslich ist (ungefähr 0,5 % bei 450 °C) [Eis01], wandert es in das Metall. Die entstehenden Lücken füllt das Aluminium auf und gelangt so in das Halbleitermaterial hinein und bildet dort mit dem Silizium ein Silizid. Aufgrund der höheren Löslichkeitsgeschwindigkeit im Kristall in Richtung der $\langle 1-0-0 \rangle$ -Ebene als in Richtung der $\langle 1-1-1 \rangle$ -Ebene formiert sich das Silizid zu Pyramiden (den sog. „Spikes“). Diese wachsen bis zu mehreren hundert Nanometer in das Material hinein [WMF88]. Liegt ein pn-Übergang unterhalb eines Kontaktes, kann ihn möglicherweise ein solcher Spike durchdringen und auf diese Weise kurzschließen. Aufgrund der Konstruktion, der in dieser Arbeit hergestellten vertikalen Transistoren, als Mesa liegt ein Kontakt immer oberhalb der epitaktisch gewachsenen npn-Schicht (siehe Abbildung 4.3). Man kann also solche eine Metallisierung nicht ohne weiteres Tempern. Temperaturstabile Kontakte sind jedoch aus folgenden Gründen dringend erforderlich:

- Formiergastempnen bei 400 bis 450 °C reduziert die Grenzflächenzustandsdichte und verbessert dadurch die Eigenschaften der Bauelemente.

¹⁶TiSi₂

- Ein Tempern des Metallisierungssystems von 300 bis 400 °C ist für die Bildung des Titansilizids, und damit geringer Kontaktwiderstände notwendig.
- Das Tempern der Aluminium-Pads zwischen 300 bis 400 °C erleichtert das Setzen der Bond-Verbindungen.
- Ein Gassensor, der auf dem Prinzip der Austrittsarbeitsänderung funktioniert, wird für die Detektion bestimmter Gase bei 200 °C und höher betrieben und das unter Umständen dauerhaft.¹⁷

Um das „Spiking“ zu verhindern, wird dem Aluminium ein Prozent Silizium beigemischt (das sog. AlSi), da diese Konzentration über der Löslichkeitsgrenze liegt, und so keine Diffusion des Siliziums in das Metall stattfindet. Versuche zeigten aber, dass das in Verbindung mit einer Titansilizid-Kontaktschicht nicht zuverlässig funktioniert.

Ein anderer Weg, die oben genannte Reaktion zwischen Aluminium und Silizium zu verhindern, ist die Verwendung einer Titanitridschicht als Diffusionsbarriere. Dieses Material ist chemisch und thermisch sehr stabil und als Diffusionsbarriere für Aluminium und Silizium und auch anderen Schichtsystemen bekannt und gut untersucht [SCMN80, ACG⁺82, Sch83, LP84, SFS85, JP92, CPG99, EKA⁺00, YLR⁺01, CMP⁺04]. Die Stöchiometrie sowie die physikalischen Eigenschaften von Titanitridschichten hängen stark von den Abscheidebedingungen ab [LP84]. Die Stöchiometrie hat wiederum Einfluss darauf, wie stabil Titanitrid unter thermischer Belastung ist und ob es die Diffusion von Aluminium in Silizium verhindert [JP92].

Diese Titanitridschichten werden mittels reaktivem Magnetron-Sputtern in einer Sputteranlage Z550 unter Verwendung eines Titan-Targets hergestellt. Dazu wird durch eine angelegte Spannung zum einen ein Plasma gezündet, und zum anderen werden Argon-Ionen auf das Target beschleunigt und schlagen dort durch Impulsübertrag einzelne Ti-Atome aus der Target-Oberfläche. Die herausgelösten Atome können dann auf dem Substrat kondensieren.¹⁸ Das Gas in der Beschichtungskammer besteht aus einem Argon-Stickstoffgemisch. Während des Filmwachstums reagieren die Titanatome mit dem Stickstoff, sodass sich eine TiN-Schicht ausbildet. Erste Experimente zeigten, aber während der Temperschnitte keine ausreichende Diffusionsbarrierewirkung. Die Folge war das Ausfallen sämtlicher Bauelemente nach Prozessschritten mit Temperaturen um 400 °C.

Um zu bestätigen, dass die Metallisierung die Ursache dieser Ausfälle ist, wurde auf einem Silizium-Wafer nach einem HF-Dip (um das natürliche Oxid zu entfernen) das herkömmliche Metallisierungssystem mit den in der Tabelle 6.5

¹⁷Ein Ziel dieser Arbeit ist es zu zeigen, dass ein Sensor durch das hier vorgestellte Konzept auch bei Temperaturen höher als 400 °C funktioniert.

¹⁸Die Beobachtungen in dieser Arbeit lassen vermuten, dass das Titan zunächst auf dem Substrat kondensiert, bevor es mit dem Stickstoff reagiert, weil die stöchiometrische Zusammensetzung mehr von der Substrattemperatur als der Gaszusammensetzung abhängt. Da das aber für diese Arbeit nicht von Belang ist, wurde es nicht weiter verfolgt.

Metall	Modus	Parameter	Zeit
Titan	DC	Substrat geerdet, 350V, $p_0 < 8 \cdot 10^{-5}$ mbar $p_{Ar} = 8 \cdot 10^{-3}$ mbar	3 min
Titannitrid	RF	Substrat erdfrei, 400W $p_{Ar} = 7,5 \cdot 10^{-3}$ mbar $p_{N_2} = 2,5 \cdot 10^{-3}$ mbar	60 min
Aluminium	DC	Substrat geerdet, 350V $p_{Ar} = 8 \cdot 10^{-3}$ mbar	36 min

alle Abscheidungen rotierend

Tabelle 6.5: Parameter für die Herstellung der bisher verwendeten Metallisierung

angegebenen Parametern mit der Sputteranlage Z550 abgeschlossen. Zwischen den Abscheidungen der einzelnen Metalle wurde der Rezipient nicht geöffnet und nicht belüftet. Die Hälfte der so hergestellten Probe wurde in einer Formiergas-Atmosphäre für 30 min einer Temperatur von 450 °C ausgesetzt. Nach dieser thermischen Belastung wurde das Aluminium mit PNA-Lösung entfernt. Das Ergebnis der anschließenden SIMS-Analyse zeigt Abbildung 6.20. Man sieht in 6.20a das Aluminium- und Titanprofil der Getemperten und in 6.20b ungetemperten Probe. Es muss hinzugefügt werden, dass die Verhältnisse der Zählraten zueinander keine direkten Schlüsse auf die stöchiometrische Zusammensetzung erlauben. Klar bestätigen die Messungen: Die so hergestellten Schichten sind nicht thermisch stabil, weil die beiden Profile doch sehr stark voneinander abweichen. Nach dem Tempern (Abbildung 6.20b) ist kein deutliches Interface mehr zusehen. Die Zählrate für das Silizium steigt ab dem Start der Messung bis auf ein konstantes Niveau an. Das Erreichen dieses Niveaus lässt auf den Beginn des Siliziumsubstrates schließen. Die Zählrate des Aluminiums bleibt bis kurz vor den Beginn des Substrates annähernd konstant und fällt dann langsam ab. Aufgrund des „Knock On“-Effektes erscheint dieser Abfall in der Messung noch langsamer als der Tatsächliche. Trotzdem zeigt diese Messung, dass Silizium bis in das Titannitrid diffundiert ist, und ebenso ist Aluminium durch das Titannitrid und Titansilizid bis in das Substrat gewandert. Diese Titannitrid-Schicht ist als Diffusionssperre ungeeignet.

XPS-Messungen ergaben, dass der Sauerstoffanteil der oben diskutierten Schichten bei 31,5 % liegt (siehe Tabelle 6.7). Bei diesen Mengen kann man nicht mehr von einem TiN sprechen. Viele Gruppen berichten, dass ihre gesputterten Titannitrid-Schichten als Verunreinigung hauptsächlich Sauerstoff enthalten [SCMN80, Sch83, LP84, SFS85, KFP⁺87]. Ziel muss es aber sein diesen so gering wie möglich zu halten, denn je größer der Anteil an Sauerstoff, desto geringer ist die Leitfähigkeit [SCMN80, LP84, KFP⁺87]. Armigliato et al. und Kumar et al. finden eine Abnahme des Stickstoffanteils mit Zunahme des Sauerstoffpartialdruckes

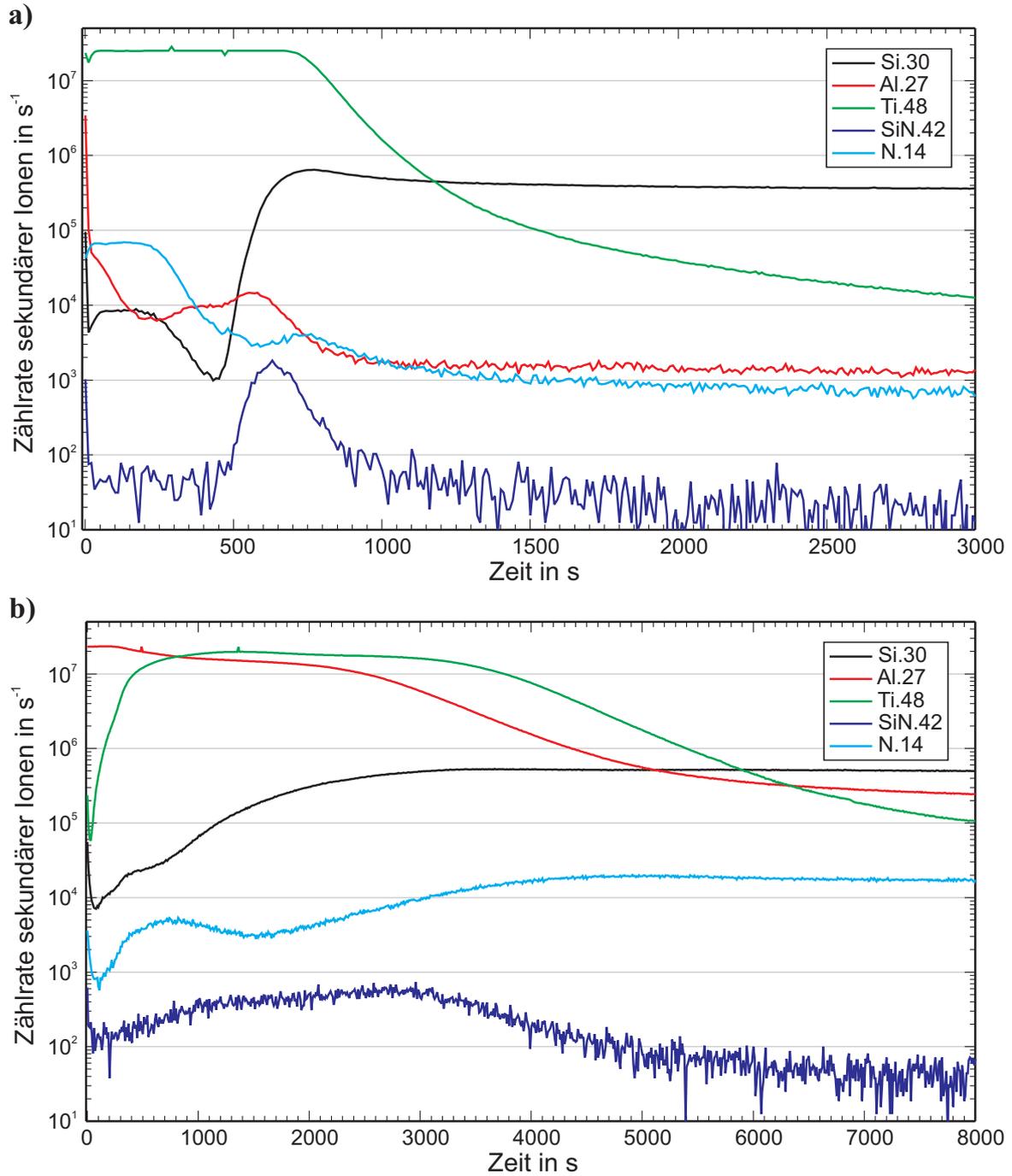


Abbildung 6.20: SIMS-Analyse des bisherig verwendeten Metallisierungssystem a) vor und b) nach dem Formiergastempnen bei 450 °C

[ACG⁺82, KFP⁺87]. Die Barrierewirkung hängt jedoch von der Stöchiometrie des Titanitrids ab [JP92]. Für das Reduzieren des Sauerstoffanteils gibt es folgende Möglichkeiten:

- Abpumpen, um den Sauerstoffpartialdruck und damit das Angebot an Sauerstoff so weit wie möglich zu verkleinern.
- Spülen des Rezipienten, sodass der Sauerstoffpartialdruck gesenkt wird.
- Rückspütern, d. h. ein kürzeres Ätzen vor Beginn der Schichtabscheidung.
- Bias-Sputter, d. h. ein Anlegen einer negativen Spannung an das Substrat.

Abpumpen alleine reicht jedoch nicht aus, weil die Reaktionswahrscheinlichkeit zwischen Titan und Sauerstoff drei Größenordnungen größer ist als zwischen Titan und Stickstoff [gme71]. Der Restgasdruck p_0 , auf den der Rezipient der Leybold-Heraeus Z550 Magnetron Sputteranlage abgepumpt werden kann, liegt um die 10^{-6} mbar. Gesputtert wird bei einem Druck p_{Sp} von ungefähr 10^{-3} mbar. Das Restgas in Vakuumanlagen bei Drücken kleiner als 10^{-4} mbar besteht zur Hälfte aus Wasserdampf [Mai83], der beim Zünden eines Plasmas in Wasser- und Sauerstoffionen bzw. Radikale zerlegt wird. Wegen der hohen Affinität zu Titan werden Letztere sofort in die Titanschichten eingebaut.

Zum Spülen des Rezipienten wurden nach dem Abpumpen die Flüsse für Argon und Stickstoff entsprechend den Arbeitsdrücken eingestellt und für 2 Tage in diesem Zustand belassen, bevor die Schicht abgeschieden wurde. Dadurch konnte der Sauerstoffanteil erheblich gesenkt werden.

Beim Rückätzen werden die Argonionen auf das Substrat und auch auf den Substratteller beschleunigt. Dort haftende Feuchtigkeit wird durch Stöße abgetragen. Aufgrund des Abpumpens werden die Wassermoleküle oder die im Plasma entstehenden Sauerstoffionen aus dem Rezipienten transportiert. Der Sauerstoffpartialdruck verringert sich. Allerdings werden auf diese Weise auch die vorher abgeschiedenen Schichten geätzt und eventuell könnte der Halbleiterkristall dabei geschädigt werden [Sed03].

Als letzte Möglichkeit wurde an das Substrat eine Bias-Spannung angelegt. Auch das reduziert den Sauerstoffanteil in der Schicht [KFP⁺87]. Dieses Festlegen des Substratpotenzials bewirkt einen Materialabtrag vom Substrat durch Ionenbeschuss wie am Target [WMF88]. Durch unterschiedlich angelegte Spannungen am Substrat und Target kann das Verhältnis zwischen Aufwachsen und gleichzeitigem Ätzen eingestellt werden.

Das Ergebnis der Kombination von Spülen und Bias-Sputtern, um Sauerstoff zu reduzieren, zeigt die XPS-Messung Abbildung 6.21. Deutlich erkennbar ist die Verringerung der Zählrate für Sauerstoff sowie die Erhöhung der Zählrate für Stickstoff.

Neben den Drücken der Gase kann für den RF-Sputter-Modus an der Magnetron Sputteranlage Z550 von Leybold-Heraeus auch die Leistung zwischen 0 und 2 kW eingestellt werden. Die Sputter-Rate für ein abgeschiedenes Material ist zur

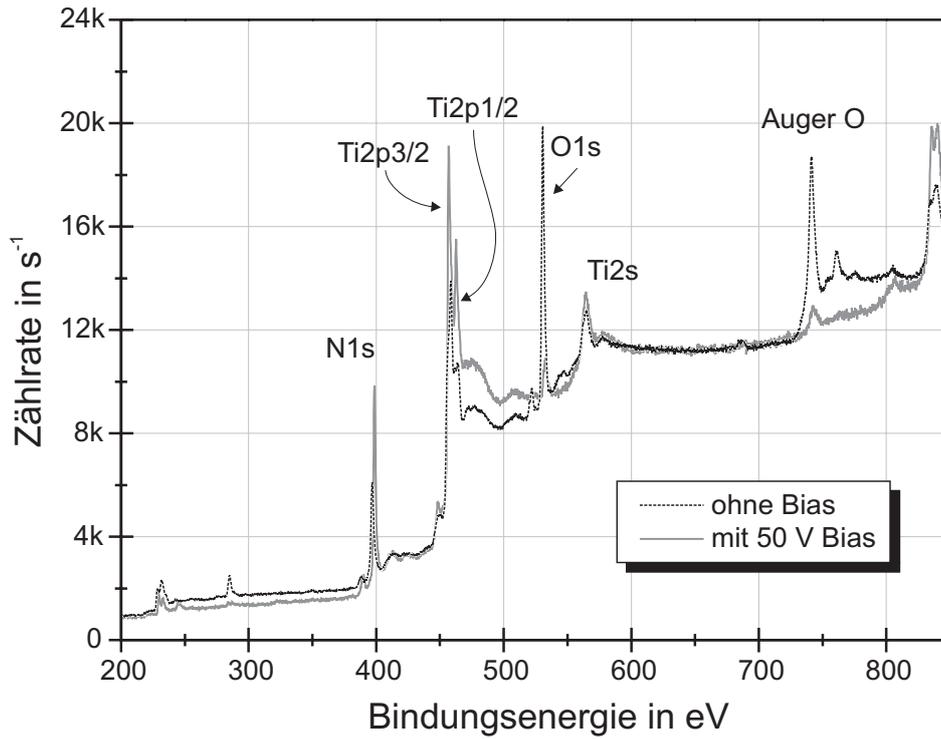


Abbildung 6.21: XPS-Messungen verschiedener TiN-Schichten

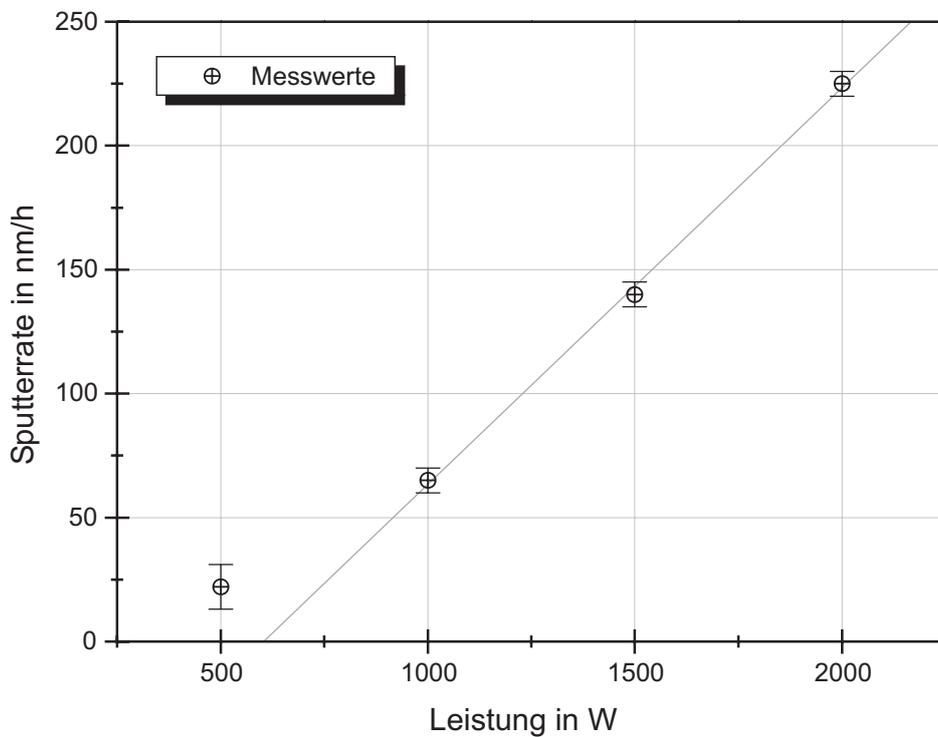


Abbildung 6.22: Abhängigkeiten der Sputterraten von der Leistung; reaktives RF-Sputtern bei 50 V Bias.

Metall	Modus	Parameter	Zeit
Titan	DC	Substrat geerdet, 350V, $p_0 < 6 \cdot 10^{-5}$ mbar $p_{Ar} = 8 \cdot 10^{-3}$ mbar	2,5 min
Titannitrid	RF	Bias 50 V, 1500W Target-Potenzial: 118 V Substrat-Potenzial: 50 V $p_{Ar} = 7 \cdot 10^{-3}$ mbar $p_{N_2} = 3 \cdot 10^{-3}$ mbar	45 min
Aluminium	DC	Substrat geerdet, 400W $p_{Ar} = 8 \cdot 10^{-3}$ mbar	22 min

alle Abscheidungen rotierend

Tabelle 6.6: Parameter für die Herstellung der Metallisierung

eingekoppelten Leistung proportional. Sputter-Raten für verschieden eingekoppelte Sputter-Leistungen sind für das reaktive Sputtern von Titannitrid bei einem Argon-Partialdruck $p_{Ar} = 7 \cdot 10^{-3}$ mbar und Stickstoff-Partialdruck $p_{N_2} = 3 \cdot 10^{-3}$ mbar in Abbildung 6.22 gezeigt. Ein linearer Zusammenhang ist nur für Leistungen oberhalb von 1000 Watt zu erkennen. Extrapoliert man die Gerade durch die Messwerte für Leistungen größer als 1000 W, dürfte sich keine Abscheiderate für 500 W ergeben. Es findet trotzdem eine Abscheidung bei 500 W statt. Das lässt den Schluss zu, dass hier zwei verschiedenen Phasen abgeschieden werden.

Mit diesen Erkenntnissen wurde das oben beschriebene Experiment wiederholt, wobei dieses Mal die Parameter entsprechend der Tabelle 6.6 verwendet wurden. Zwischen dem Abscheiden des Titannitrids und des Aluminiums wurde der Rezipient geöffnet und die Proben für einen Tag unter Normalbedingungen dem Sauerstoff der Raumluft ausgesetzt.¹⁹ Wieder wurde die eine Hälfte wie abgeschieden untersucht und die andere nach 30ig minütigem Tempern bei 450 °C in Formiergas, um das Diffundieren des Aluminiums zu provozieren.

¹⁹Prozesstechnisch ist dieser Schritt notwendig. Die Wafer müssen ausgebaut werden, um das Aluminium-Target durch Sputterätzen reinigen zu können. Sinke et al. und Kumar et al. sind der Meinung, ein Aussetzen an der Luft bei Normalbedingungen sei sogar für eine Diffusionsbarrierewirkung förderlich [SFS85, KFP⁺87]. Sie erklären die Verbesserung der Diffusionsbarriere damit, dass Aluminium nicht durch das Titannitrid wandern kann, sondern nur entlang der Korngrenzen des polykristallinen Filmes. Der Sauerstoff der Raumluft oxidiert diese Korngrenzen auf und erschwert so die Diffusion des Aluminiums. Die Effektivität einer solchen Nachbehandlung wird jedoch von Chun et al. angezweifelt [CPG99]. In dieser Arbeit konnte das nicht abschließend geklärt werden. Es ist aber anzumerken, dass die ersten Versuche bei denen das Vakuum nicht unterbrochen wurde (Tabelle 6.5), keine Barrierewirkung aufwiesen. Mit den veränderten Prozessparametern, unter anderem dem Unterbrechen des Vakuums, hingegen wurden temperaturstabile Schichten erzeugt (Tabelle 6.6).

Eine weitere mögliche Nachbehandlung von Titannitridschichten, die kein stöchiometrisches Verhältnis von 1,0 aufweisen, ist das Ausheilen bei 600 bis 700 °C in N₂- oder NH₃-Gas [JP92].

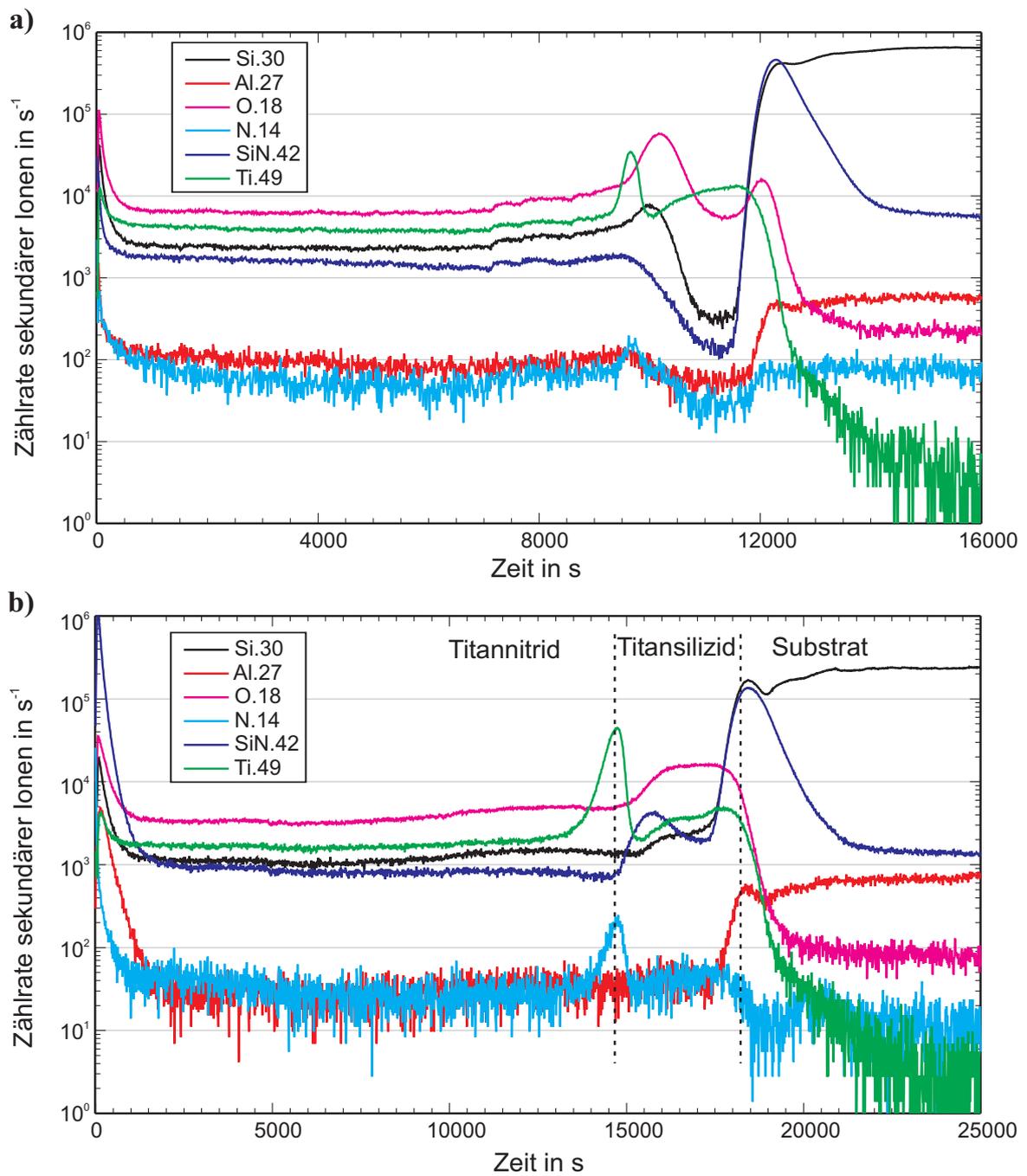


Abbildung 6.23: SIMS-Messung der abgeschiedenen Titanitrid-Schicht a) vor und b) nach dem Formiergastempnen bei 450 °C

Der Vergleich der SIMS-Messungen der Getemperten (Abbildung 6.23a) und ungetemperten Proben (Abbildung 6.23b) zeigt nur minimale Unterschiede. Der einzige Bereich, in dem eine wirkliche Veränderung eintritt, ist im Bereich des Titansilizids. Der Siliziumgehalt nimmt um ungefähr eine Größenordnung zu und ist nach dem Tempern, wie das Titan, fast konstant verteilt. Dieses erwartete und auch erwünschte Verhalten ist durch die Silizidbildung zu verstehen. Der Sauerstoff, den man zunächst hauptsächlich an den Grenzflächen findet, ist nach dem Tempern im Silizid gleichmäßig verteilt. Der Abfall des Sauerstoffes deckt sich nun mit dem des Titans. Die Häufung des Sauerstoffes an den Grenzflächen ist durch die Feuchtigkeit im Rezipienten bedingt. Während das Plasma brennt, wird diese verbraucht. Da kein Sauerstoff nachgeliefert wird, sinkt seine Konzentration. Der gleichzeitige Abfall der Sauerstoff- und Titankonzentration in diesem Bereich gibt einen Hinweis, dass das Titan eine dünne Siliziumdioxidschicht aufbraucht, was der Verbesserung der Kontakte dient.

Eine weitere Veränderung zwischen Abbildung 6.23a und Abbildung 6.23b ist am Anfang der Titannitridschicht zu sehen. Die Aluminium-Konzentration fällt nach dem Tempern nicht mehr direkt auf das Niveau des Grundrauschens ab, sondern erst nach einer kurzen Weile. Diese Spuren mögen von einem nicht vollständig durchgeführten PNA-Ätzschritt stammen, bei dem Spuren an Aluminium an der Oberfläche zurückbleiben. Es könnte aber auch eine tatsächliche Diffusion stattgefunden haben. Trotzdem reicht die Diffusionstiefe nicht weit in das Titannitrid hinein. Sonst verschwindet die Konzentration im Grundrauschen. Ein nach Tabelle 6.6 hergestelltes Metallisierungssystem bietet also unbeschadet thermischer Belastung von 450 °C genügend Barrierewirkung gegen Aluminiumdiffusion in das Silizium.

Die beiden SIMS-Messungen der Abbildung 6.23 zeigen eine Zunahme der Aluminiumzählrate beim Übergang vom Titan bzw. Titansilizid zum Substrat um zirka eine Größenordnung. Dieser hohe Anteil dieses Metalls ist nicht real vorhanden, sondern vielmehr durch eine Messungenauigkeit des Simsmessgerätes zu erklären. Die Massenseparation ist nicht genau genug, um Silizium (Nukleonenzahl 28) von Aluminium (Nukleonenzahl 27) scharf und exakt zu trennen. In der Folge werden einzelne Siliziumatome als Aluminium detektiert. Erreicht die SIMS-Messung das Substrat, steigt die Menge des Siliziums so stark an, dass sie die Zählrate des Aluminiums massiv beeinflusst.

Neben dem Verringern des Sauerstoffpartialdruckes ist sicher gleichermaßen die Erhöhung der Sputterleistung entscheiden für dieses Ergebnis. Je höher diese ist, desto wärmer ist die Substrattemperatur, weil die größere Intensität des Ionenbeschusses das Substrat stärker aufheizt. Mit der Zunahme der Temperatur nimmt auch die Oberflächenbeweglichkeit zu. Damit steigt wiederum die Wahrscheinlichkeit einer kristallinen Bindung zwischen Titan und Stickstoff. Erst ab einer gewissen Substrattemperatur die mit ungefähr 1000 Watt Sputterleistung korreliert, fängt ein Titannitridfilm an aufzuwachsen. Unterhalb dieser Temperatur wächst ein Titanfilm auf, dessen Wachstumsraten aufgrund des besseren Haftungskoeffizienten größer sind (man vergleiche dazu Abbildung 6.22). Dieser Effekt wird durch

	Ti	N	O	$\frac{N}{Ti}$	$\frac{O}{Ti}$
alte Metallisierung	52,81 %	15,69 %	31,50 %	0,29	0,57
neue Metallisierung	62,13 %	32,96 %	4,91 %	0,53	0,08

Tabelle 6.7: Stöchiometrie der untersuchten Schichten

ein Anlegen einer Bias-Spannung noch verstärkt. Allerdings darf diese Temperatur genauso wenig zu heiß werden, weil sonst die sich ergebenden Verspannungen zu Rissen in dem Film führen [KFP⁺87]. Außerdem trägt auch das Anlegen einer Bias-Spannung selbst zur Qualitätsverbesserung der Diffusionsbarriere bei. Durch das Ätzen werden das vorhandene Titanoxid und Titan schneller abgetragen als das chemisch und mechanisch sehr stabile Titannitrid. Wie aus Tabelle 6.7 hervorgeht, vergrößert sich das Verhältnis aus Titan zu Stickstoff, während sich das aus Titan zu Sauerstoff verringert. Zudem wird die Schicht durch den verstärkten Ionenbeschuss verdichtet.

Mit XPS-Messungen wurde die Zusammensetzung des Titannitrides entsprechend der Tabelle 6.7 bestimmt. Widerstandsmessungen ergaben für dieses Material einen Schichtwiderstand von $42,6 \pm 2,4 \mu\Omega\text{cm}$. Dieser sehr gute Leitwert ist auf den hohen Anteil Titan zurückzuführen.

Die morphologische Struktur jeweils vor und nach dem Tempern sieht man in den REM-Aufnahmen in Abbildung 6.24. Während des Sputterns wächst das Titannitrid säulenartig auf dem Titan auf. Nach dem Temperenschritt ist das Titan aufgrund der Silizidbildung nicht mehr zu sehen. Die Kristalle des Titannitrids sind kleiner und nicht mehr stäbchenförmig. Die nach dem neuen Verfahren erzeugte Schicht ist trotz der thermischen Belastung immer noch kompakt und zeigt keine Risse.

Das Metallisierungssystem

Für den Sensor, der in dieser Arbeit gefertigt wird, soll nun das im Folgenden beschriebene Schichtsystem verwendet werden:

Als Erstes wird Titan abgeschieden. Wird diese Schicht zu dick gewählt, verbraucht sie zu viel Silizium auf. Unter Umständen werden davon Raumladungszonen beeinflusst. Das Titan muss jedoch genügend dick sein, um mechanische Spannungen des Titannitrids ausgleichen zu können. Versuche ohne diesen Haftvermittler führten zu keinem brauchbaren Ergebnis. Das Titannitrid reißt und wird brüchig. Infolgedessen werden 35 nm Titan gesputtert.

Folgend wird eine Titannitridschicht abgeschieden. Schon eine 75 nm dicke Titannitridschicht, die mit den angegebenen Parametern hergestellt wurde, bewies eine ausreichende Barrierewirkung für das Tempern bei 450 °C. Um diese mit Sicherheit zu erhalten und nicht wegen Inhomogenitäten der Abscheiderate zu unterschreiten, wird eine Schichtstärke von 100 nm gewählt.

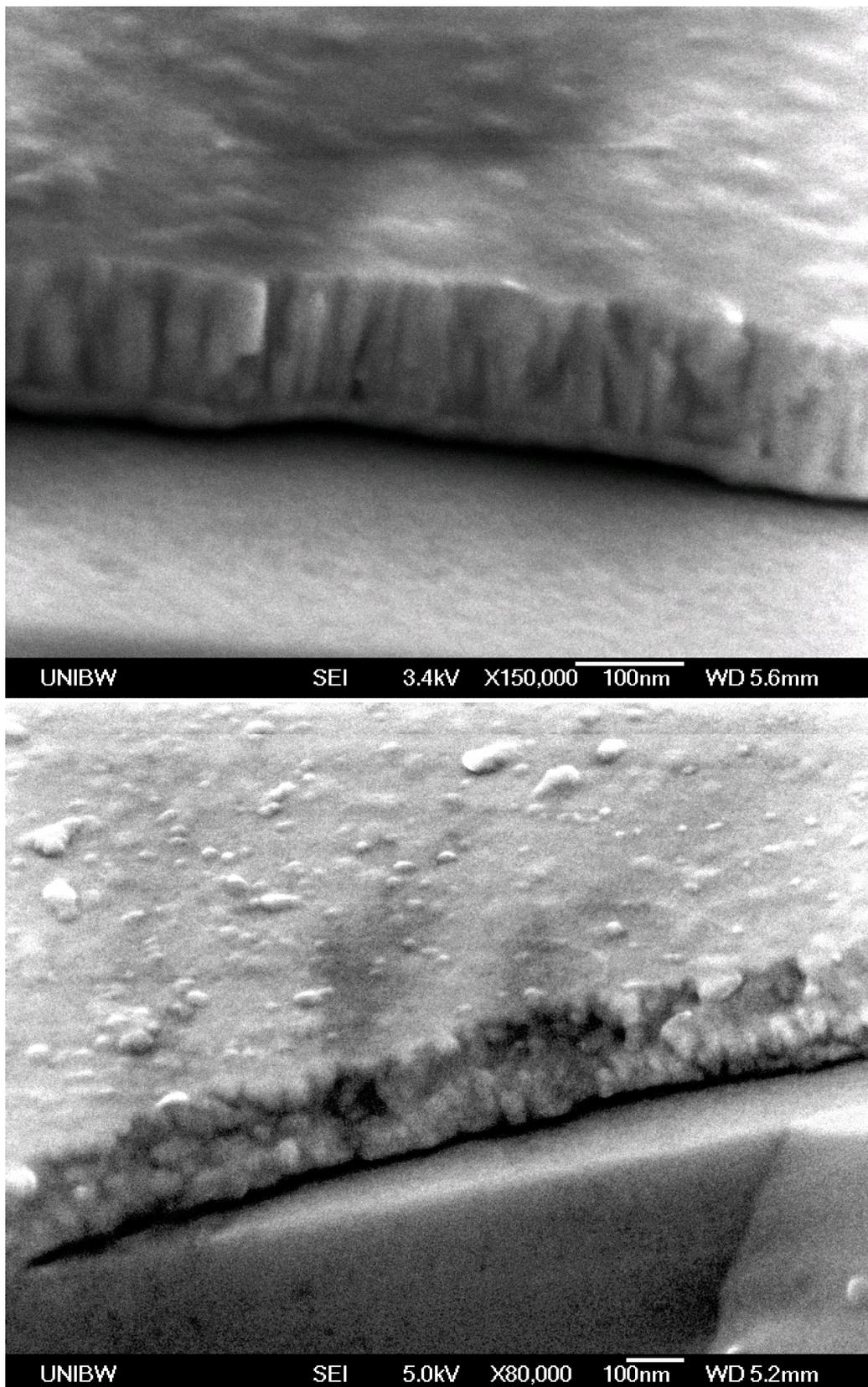


Abbildung 6.24: REM-Aufnahmen einer Titan/Titannitrid-Schicht (oben) vor und (unten) nach dem Tempern bei 450 °C

Nachdem der Rezipient der Sputteranlage belüftet wurde und die Wafer aus- und wieder eingebaut wurden,¹⁹ wird erneut ein Vakuum erzeugt und Aluminium²⁰ abgeschieden. Das ist das eigentliche Strom führende Material. Der Leitwert ist in erster Näherung proportional zu Schichtdicke. Auch für das Bonden muss das Aluminium eine genügende Stärke aufweisen. In der Praxis stellten sich 300 nm dafür als ein brauchbarer Wert heraus.

6.6.2 Strukturierung der Metallisierung

Die veränderten Parameter lassen die Verwendung des bisher an der Universität der Bundeswehr benutzten Strukturierungsprozesses [Fin00, Sed03] nicht zu. Allein schon die notwendige hohe Sputterleistung von 1500 W zur Herstellung der TiN-Schicht verbietet sich für den bisherigen Prozess [Sed03]. Die dabei auftretenden Temperaturen zerstören schon vor dem Lift-Off-Prozess den Fotolack. Wegen des speziellen Designs, der in dieser Arbeit verwendeten Masken, insbesondere den langen und dichtgepackten Leiterbahnen, ist die Ausbeute mit einem Lift-Off-Prozess nicht sehr hoch.

Die Einführung eines neuen Strukturierungsprozess ist demzufolge notwendig. Weiterhin kann die nasschemische Strukturierung des Aluminiums mit PNA¹⁵ genutzt werden. Dieser Prozessschritt bietet neben einer hohen Selektivität sehr reproduzierbare Ergebnisse. Zudem kann dieser in Bechergläsern durchgeführt werden. Das Titan und Titanitrid wird nasschemisch strukturiert werden.

Diese Materialien werden von Flusssäure²¹ oder Königswasser²² angegriffen. Auch Basen wie KOH²³ und TMAH²⁴ oder der RCA-1-Clean lösen die beiden Materialien. Letzterer bietet die höchste Selektivität zu Aluminium und Siliziumnitrid und soll deshalb verwendet werden.

Schwierig ist bei der Verwendung von RCA-1 die Selektivität zum Fotolack, da diese Lösung, wie in 6.2 besprochen, ja gerade zum Entfernen organischer Verunreinigungen eingesetzt wird. Umkehrlacke²⁵ wie der AR-U 4030 halten im alkalischen Milieu lange genug stand, wenn sie als Umkehrlack betrieben werden, weil die zu ätzenden Schichten sehr dünn (insgesamt nur 140 nm) sind. Allerdings darf die Temperatur der RCA-1-Lösung nicht über 35 °C steigen, da sonst der Lack zu schnell angegriffen wird.

Bei der Kombination der Strukturierung einer Schicht aus Titanitrid und Aluminium mit einem PNA- und anschließenden RCA-1-Schritt tritt jedoch folgendes

²⁰Hier findet ein AlSi-Target (Al mit 1 % Si) Verwendung. Aufgedampft reines Aluminium führte zu den gleichen Ergebnissen.

²¹HF; oft wird noch zusätzlich H₂O₂ beigemischt, auch bekannt als Titanium Etchant

²²entsteht beim Kochen von 3 Teilen Salzsäure und 1 Teil Salpetersäure, bis die Flüssigkeit und der Dampf erkennbar eine gelbe Färbung zeigen (dauert etwa 1 Stunde)

²³Kaliumhydroxid-Lösung

²⁴Tetramethylammoniumhydroxid-Lösung

²⁵Das sind spezielle Fotolacke, die sowohl positiv als auch negativ betrieben werden können. Damit die Chrommaske negativ abgebildet wird, werden sie nach der ersten Belichtung noch einmal gebacken. Danach folgt eine zweite Belichtung, die Umkehrbelichtung. So behandelt verkettet die Polymermatrix sehr stark und die Lacke zeigen sich als chemisch besonders resistent.

Problem auf: Obwohl Aluminium von RCA-1 nur sehr langsam angegriffen wird, erhöht sich die Ätzrate, aufgrund elektrochemischer Effekte, wenn das Aluminium auf einer TiN-Schicht liegt. Die Ätzrate des Aluminiums nimmt in diesem Fall noch zu, sobald das Titannitrid geätzt ist, sodass die Strukturen freigestellt sind und isoliert auf der Passivierung stehen. Das macht insbesondere Schwierigkeiten in Verbindung mit stark topologisch strukturierten Substraten. Hier ist ein Überätzen notwendig, um auch in Kavitäten genügend lange zu ätzen. Ein Beispiel sieht man in Abbildung 6.27. Das Aluminium der Metallisierung ist bereits sehr weit überätzt, während die goldene TiN-Schicht maßhaltig übertragen ist. Es ist zu betonen, dass die Überätzungen des Aluminiums dieser Probe nicht von einem zu langen PNA-Prozessschritt stammen und sich ebenso wenig der Fotolack vorzeitig abgelöst hat.

Um diesen Effekt zu verhindern, muss das Aluminium passiviert werden. Für den basischen RCA-1-Clean bietet sich Kieselsäure²⁶ an. Sie bildet zusammen mit Aluminium im alkalischen Milieu Aluminiumsilikate, die in diesem Milieu beständig sind. Durch den Zusatz von Kieselsäure kann somit eine Aluminium-Hardmask verwendet werden. Sie ist aufgrund der Aluminiumsilikat-Passivierung in der Ätzlösung beständig.

Versuche zeigten, dass die passivierende Eigenschaften der Kieselsäure schon ab einer Konzentration von 1 g/l einsetzt. Vollständig wird das Aluminium jedoch erst ab 2 g/l geschützt. Um sicher eine Silikatbildung zu erhalten, ohne dass durch eventuelle lokale Konzentrationsverringerungen, z. B. in Kavitäten, diese Mengen unterschritten wird, werden für diese Arbeit 3 g/l als Standard gewählt.

Mit dieser Passivierung gewinnt man zusätzlich den Vorteil, das bereits strukturierte Aluminium selbst als Hardmask für das Ätzen der restlichen Metallisierung verwenden zu können. Damit wird man wieder unabhängiger in der Wahl des Fotolackes. Dieser muss lediglich der PNA-Lösung standhalten. Es sind demnach auch Positivlacke, wie der AR-P 3840 einsetzbar, die eine kleinere Strukturauflösung und einfachere Lithografieverfahren erlauben. Außerdem kann die Prozessierungszeit verkürzt werden. Die Ätzlösung kann über die 35 °C hinaus erwärmt werden, weil kein Fotolack während der Strukturierung des Titannitrids dem RCA-1 standhalten muss.

Mit einer Lösung entsprechend Tabelle 6.8, im folgenden APK-Ätze (Ammoniak, WasserstoffPeroxid, Kieselsäure) genannt, und dem gleichen Ansatz ohne Kieselsäure wurden Ätzraten für Titan und Titannitrid in Abhängigkeit von der Temperatur bestimmt. Die Ergebnisse werden in Abbildung 6.25 miteinander verglichen. Deutlich zeigt sich die starke Temperaturabhängigkeit. Die Kieselsäure reduziert die Ätzraten sowohl für Titan als auch für Titannitrid. Die berechneten Ätzzeiten für Titan und Titannitrid mit und ohne Kieselsäure als Funktion der Temperatur sind in Abbildung 6.26 dargestellt. Die Zeiten beziehen sich auf das Ätzen des Titannitrids und des Titans der neu entwickelten Metallisierung, d. h. 100 nm bzw. 35 nm Schichtdicke. Auch in dieser Darstellung wird der Einfluss

²⁶ H_4SiO_4 (Monokieselsäure), sie polykondensiert zu Polykieselsäuren $\text{H}_{2n+2}\text{Si}_n\text{O}_{3n+1}$

Substanz	Menge
NH ₃ -Lösung (29%ig)	100 ml
H ₂ O ₂ -Lösung (30%ig)	200 ml
H ₂ O	700 ml
H ₄ SiO ₄ ^a	3 g/l ^b

^aleichte Kieselsäure (gefällt, reinst)

^bentspricht 50 ml Schüttvolumen

Tabelle 6.8: Zusammensetzung der APK-Ätze

der Kieselsäure auf die Ätzrate deutlich. Trotzdem ist eine Verringerung der Gesamtprozesszeit möglich. Wie oben angesprochen, ist bei der Verwendung von Kieselsäure kein Fotolack notwendig. Deshalb kann die Temperatur zur Steigerung der Ätzgeschwindigkeit weiter erhöht werden, sodass sich die Ätzzeit insgesamt gegenüber der anderen Variante verkürzt. Die Durchführung dieses Prozesses ist einfach und es genügen Bechergläser. Bei der Verwendung von Kieselsäure ist die Temperatur nicht kritisch, weil kein Fotolack verwendet wird. Das Überätzen ist ebenso wenig problematisch, da die Ätzlösung eine sehr hohe Selektivität zu Siliziumnitrid aufweist.

Der Unterschied zwischen zwei Metallisierungen, die mit dem RCA-1 (in Abbildung 6.28) bzw. mit der APK-Ätze (in Abbildung 6.29) strukturiert wurden, ist optisch leicht zu erkennen. Die Herstellungsbedingungen der dargestellten Proben waren identisch, bis auf die Zugabe von Kieselsäure in die Ätzlösung. Gut sichtbar ist in Abbildung 6.28 die Verkleinerung der lateralen Abmessungen des Aluminiums gegenüber denen des golden schimmernden Titannitrids. An den Rändern der Leiterbahnen in Abbildung 6.29 sind Abschrägungen zu erkennen. Diese stammen von dem isotropen Ätzverfahren für die Aluminiumstrukturierung. Es ist aber kein Titannitrid zu erkennen.

6.6.3 Kontaktwiderstände

Mit der entwickelten Metallisierung hergestellte Kelvinstrukturen [LLYL87] wurden mit Vierspitzenmessungen charakterisiert. Der Kontaktwiderstand ist streng von der Dotierkonzentration des kontaktierten Siliziums abhängig. Erst ab der Dotierhöhe von $9 \cdot 10^{19} \text{ cm}^{-3}$ entstehen echte ohmsche Kontakte. Die Spannung steigt bei kleineren Kontaktlöchern schneller als bei Größeren, weil der Widerstand wegen der größeren Fläche kleiner ist. Das durch Tempern entstehende Titansilizid erhöht den Leitwert. Trotzdem bleiben die Kontakte ohmisch. Dies bestätigt die Vierspitzenmessung in Abbildung 6.30; die Kurven der getemperten Proben verlaufen flacher. Die Messung der Kontakte auf kristallinem n-Silizium erfolgte jeweils vor und nach dem Formiergastempnen bei 450 °C. Die Kontaktlö-

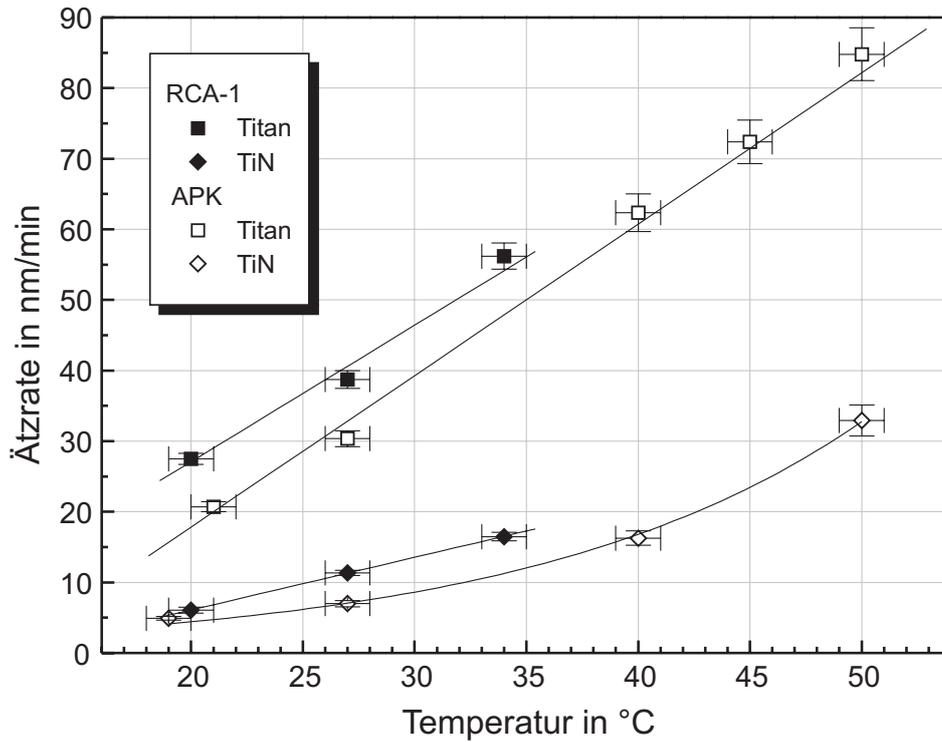


Abbildung 6.25: Ätzraten für Titan bzw. Titannitrid bezüglich RCA-1 (ohne Kieselsäure) und APK (mit Kieselsäure).

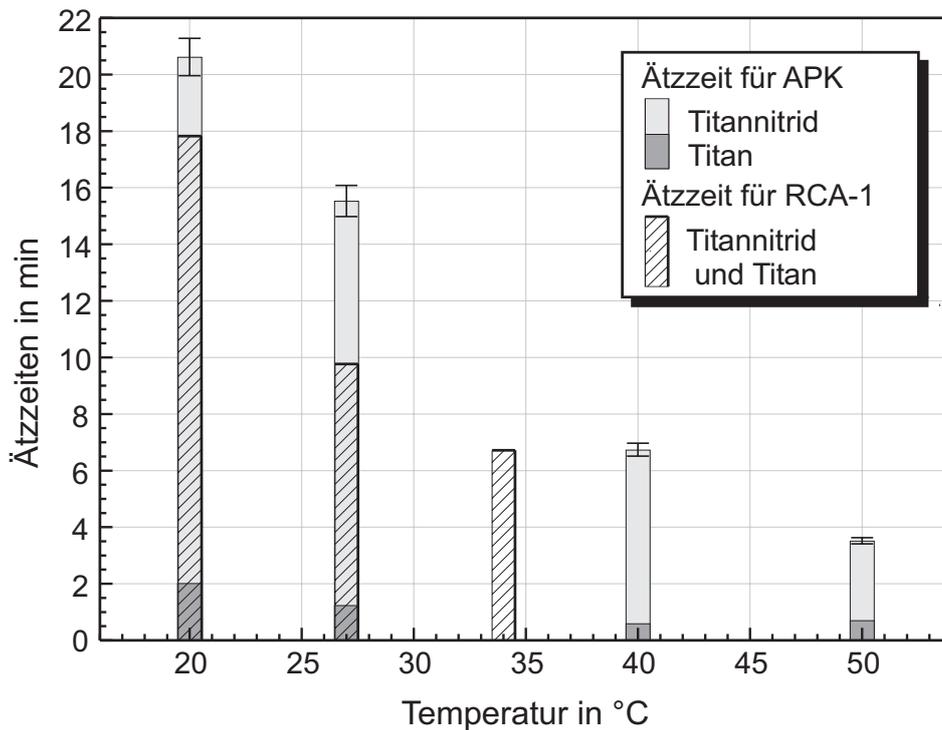


Abbildung 6.26: Ätzzeiten für das Titan/Titannitridsystem. Für das Ätzen mit RCA-1 (ohne Kieselsäure) wurde eine Fotolackmaske, für das Ätzen mit APK (mit Kieselsäure) wurde eine Aluminium-Hardmask mit Aluminiumsilikat-Passivierung verwendet.

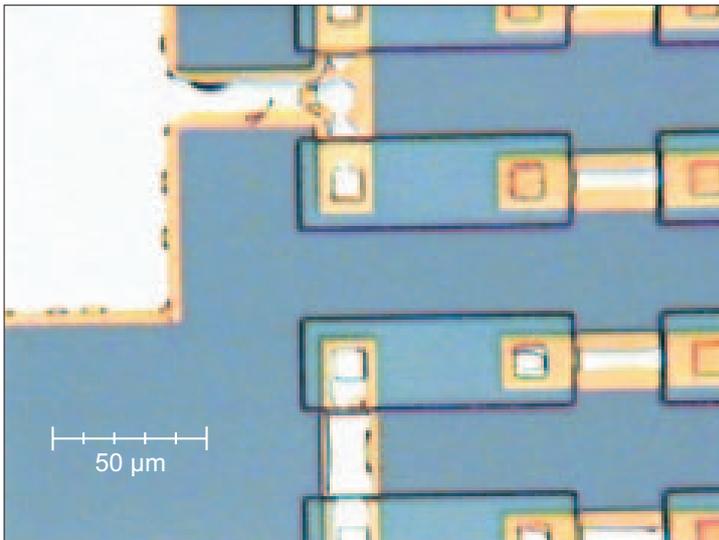


Abbildung 6.27:
mit RCA-1 überätzte Aluminiumstrukturen (weiß) auf Titan-
nitrid (golden) in Draufsicht. Aufgrund elektrochemischer Ef-
fekte erhöht sich die Ätzrate für Aluminium, wenn es sich
in Kontakt mit Titanitrid be-
findet.

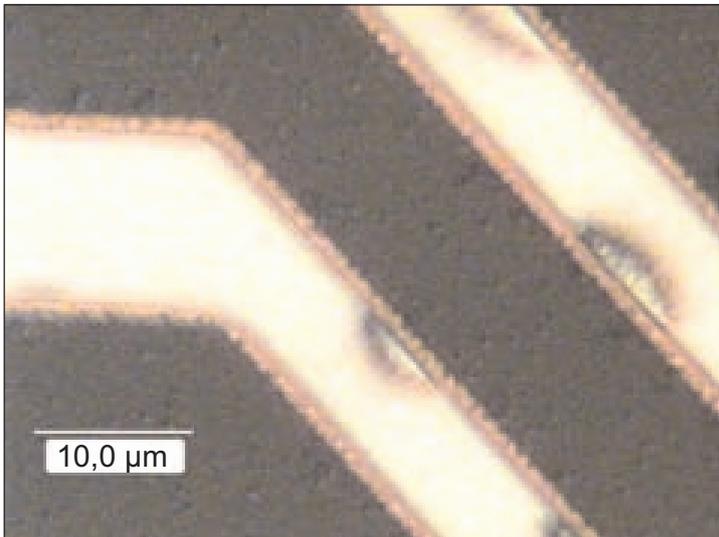


Abbildung 6.28:
die Metallisierung (in Drauf-
sicht) nach dem Ätzen mit RCA-
1 (ohne Kieselsäure). Die later-
alen Abmessungen der (weißen)
Aluminiumstrukturen sind ge-
genüber denen des (goldenen)
Titanitrids verkleinert.

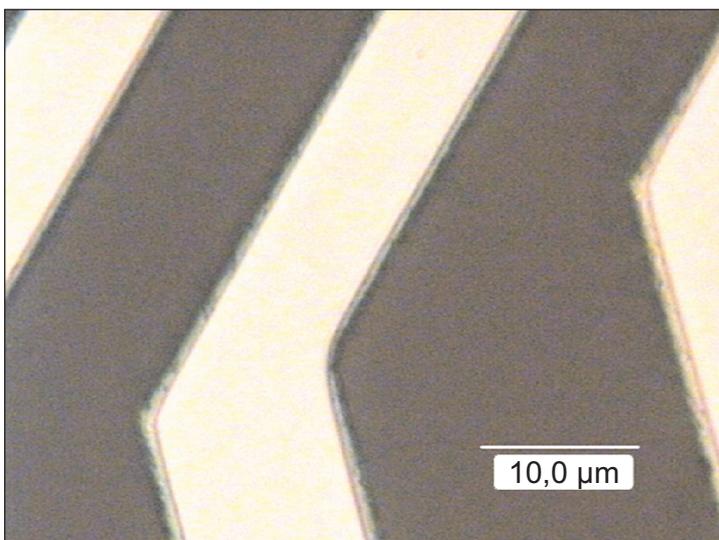


Abbildung 6.29:
die Metallisierung (in Drauf-
sicht) nach dem Ätzen mit APK
(mit Kieselsäure). Das Alumini-
um wird durch die Aluminiumsi-
likate passiviert und nicht mehr
von der Ätze angegriffen. Die
schrägen Kanten stammen vom
isotropen PNA-Ätzen.

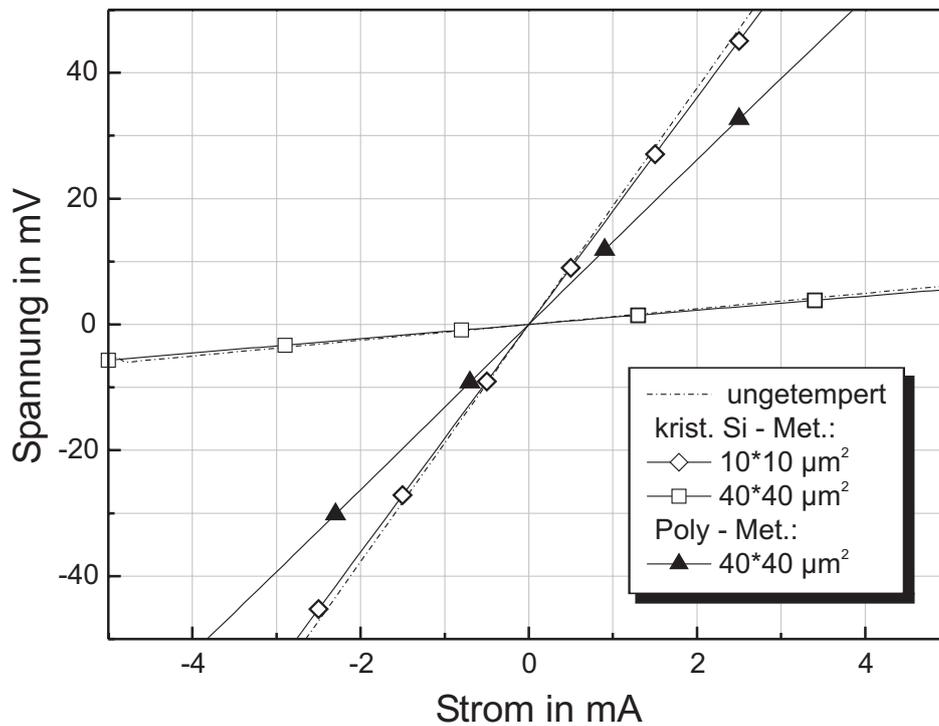


Abbildung 6.30: U-I Messung von Kelvinstrukturen zur Bestimmung der Kontaktwiderstände.

cher besitzen eine Fläche von $10 \cdot 10 \mu\text{m}^2$ bzw. $40 \cdot 40 \mu\text{m}^2$. Aus den Ergebnissen errechnet sich der spezifische Kontaktwiderstand zu $1,722 \pm 0,025 \cdot 10^{-5} \Omega\text{cm}^2$ für Kontakte auf kristallinem Silizium mit einer Dotierkonzentration von $9 \cdot 10^{19} \text{cm}^{-3}$. Die hergestellten Metall-Polysilizium-Kontakte sind im Vergleich hochohmiger. Für diese ergibt sich der spezifische Kontaktwiderstand zu $20,808 \pm 0,016 \cdot 10^{-5} \Omega\text{cm}^2$. Trotzdem zeigen auch diese Ohm'sches Verhalten. Die Messungen weichen in jedem Messpunkt weniger als 1% vom Mittelwert ab, was noch innerhalb der Messgenauigkeit des Parameteranalysers liegt. Das beweist die sehr gute Linearität der Strom-Spannungs-Charakteristik.

6.6.4 Möglichkeiten der weiteren Optimierung

Wie angesprochen, liegt das stöchiometrische Verhältnis der TiN-Schichten noch nicht bei eins. Durch Verändern der Bias-Spannung während des Sputterns könnte das Verhältnis noch weiter verbessert werden. Auch kann durch eine Modifikation an der Sputteranlage bei gleichzeitigem DC-Sputtern das Substrat im RF-Modus geätzt werden. Dadurch lässt sich die Bias-Spannung noch leichter variieren.

Sauerstoff findet sich nicht nur im Titanitrid sondern genauso im Titan. Auch in dieser Schicht verringert er den Leitwert. Die einfachste Möglichkeit ist, den Sauerstoff bzw. die Feuchte gar nicht erst in den Rezipienten zu lassen. Abhilfe könnte dazu eventuell ein Loadlock schaffen.

Ebenso muss das Rückspattern noch eingehender untersucht werden. Die Energie der beschleunigten Ionen hilft, die Feuchtigkeit schneller zu desorbieren. Es stellt sich die Frage, inwieweit die Plasmaschäden einer sauerstofffreien Legierung entgegen stehen.

6.7 Herstellung der Deckel

6.7.1 Ätzen der Klebelaschen

Als letzter Prozessschritt sollen der Deckel und der Chip mit einem Epoxidharzkleber verklebt werden. Dazu werden seitlich am Deckel kleine Vertiefungen geätzt, um den Klebstoff aufzunehmen. Deren Tiefe soll 30 μm betragen. Da als Substratmaterial für den Deckel Silizium verwendet wird, findet hier wieder ein RIE-Ätzprozess mit den Gasen Schwefelhexafluorid, Argon und Sauerstoff Verwendung. Die Ätzraten dieses Prozesses für Fotolack liegen bei 180 nm und für Silizium bei 1200 nm. Rechnerisch werden also für die Zeit, die für das Ätzen eines Grabens der gewünschten Tiefe benötigt wird, 4,5 μm Fotolack abgetragen. Mit den verwendeten Fotolacken ist eine solche Lackmaske nur schwer herstellbar (man vergleiche Abbildung 6.1). Darum wird hier eine Hardmask aus Aluminium verwendet. Zunächst wird das Metall aufgesputtert und mit PNA strukturiert. 300 nm stellen sich als ausreichend heraus. Nach dem RIE-Ätzen wird die Hardmask wieder nasschemisch vollständig entfernt.

Die Klebetröpfchen werden auf ein Feld von 500 μm x 1000 μm aufgetragen. Das Design ist für ein Gelingen ausschlaggebend. Dass die Vertiefungen als Einbuchtungen ausgeführt werden, hat zwei Gründe:

- Die Justierung der Wafersäge erfolgt, ähnlich wie bei der Fotolithografie, optisch, anhand der Strukturen, die in den Wafer eingearbeitet sind. Diese dürfen aber ein gewisses Maß nicht überschreiten, weil sie sonst nicht mehr in das Sichtfeld des Mikroskops passen. Auf dem Wafer liegen diese Deckel, wie Abbildung 6.31 zu sehen, dicht aneinander. Wenn die Klebelaschen nicht als Einbuchtungen ausgeführt sind, liegen die Kanten, an denen die Sägelinien ausgerichtet werden, über einen Millimeter auseinander. Mit Einbuchtungen wird dieses Problem umgangen.
- Außerdem treten an den 30 μm hohen Stufen beim weiteren Prozessieren der Deckel die in 6.4 beschriebenen Probleme bezüglich des Fotolacks und der Kanten in sogar noch gravierenderem Ausmaß auf. Zur Abhilfe müssen die Abstandhalter so weit wie möglich entfernt von der Kante der Vertiefung platziert werden, sodass die zu entwickelnden Strukturen nicht im Bereich des sich verdünnenden Fotolackes liegen. Auf der anderen Seite sollen sich diese entfernt vom sensitiven Bereich befinden. Auch das wird mit dem Design erreicht.

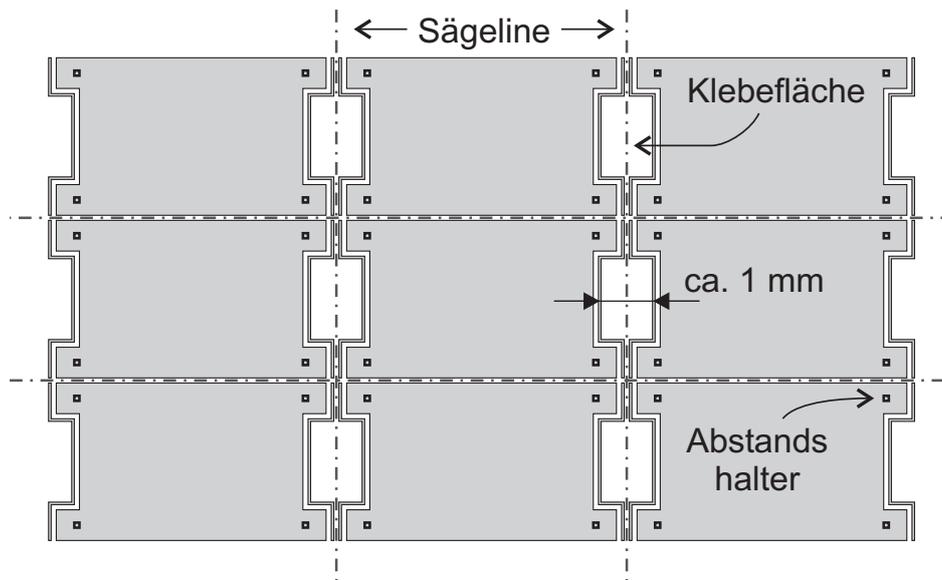


Abbildung 6.31: Anordnung der Deckel auf dem Wafer. An den eng aneinander liegenden Strukturen kann die Säge ausgerichtet werden.

6.7.2 Abstandhalter

Die Deckel tragen neben der sensitiven Schicht auch die Abstandhalter, die dafür sorgen, dass beim Zusammenfügen zwischen Schicht und Gate-Elektrode ein definierter Luftspalt bleibt, in den Gas eintreten kann. Für deren Herstellung werden 45 nm Titan als Haftvermittler und 3,6 μm Aluminium abgeschieden. Anschließend werden diese Schichten mittels Fotolithografie und Nasschemie strukturiert.

In Zukunft könnte anstatt des Aluminiums Kupfer verwendet werden. Das Edelmetall ist weit beständiger gegenüber Basen und Säuren. Vor allem widersteht es auch Flusssäure. Die ist notwendig, um das natürliche Oxid des Siliziums zu entfernen. Des Weiteren bräuchten damit die Deckel nicht mehr auf die Chips geklebt, sondern könnten gelötet werden. Mit dieser Technik ist es möglich, die Montage von Wafer zu Wafer durchzuführen [DV03], d. h. sie könnte noch vor dem Sägen stattfinden.

6.7.3 Design der Klebersperre

Die Verbindung zwischen dem Deckel und dem Chip wird mit einem Epoxidharzkleber hergestellt. Nachdem Klebetropfen auf den Deckel aufgebracht sind, wird er mit einem Flip-Chip-Bonder justiert und montiert. Beim Zusammenpressen kann es vorkommen, dass der noch flüssige Kleber in den Bereich des Luftspaltes gelangt. Um diesen Vorgang zu veranschaulichen, wurde ein Deckel auf einen Glas-träger geklebt, weil das durchsichtige Glas eine Draufsicht unter dem Mikroskop erlaubt. Der Abstand zwischen diesen beiden entspricht dem zwischen Deckel und Chip. In Abbildung 6.32 kann man den Weg des Klebers verfolgen. Berührt der Tropfen den Bereich des Luftspaltes, wird er in diesen durch den Kapillar-Effekt

hineingezogen. In diesem Fall ist der Sensor nicht mehr zu gebrauchen. Das kann mit Hilfe einer Klebersperre, wie in Abbildung 6.33 gezeigt ist, verhindert werden. Deren Funktionsweise soll kurz erläutert werden.

Wenn eine Flüssigkeit einen Festkörper berührt, wirken aufgrund der Adhäsions- und Kohäsionskräfte drei Randspannungen, die von den 3 Grenzflächen flüssig-fest, flüssig-gasförmig und fest-gasförmig stammen. Damit sich diese im Gleichgewicht befinden, bildet sich ein Randwinkel α zwischen der Festkörper- und Flüssigkeitsoberfläche im Berührungspunkt aus [GKV86]. Das ist z. B. an benetzenden Flüssigkeiten in Gläsern zu beobachten, bei denen die Oberfläche der Flüssigkeit nicht im rechten Winkel auf die Behälterwand trifft, sondern scheinbar an ihr hängt.

Befindet sich ein Flüssigkeitstropfen, in diesem Fall der Klebetropfen, im Luftspalt, so verspürt er von zwei Seiten her, von dem Deckel oben und dem Chip unten, Adhäsionskräfte. Damit sich die Randwinkel auf beiden Seiten einstellen können, muss die Flüssigkeit, wie in Abbildung 6.34 skizziert, ihre Oberfläche krümmen. Damit geht ein Druckanstieg

$$\Delta p = \frac{2\sigma}{r} \quad (6.17)$$

in der Flüssigkeit einher, der proportional zur Oberflächenspannung σ und umgekehrt proportional zum Krümmungsradius r ist [GKV86]. Je kleiner der Luftspalt ist, desto kleiner muss aus geometrischen Gründen dieser Krümmungsradius sein und desto stärker wird der Tropfen in den Spalt hineingezogen.

Damit so eine Auswirkung verhindert wird, wurde eine Klebersperre entwickelt und in den Deckel integriert: Trifft die Flüssigkeitsfront auf die Stufe, kann sich der Krümmungsradius r in erster Näherung verdoppeln (Abbildung 6.34 rechts), was den Druckanstieg Δp halbiert. Je nach Beschaffenheit der Kante kann aber die Oberfläche des flüssigen Klebers gleichzeitig konkav und konvex geformt sein, sodass sich die durch die Krümmungen hervorgerufenen Druckänderungen gegenseitig aufheben. Deshalb wurden vor dem sensitiven Bereich lang gestreckte Mesen platziert. Sie werden im gleichen Arbeitsschritt wie die Vertiefungen zur Aufnahme des Klebers aus dem Silizium herausgeätzt. Da nun der Druckanstieg in der Flüssigkeitsfront (hier des Klebers) senkrecht zur Stufe größer ist als derjenige an der Stufenkante selbst, breitet sich die Flüssigkeit zunächst nur auf der Stufe aus. Genau das demonstriert Abbildung 6.33. Die Klebetropfen berühren deutlich die Mesa, weil sehr viel Klebstoff verwendet wurde. Der Kleber fließt nur entlang der Klebersperre (erkennbar an der dunkleren Farbe), nicht aber in Richtung der sensitiven Fläche.

Die Versuche ergaben, dass 30 μm Vertiefungen ausreichen, um den Klebstoff aufzunehmen und trotzdem Chip und Deckel zu benetzen. Man vergleiche dazu wieder die Abbildung 6.32. Die Klebepunkte zeigen einen scharfen Farbübergang, was bedeutet, dass der Klebstoff das Glas berührt. Befände sich der Tropfen nur auf dem Silizium, zeigte sich an diesen Stellen ein Farbverlauf.

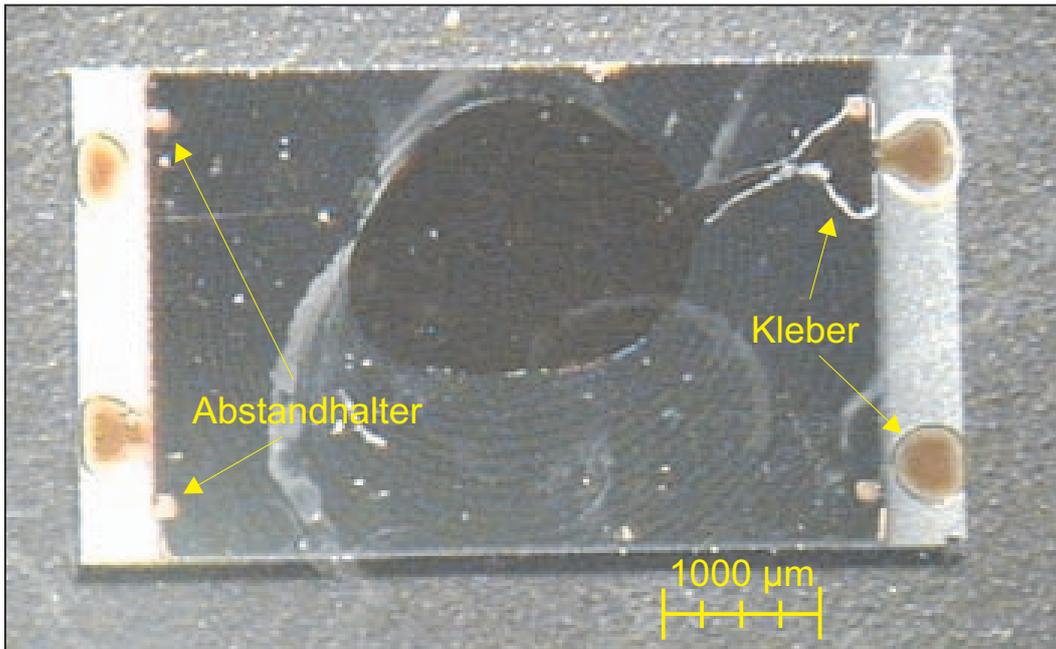


Abbildung 6.32: Der Klebetropfen berührt beim Zusammenpressen den Bereich des Luftspaltes. Durch den Kapillar-Effekt wird er immer weiter hineingezogen und sammelt sich im sensitiven Bereich.

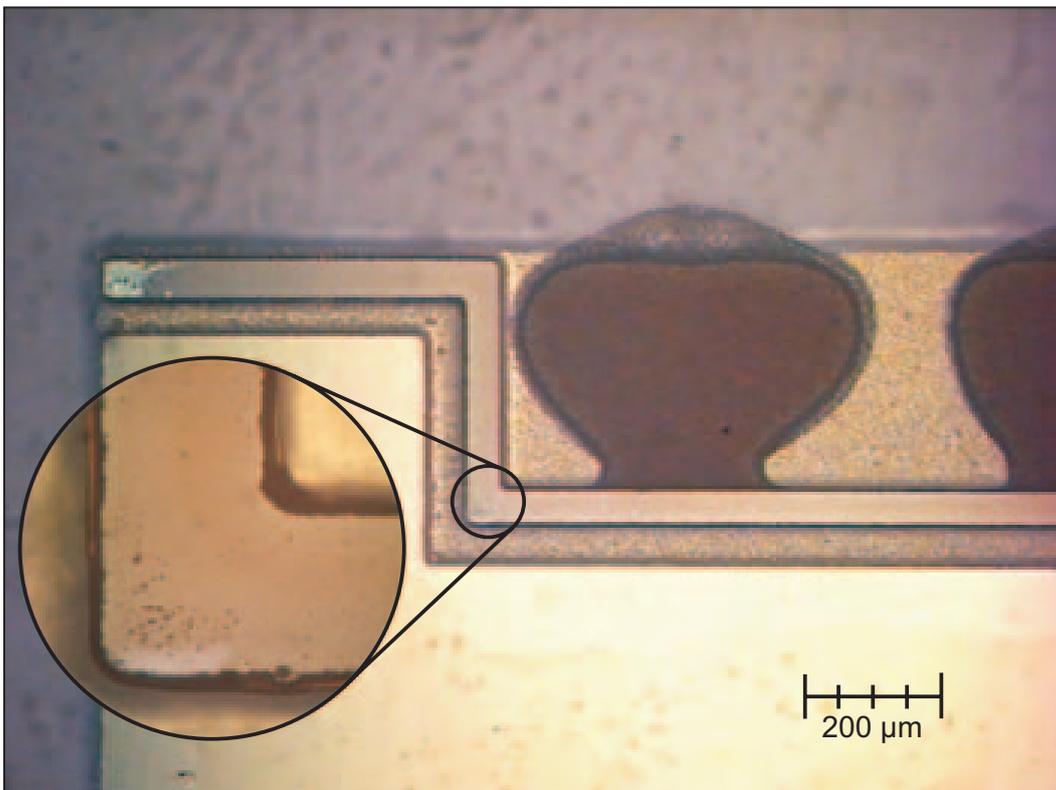


Abbildung 6.33: Mit einer Klebersperre kriecht der 2-Komponenten-Kleber nicht in den sensitiven Bereich.

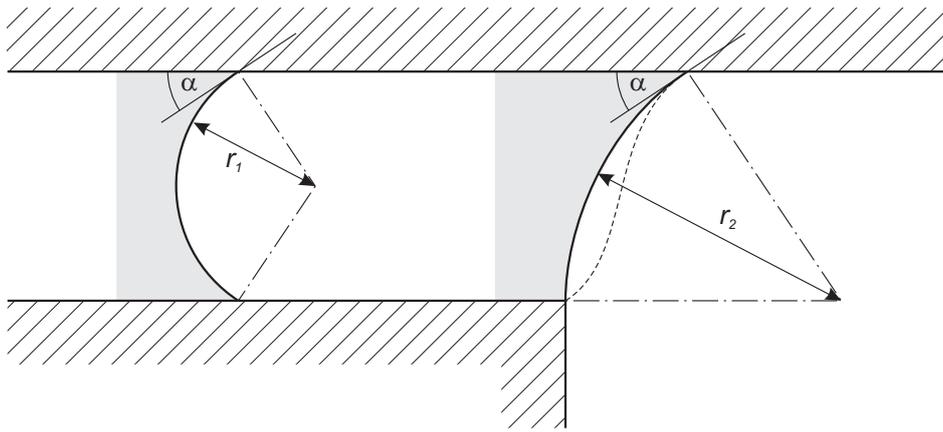


Abbildung 6.34: Flüssiger Kleber im Luftspalt

6.7.4 Elektrische Kontaktierung des Deckels

Bei den bisher aufgebauten Sensoren wurde üblicherweise direkt auf den Deckel gebondet, um die sensitive Schicht zu kontaktieren. Das setzt ein elektrisch leitendes und bondbares Material als Träger der sensitiven Schicht voraus. Gewöhnlich wurde dazu bisher Silizium als Substratmaterial verwendet. Um den Sensor noch flexibler einsetzen zu können, wurde die Möglichkeit untersucht, die sensitive Schicht über die Abstandhalter zu kontaktieren und schließlich über ein Bondpad auf dem Chip anzuschließen. So wäre es denkbar ein Glassubstrat als Träger der sensitiven Schicht heranzuziehen. Aus Kostengründen wurde auch schon das billige Aluminiumoxid als Material angedacht und getestet [SPL⁺04].

Dazu wurde ein Sensor fertig aufgebaut. Jeweils zwei Abstandhalter aus Aluminium stehen auf zwei metallisierten Flächen, die je mit einem Bondpad verbunden sind. Der Sensor wurde hierfür so präpariert, dass ein Stromfluss ausschließlich von einem Bondpad über die Abstandhalter, durch die sensitive Fläche und wieder über den Abstandhalter zum anderen Pad möglich ist. Anschließend wurde mit einer Vierspitzenmessung die U - I -Kennlinie aufgenommen.

Das Ergebnis wird in Abbildung 6.35 wiedergegeben. Es lässt keinen reinen ohmschen Kontakt erkennen. Immerhin ist er nahezu linear. Der Kontaktwiderstand errechnet sich zu $9,5 \pm 0,5 \Omega$. Es reicht aber durchaus, um die sensitive Schicht auf ein vorgegebenes Potenzial zu legen. Um die Einsatztauglichkeit zu bestätigen, müssen noch weitere Langzeitversuche gemacht werden.

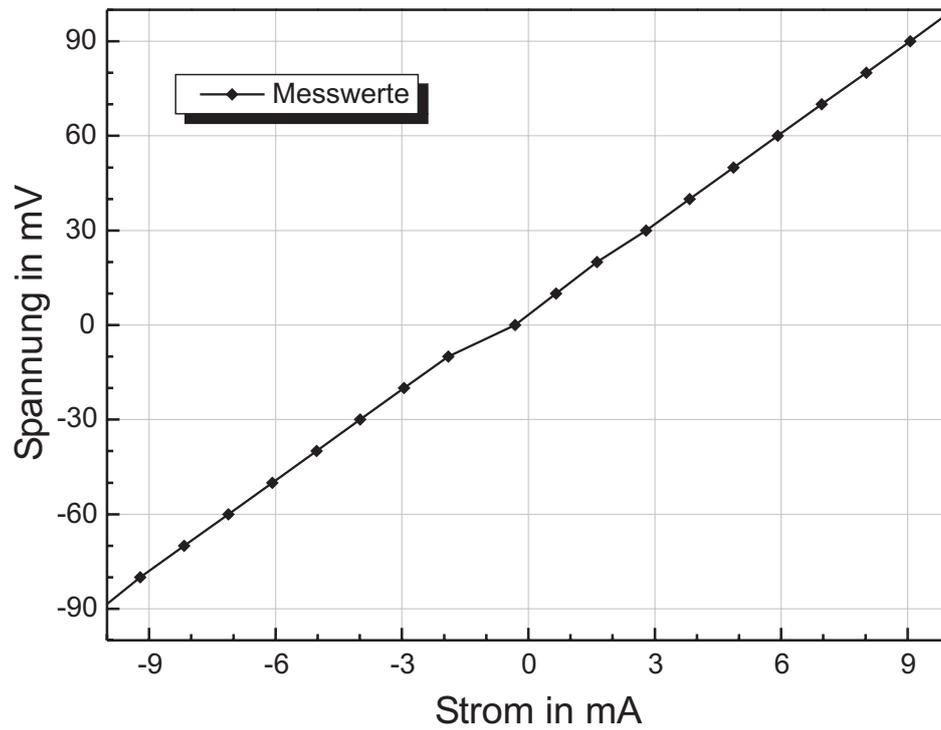


Abbildung 6.35: Elektrische Kontaktierung des Deckels über die Abstandhalter

7 Charakterisierung

Entsprechend der in Kapitel 6 entwickelten Technologie werden Chip und Deckel gefertigt. Damit die beiden Teile miteinander verklebt werden können, werden sie vereinzelt. Das geschieht in einer Wafer-Säge. Um die Sensoren leichter während der Messungen handhaben zu können, fixiert man sie in einem TO-8 Gehäuse. Als Letztes werden die Pins dieses Sockels mit den Bondpads unter Verwendung eines Ultraschallbonders kontaktiert. In diesem Status befindet sich der Sensor der Abbildung 7.1.

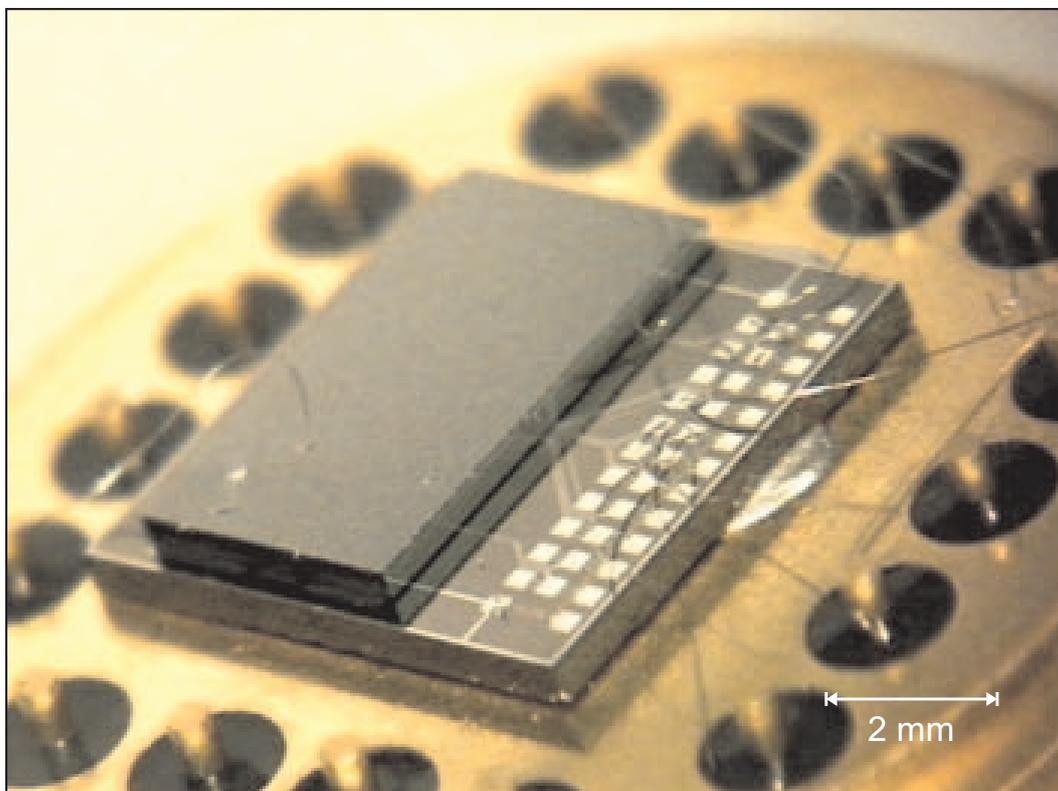


Abbildung 7.1: Sensor ist mit Leitsilber auf einen TO-8 Sockel geklebt, um ihn leichter handhaben und kontaktieren zu können.

7.1 Charakterisierung der Transistoren

Die drei verschiedenen Typen der Transistoren auf den Chips sollen vermessen und charakterisiert werden. Daraus lässt sich ableiten, welcher Typ der Transistoren am besten für den Einsatz im Sensor geeignet ist. Dazu wurden die unterschiedlichen

Transistoren vermessen. Die Abbildungen 7.2 bis 7.4 zeigen jeweils die Transfer- und Ausgangskennlinien der drei Typen.

7.1.1 Transfer- und Ausgangskennlinienfeld

Offensichtlich ist die Steuerungswirkung der Gate-Spannung U_G . Die Sperrströme entnimmt man dem Bereich der Transferkennlinie, in dem der Transistor noch ausgeschaltet ist, hier z. B. zwischen 0 und 1 V. Im Vergleich hat der Transistor Typ 2 den geringsten Sperrstrom, gefolgt von Transistor Typ 1. Diese Reihenfolge entspricht der Erwarteten. Ein vertikaler n-Kanal-Transistor ist aufgrund seiner Bauform immer auch eine npn-Diode, die im Folgenden als Volumendiode bezeichnet wird. Deren Sperrstrom skaliert mit der Fläche der Mesa. Diese Volumendiode ist im Ersatzschaltbild dem Kanalwiderstand parallel geschaltet. Im angesprochenen Bereich der Kennlinien rührt der größte Anteil des Stromes eben von dieser Volumendiode her, weil der Transistor noch sperrt und so der Kanalwiderstand viel größer ist als derjenige der Volumendiode. Die Fläche der Mesa ist für den Transistor Typ 1 am kleinsten und für Transistor Typ 3 am größten.

Im eingeschalteten Zustand führt der Transistor Typ 1 den größten Strom I_{DS} . Dieser beträgt zirka das Doppelte des Stromes I_{DS} des Transistor Typs 2. Dieser Faktor entspricht auch ungefähr dem Verhältnis der jeweiligen Kanalweiten von 110 μm zu 50 μm . Da I_{DS} proportional zur Kanalweite W ist (siehe Gleichung (3.4)) bestätigt dieses Ergebnis auch die Theorie. Im Gegensatz dazu ist der Ausgangsstrom des Transistors Typs 3 am geringsten, obwohl die Kanalweite mit nominell 400 μm vom Design her am größten ist. Das Ausgangskennlinienfeld bleibt für alle Gate-Spannungen U_{GS} unterhalb von einem Milliampere. Bei diesem Transistor sind die Ausgangskennlinien (Abbildung 7.4) für 7 und 8 V Gate-Spannung U_G fast identisch. Auch aus der Transfercharakteristik ist kein Anstieg des Stromes I_{DS} ab 7 V Gate-Spannung U_{GS} mehr zu erkennen. Das deutet auf eine Begrenzung des Stromes I_{DS} nicht durch den Kanalwiderstand selbst, sondern durch den Zuleitungswiderstand hin.

7.1.2 Zuleitungswiderstand

Diese Größe kann aus der I - U -Messung der Volumendioden errechnet werden. Dazu wird der Drain-Strom I_{DS} in Abhängigkeit der Drain-Source-Spannung U_{DS} gemessen, während an Gate und Source konstant das Massepotenzial angelegt wird. Dabei spielt der Transistor keine Rolle, der Strom fließt durch das Volumen der Mesa. Der Zuleitungswiderstand liegt mit der Volumendiode in Serie. Mit ansteigender Spannung verringert sich der Widerstand der Volumendiode, sodass näherungsweise nur noch der Zuleitungswiderstand messbar ist. Dementsprechend entnimmt man dem Messgraphen diejenige Steigung, gegen welche die Kennlinie für große Spannungswerte strebt. Diese entspricht dem Leitwert der Zuleitung. Abbildung 7.5 visualisiert die Messungen und das Vorgehen. Die sich daraus ergebenden Werte sind in Tabelle 7.1 noch einmal zusammengefasst.

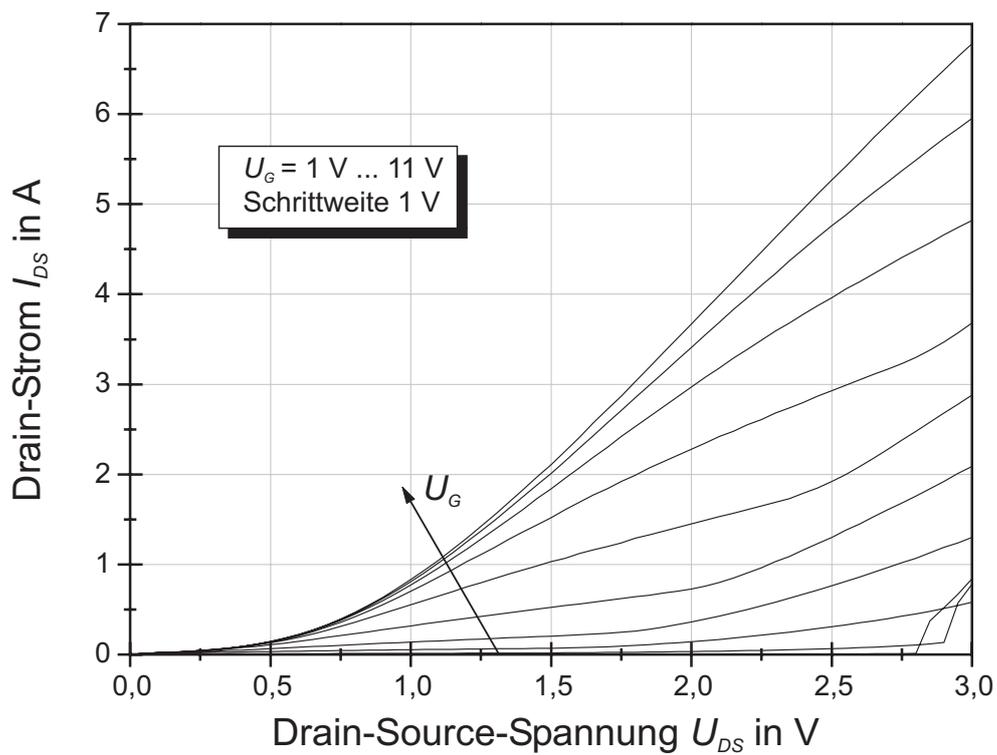
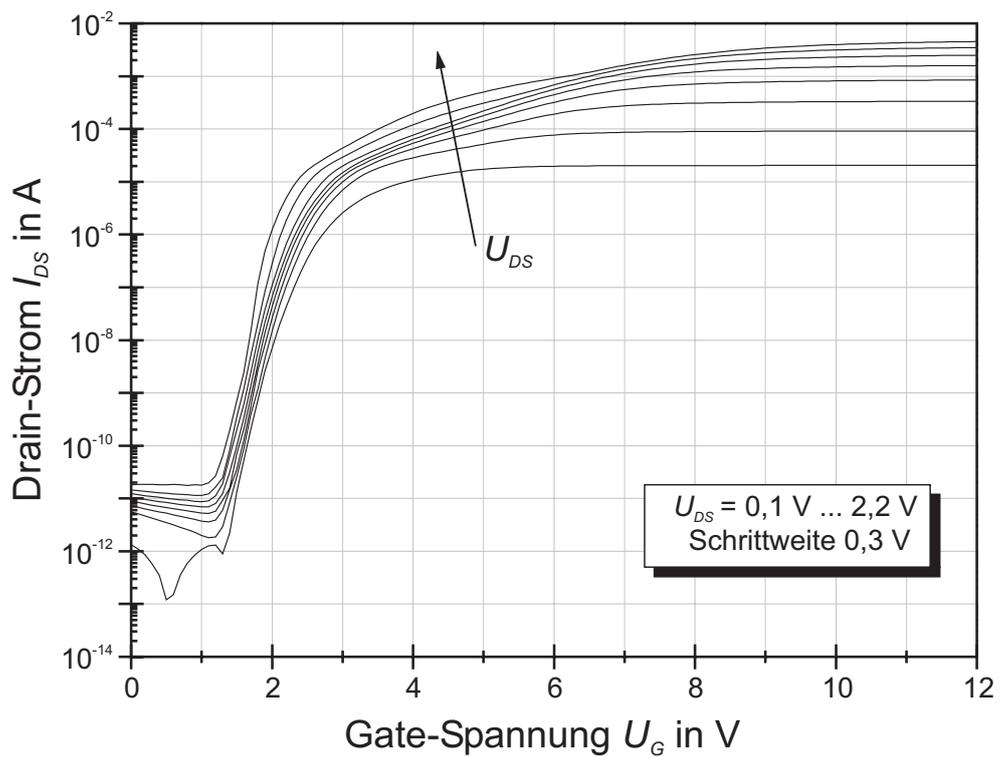


Abbildung 7.2: Transfer- (oben) und Ausgangskennlinie (unten) des Transistors Typ 1

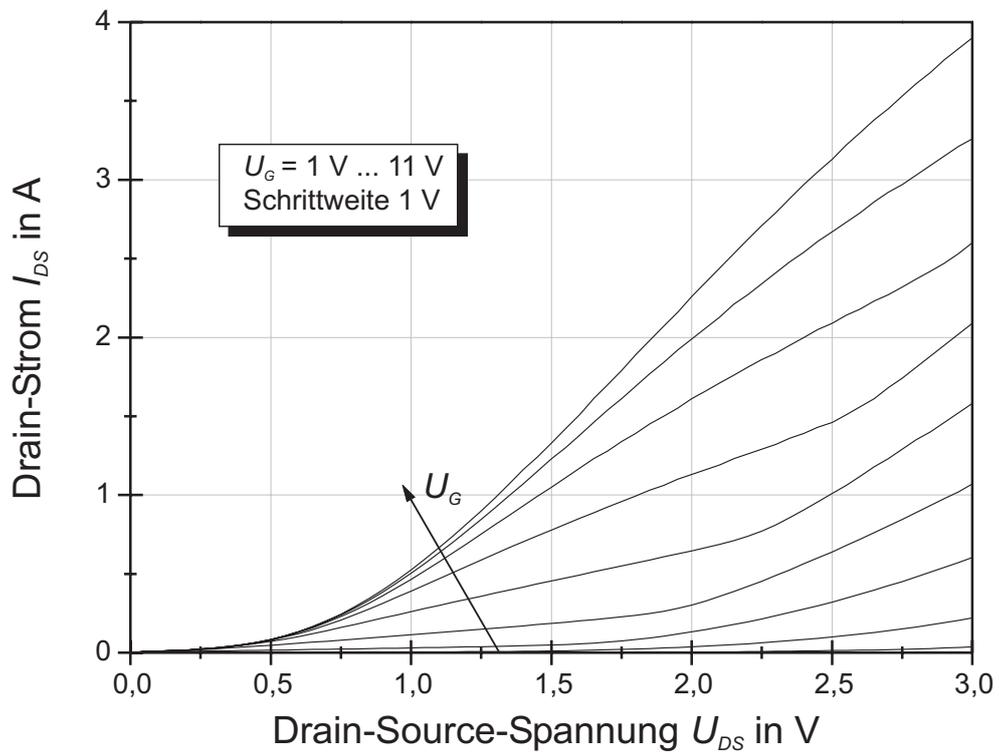
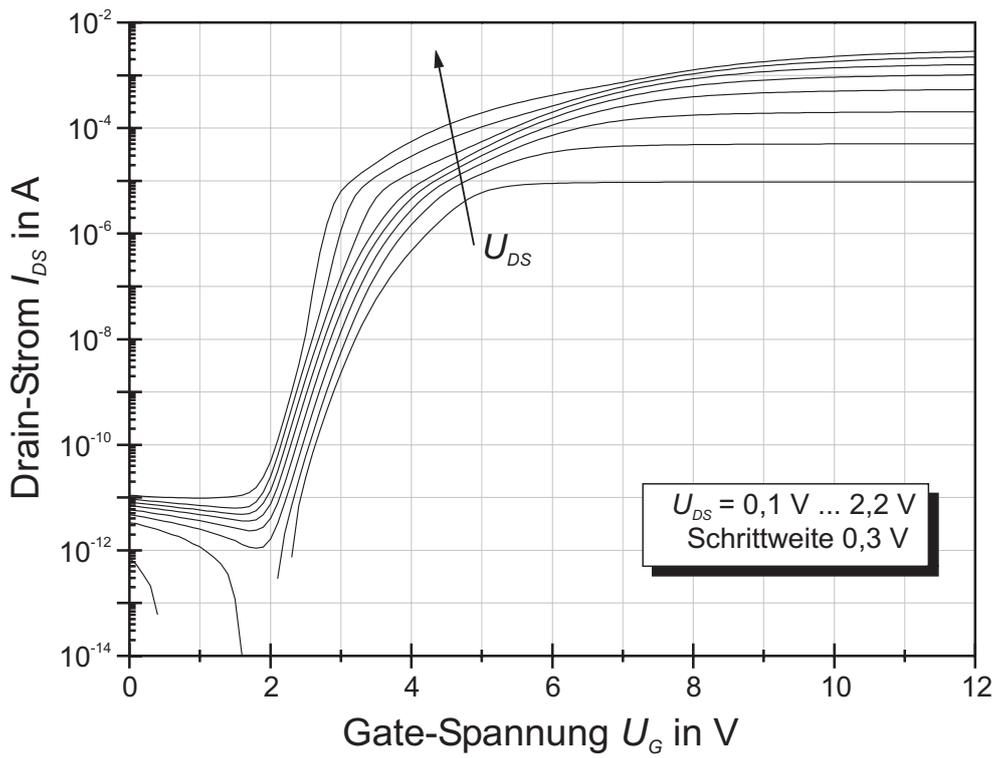


Abbildung 7.3: Transfer- (oben) und Ausgangskennlinie (unten) des Transistors Typ 2

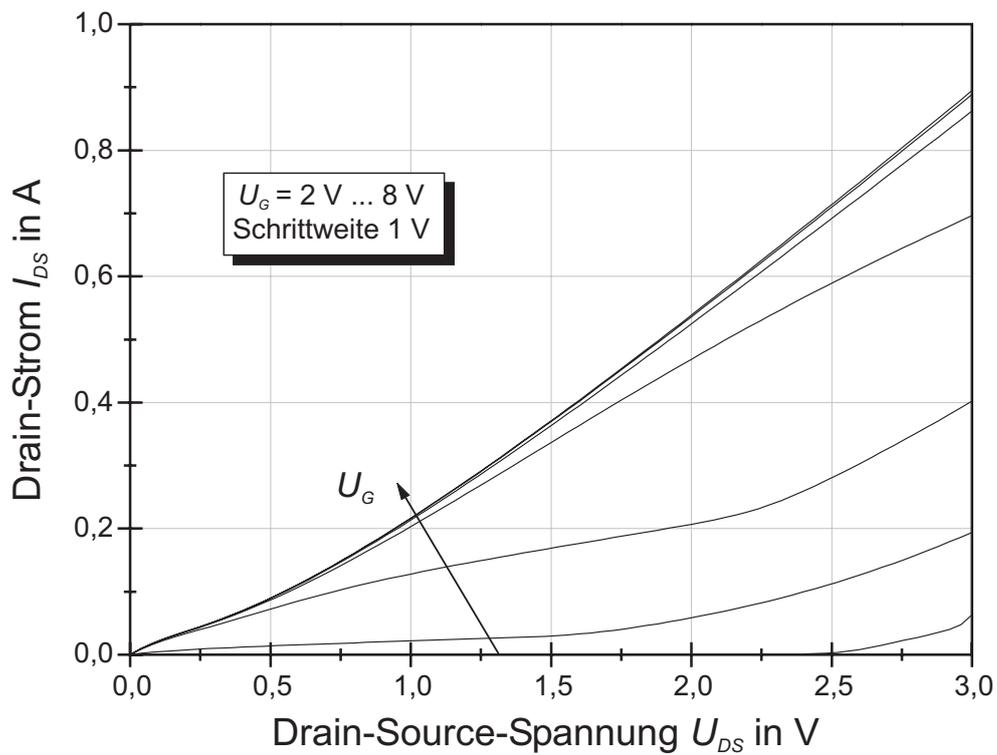
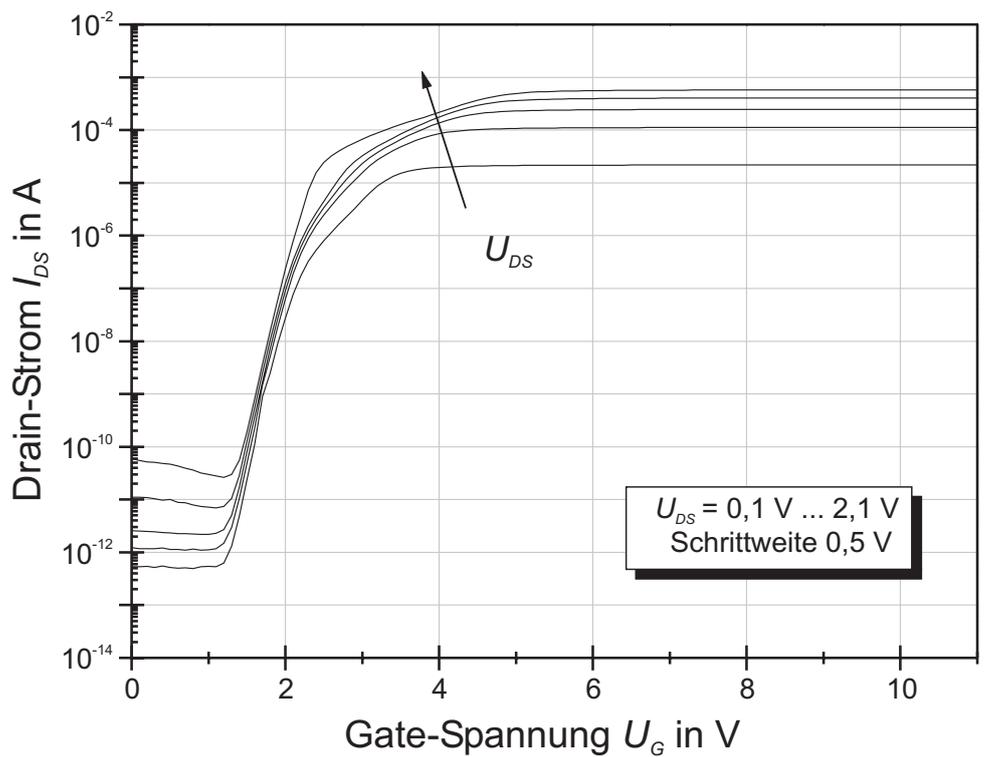


Abbildung 7.4: Transfer- (oben) und Ausgangskennlinie (unten) des Transistors Typ 3

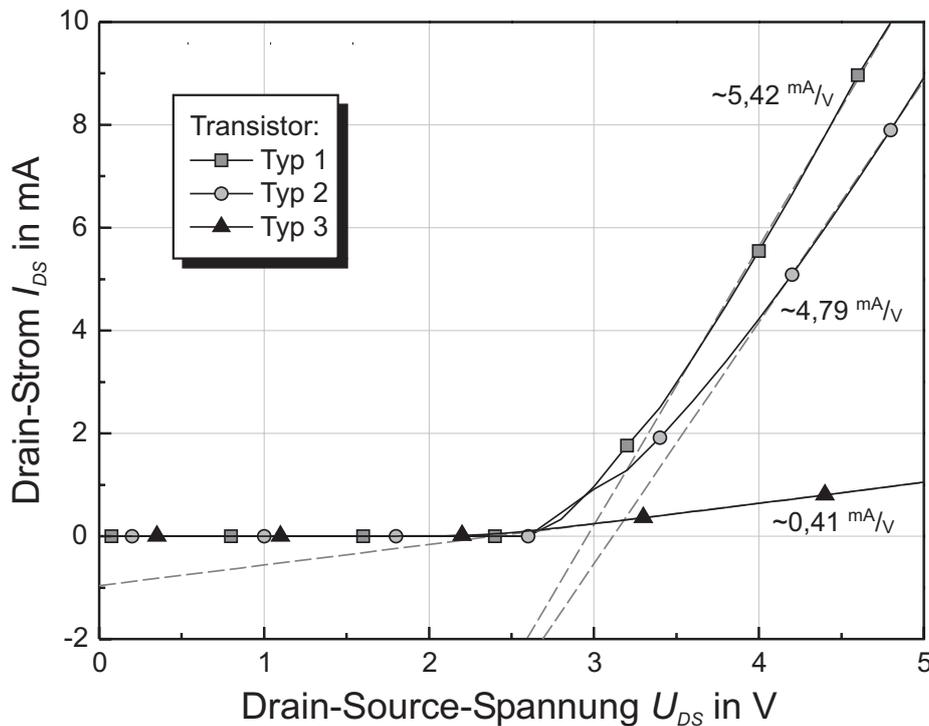


Abbildung 7.5: I-U-Kennlinien der Volumendioden

Die Ergebnisse bestätigen die oben geäußerte Vermutung des hohen Zuleitungswiderstandes für den Transistor Typ 3. Dieser Transistor ist mit einer zusätzlichen Verdrahtungsebene realisiert (siehe Abbildung 4.17). Diese ist notwendig, weil sich die Zuleitungen für Source und Drain einmal kreuzen. Der Kreuzungspunkt ist mit einer Aluminiumleiterbahn über einer Schicht Polysilizium ausgeführt. Dazwischen befindet sich Siliziumoxid und -nitrid. Der spezifische Widerstand ρ_{Poly} für das Polysilizium wurde mit ca. 20 m Ω cm gemessen.

Das führt mit der verwendeten Geometrie (Länge $l = 40 \mu\text{m}$, Breite $b = 30 \mu\text{m}$ und Schichtdicke $h = 250 \text{nm}$) gemäß

$$R = \rho_{Poly} \frac{l}{bh} \quad (7.1)$$

zu einem Widerstand von ungefähr 1000 Ω . Der Hauptanteil des Zuleitungswiderstandes des Transistors Typ 3 ist auf die vergrabene Verdrahtungsebene zurückzuführen. Am Kreuzungspunkt des Transistors Typ 3 befinden sich zusätzlich zwei Polysilizium-Metallkontakte in der Zuleitung. Diese sind, wie schon im Abschnitt 6.6.3 erörtert, überdies größer als die zwischen Metallisierung und kristallinem Silizium.

Der etwas geringere Zuleitungswiderstand des größeren Typs 1 ist mit den ebenso größeren Kontaktlöchern zu erklären, weil der Kontaktlochwiderstand mit der Fläche skaliert. Je größer die Fläche, desto kleiner der Widerstand.

Transistor	Leitwert in mS	Widerstand in Ω
Typ 1	$5,42 \pm 0,06$	$184,5 \pm 2,0$
Typ 2	$4,79 \pm 0,06$	$208,8 \pm 2,6$
Typ 3	$0,41 \pm 0,01$	$2442,5 \pm 1,8$

Tabelle 7.1: Zuleitungswiderstände der verschiedenen Transistoren

7.1.3 Volumendiode

Die Diodenmessungen in Abbildung 7.5 bietet auch eine Abschätzung, bis zu welcher Drain-Source-Spannung U_{DS} die Transistoren betrieben werden können. Ab dem Punkt $U_{DS} \approx 2,5 \text{ V}$ sperrt die Volumendiode nicht mehr ausreichend und ihr Anteil am Gesamtstrom nimmt stark zu. Dieser Anteil des Drain-Stroms I_{DS} kann nicht durch das Gate gesteuert werden. Aus diesem Grund ist es sinnvoll, die Transistoren unterhalb dieses Punktes zu betreiben.

7.1.4 Einsatzspannung

Die logarithmischen Transferkennlinien (siehe erneut Abbildungen 7.2 bis 7.4) zeigen, dass sich die Einsatzspannung mit ansteigender Drain-Spannung U_{DS} zu kleineren Werten hin verschiebt. Eine Bestimmung dieser Größe aus der logarithmischen Transferkennlinie über den Schnittpunkt der Tangenten an den Graphen im Unterschwellstrombereich und linearen Bereich bleibt aber ungenau. Die sonst übliche Extraktion der Einsatzspannung aus der $\sqrt{I_D}$ - U_{GS} -Darstellung [Hof90] ergab für diese Transistoren ebenso keine eindeutigen Werte. Einfacher ist es, einen geometrieunabhängigen Einsatzstrom festzulegen. Erreicht der Drain-Strom I_D diesen Wert, so gleicht die dazugehörige Gate-Spannung U_{GS} der Einsatzspannung U_T . Man definiert sie für n-Kanal-Transistoren wie folgt:

$$U_T = U_{GS} \quad \forall \quad I_D(U_{GS}) = \frac{W}{L} 8 \cdot 10^{-8} \text{ A} , \quad (7.2)$$

wobei L und W die Kanallänge bzw. -weite symbolisieren [Eis01]. Abbildung 7.6 erfasst die Einsatzspannungen für die drei Transistortypen nach Gleichung (7.2). Hieraus wird ein Sinken der Einsatzspannung U_T mit steigender Drain-Spannung U_{DS} noch deutlicher.

7.1.5 Steilheit

Da der Sensor letztendlich kleine Potenzialänderungen auslesen soll, stellt sich die Frage, in welchem Arbeitspunkt der Sensor den größten Signalhub, sprich, der Transistor die größte Änderung des Drain-Stroms I_{DS} mit Variation der Gate-Spannung U_{GS} liefert. Um günstige Einstellungen der Drain- und Source-Potenziale für die Sensoren zu finden, muss man die Steilheit g_m der Transistoren betrachten.

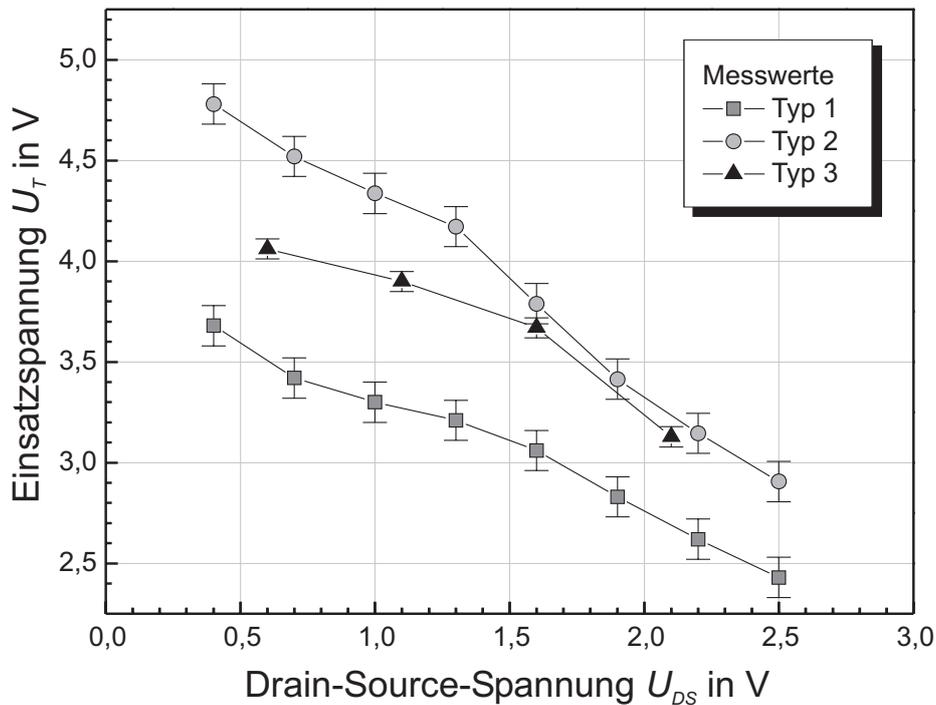


Abbildung 7.6: Einsatzspannungen der verschiedenen Transistoren

Man erhält sie durch Differenzieren der Transferkennlinie nach der Gate-Spannung U_{GS} :

$$g_m = \left. \frac{\delta I_{DS}}{\delta U_{GS}} \right|_{U_{DS}=\text{const.}} \quad (7.3)$$

Sie ist ein Maß dafür, wie gut die Gate-Spannung U_{GS} den Drain-Strom I_{DS} zu steuern vermag.

Die Berechnungen nach Gleichung (7.3) finden sich in der Abbildung 7.7 wieder. Für die Transistoren Typ 1 und 2 bewegt sich das Maximum der Steilheit zwischen 6 und 8 V. Die Kurven ähneln sich sehr, nur der Betrag ist wieder bei ersterem Transistor knapp das Doppelte des Zweiten. Das ist erneut über das W/L -Verhältnis zu erklären, weil es in die Stromverstärkung und damit auch in die Steilheit eingeht (siehe Gleichung (3.12)). Deshalb müsste der Transistor Typ 3 mit der größten Kanalweite die höchste Steilheit aufweisen. Die Maxima fallen für diesen jedoch am geringsten aus. Sie liegen um 4 V. Wieder verhindert der hohe Zuleitungswiderstand die theoretisch möglichen Werte.

Der Arbeitspunkt des Sensors ist nun so einzustellen, dass sich der Transistor in diesem Maximum befindet. Das bedeutet die Gate-Spannung U_{GS} sollte bei den Transistoren Typen 1 und 2 zwischen 6 und 8 V und bei Transistor Typ 3 um die 4 V liegen.

Da jedoch jeder Potenzialunterschied im Bereich des kapazitiven Spannungsteilers zu feuchteinduzierter Ladungsverschiebung auf der Oberfläche führt, ist es daher besser die Gate-Spannung U_{GS} auf Masse zu legen und die Source-

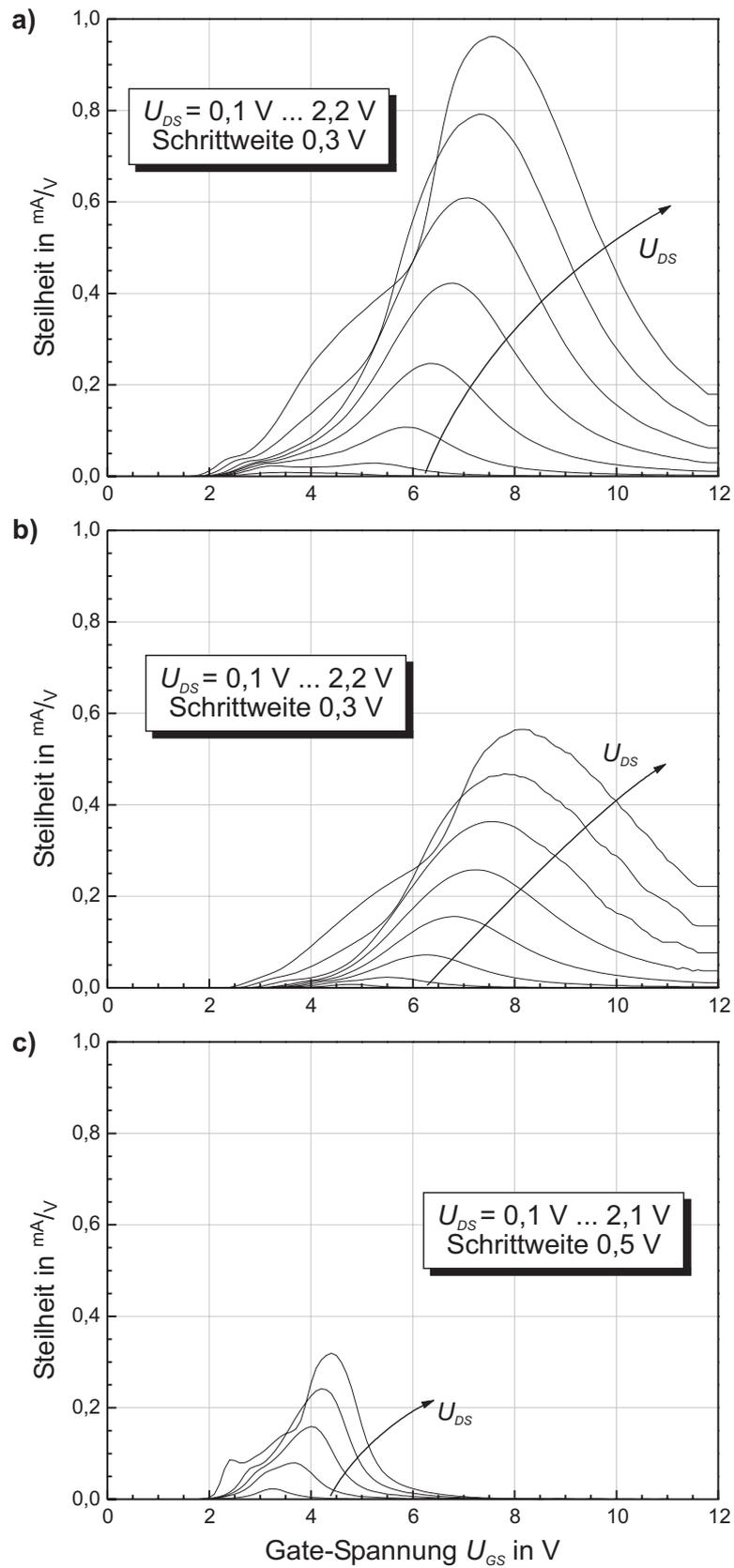


Abbildung 7.7: Steilheiten der verschiedenen Transistoren: a) Transistor Typ1; b) Transistor Typ2; c) Transistor Typ3

Spannung entsprechend negativ zwischen -6 und -8 V bzw. -4 V zu wählen. Das Drain-Potenzial ist um U_{DS} größer, aber ebenfalls negativ einzustellen.

7.1.6 Unterschwellsteigung

Eine weitere berechenbare Größe, welche die Qualität eines Transistors beschreibt, ist die Unterschwellsteigung S (engl. Subthreshold Swing), auch Parameter S genannt (nicht zu verwechseln mit den S -Parametern der Mehrorttheorie!). Das ist ein Maß dafür, wie schnell das Bauelement einschaltet. Man definiert ihn als Kehrwert der 1. Ableitung der logarithmischen Transferkennlinie:

$$S = \left(\frac{\delta \log I_{DS}(U_{GS})}{\delta U_{GS}} \right)^{-1} \Big|_{U_{DS} = const.} \quad (7.4)$$

Dieser Parameter gibt die Spannungsänderung ΔU_{GS} an, die, an der Gate-Elektrode beaufschlagt, den Drain-Strom I_{DS} um eine Dekade steigert. Analytisch ist S über

$$S = \ln(10) \frac{nkT}{q} \quad \text{mit} \quad (7.5)$$

$$n = 1 + \frac{C_i}{C_j} \quad (7.6)$$

gegeben. Im Extremfall, gleichzeitig auch das Optimum, ist die Kapazität des Isolators C_i viel kleiner als die Sperrschichtkapazität C_j , sodass S rechnerisch minimal $60 \text{ mV}/\text{Dek.}$ annehmen kann.

Dieser theoretische Wert für den Parameter S lässt sich mit diesen Transistoren nicht erreichen, weil die Sperrschichtkapazität nicht vernachlässigt werden kann. Setzt man die Gleichung (3.9) in Gleichung (7.6), so ergibt sich nach Umformen:

$$n = 1 + \frac{\varepsilon_{Si} \cdot d_{Ox}}{\varepsilon_{Ox} \cdot x_D} \quad (7.7)$$

Die hier vermessenen Transistoren haben eine Oxiddicke d_{Ox} von $16,9$ nm. Die Dotierung N_A des p-Gebietes beträgt nominell $1 \cdot 10^{18} \text{ cm}^{-3}$. Damit lässt sich die Weite der Sperrschicht x_D nach [Sze69] zu

$$x_D \approx \sqrt{\frac{4\varepsilon_0 \varepsilon_k T \ln(N_A/n_i)}{q^2 N_A}} = 35 \text{ nm} \quad (7.8)$$

nähern. Setzt man diese Werte (zusammen mit $\varepsilon_{Si} = 11,9$; $\varepsilon_{Ox} = 3,9$; $kT/q = 26 \text{ mV}$) in die Gleichungen (7.5) und (7.7) ein, erhält man den Parameter $S = 148 \text{ mV}/\text{Dek.}$. Berechnungen nach Gleichung (7.4) wurden für alle drei Transistoren anhand der experimentellen Kennlinien (Abbildung 7.2-7.4) durchgeführt und sind in Abbildung 7.8 dargestellt. Die Kurven belegen die Richtigkeit der oben gemachten Abschätzung. Die nur geringfügig höheren gemessenen Ergebnisse sind auf Überlappkapazitäten zurückzuführen.

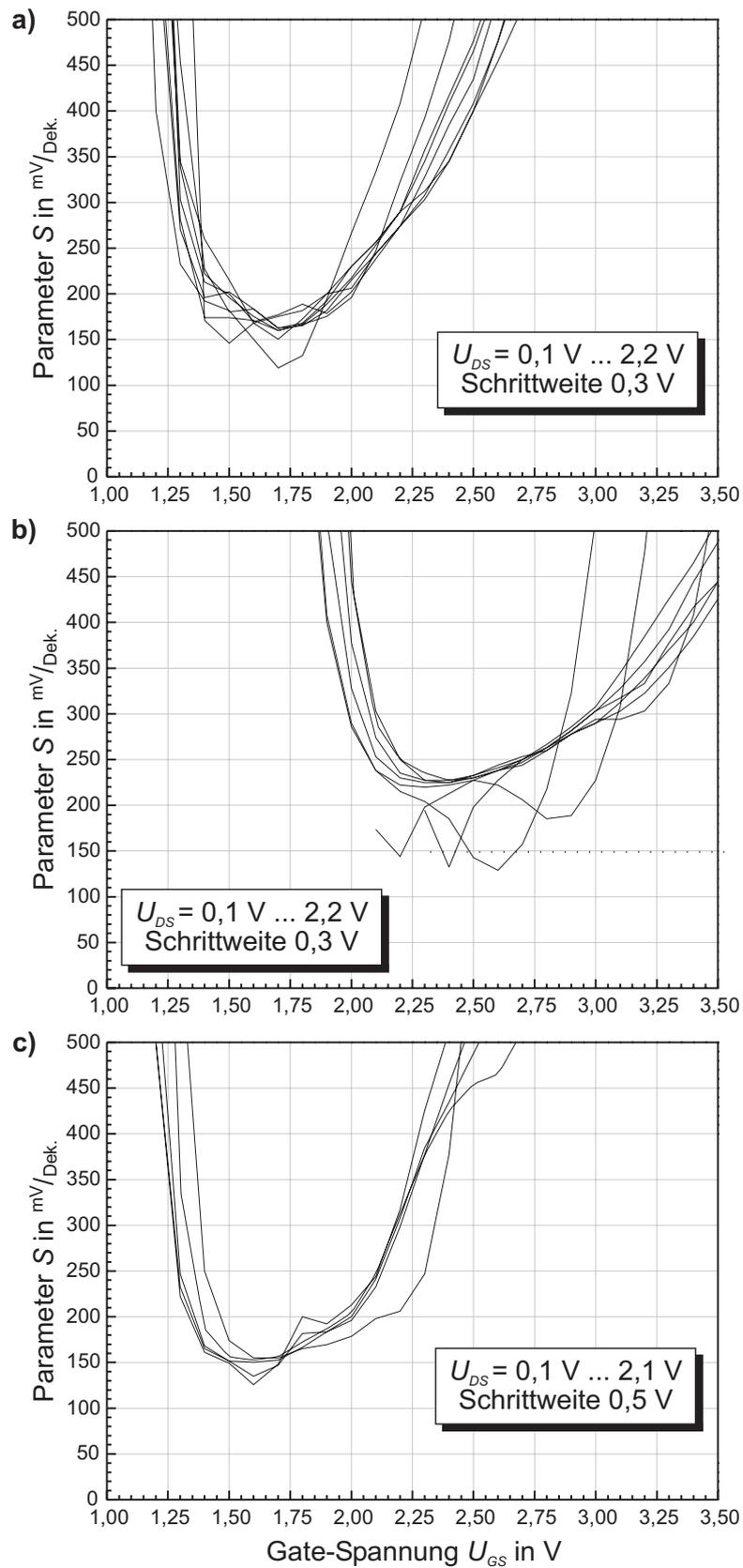


Abbildung 7.8: Parameter S (engl. Subthreshold Swing) der verschiedenen Transistoren: a) Transistor Typ1; b) Transistor Typ2; c) Transistor Typ3

7.2 Charakterisierung des Sensors

Der fertig prozessierte Sensor wurde charakterisiert und erste Messungen unter Gasbeaufschlagung durchgeführt. Untersucht wurde dazu ein Chip mit dem Transistor Typ 1, weil dieser, wie im vorherigen Abschnitt gezeigt, die günstigsten Eigenschaften aufwies. Dessen Kanalweite und -länge beträgt $110\ \mu\text{m}$ bzw. $200\ \text{nm}$. Somit besitzt der Transistor ein W/L -Verhältnis von 550.

Für die folgenden Erklärungen der Beschaltungen wird auf die Nummerierung aus Abbildung 4.1 zurückgegriffen.

7.2.1 Transferkennlinien

Zum Finden des Arbeitspunktes wird eine Transferkennlinie benötigt. Es stellt sich die Frage, wie groß die Potenzialdifferenz U_{GS} zwischen Drain und der Gate-Elektrode abzustimmen ist. Letztere ist im Falle des vollständig aufgebauten Sensors nicht mehr zugänglich. Trotzdem können Transferkennlinien aufgenommen werden. Dazu wird an beiden Seiten des Spannungsteilers, Eingang 1 und 2 aus Abbildung 4.1, $0\ \text{V}$ angelegt, denn so werden der Mittelabgriff und damit die floatende Gate-Elektrode auf das gleiche Potenzial festgelegt. Für die Aufnahme einer Transfercharakteristik fährt man nun die Spannungen U_{SG} Anschluss 4 und U_{DG} Anschluss 5 von $0\ \text{V}$ in den negativen Bereich und hält für jeweils eine Kennlinie die Drain-Source-Spannung U_{DS} konstant. Man nimmt dabei den Strom I_{DS} auf, der zwischen den Klemmen 4 und 5 fließt. Der Betrag der Gate-Spannung U_{GS} , die beim herkömmlichen Messen der Transferkennlinien durchlaufen wird, entspricht so in erster Näherung U_{SG} . Eine solche Kennlinienschar für den fertig gestellten Sensor zeigt Abbildung 7.9. Auch hierfür lässt sich die Steilheit nach Gleichung (7.3) ermitteln. Aus den errechneten Werten leitet sich die maximale Steilheit für eine Potenzialdifferenz U_{SG} von $-6\ \text{V}$ ab. Dieses Ergebnis deckt sich mit denen der vorhergehenden Untersuchungen des Transistors. Der Arbeitspunkt wird daher für die weiteren Untersuchungen folgendermaßen festgelegt:

$$\begin{aligned}
 U_{SG} &= U_4 = -6\ \text{V} \\
 U_{DG} &= U_5 = -4,6\ \text{V} \\
 U_{DS} &= U_{54} = 1,4\ \text{V} \\
 I_{DS} &= I_5 = 1,25\ \text{mA} \\
 g_m &= (713,64 \pm 0,29)\ \mu\text{A/V}
 \end{aligned} \tag{7.9}$$

Die Werte I_{DS} und g_m ergeben sich aus den angelegten Spannungen und sind dem Graphen in Abbildung 7.9 entnommen.

Neben diesen Transferkennlinien ist es ebenso erwägenswert, die Charakteristiken für die Eingänge über den Spannungsteiler Anschluss 1 und 2 (Abbildung 4.1) aufzunehmen. Diese, im Arbeitspunkt durchgeführten Messungen, sind in Abbildung 7.10 dargestellt. Auch daraus lässt sich nach Gleichung (7.3) die Steilheit für die beiden Eingänge berechnen. Man erhält $(21,04 \pm 0,13)\ \mu\text{A/V}$ und

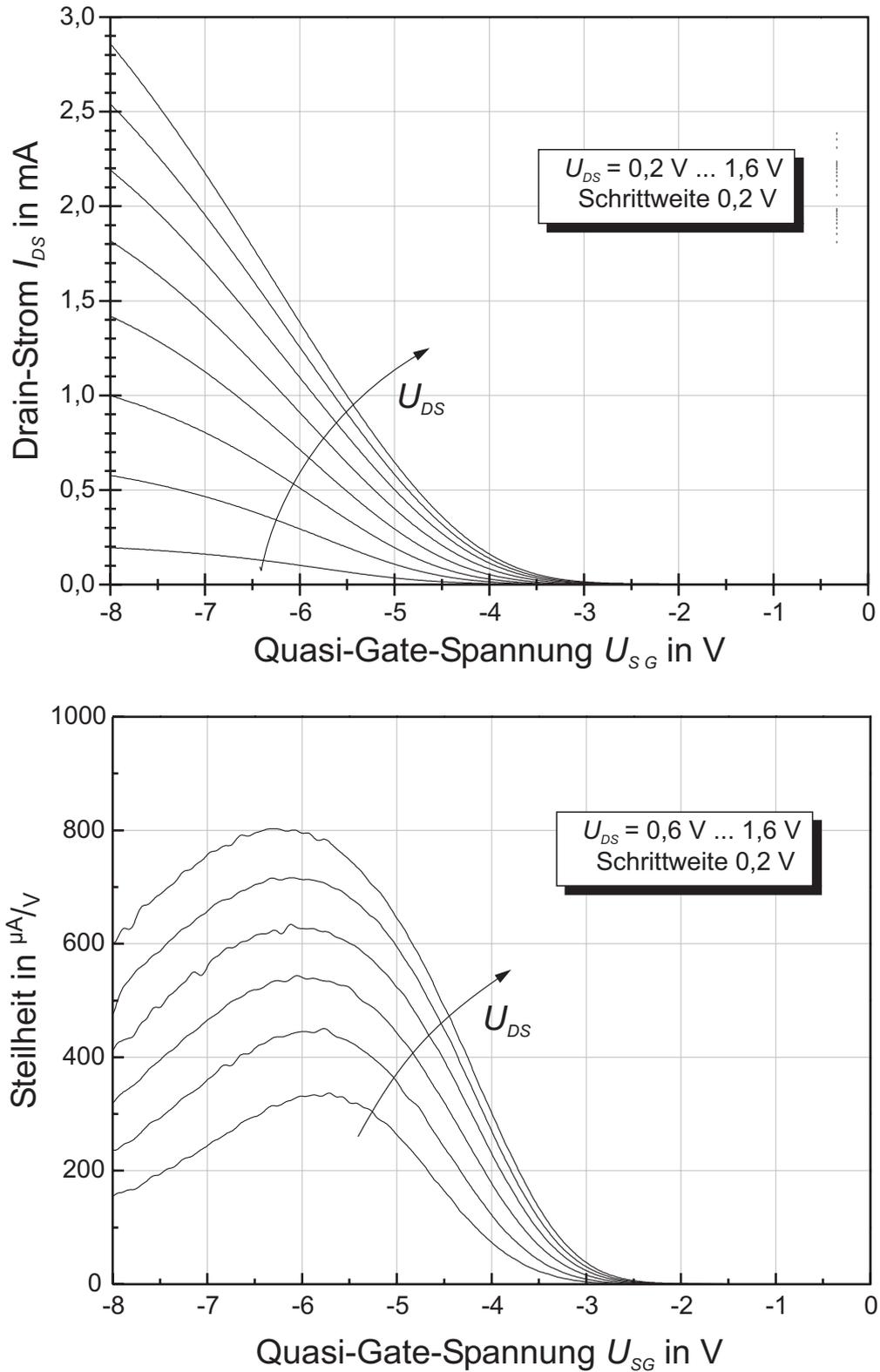


Abbildung 7.9: Transferkennlinie und Steilheit des Sensors gemessen mit der Quasi-Wannensteuerung, $U_1 = U_2 = 0 \text{ V}$

$(685,07 \pm 0,70) \mu\text{A}/\text{V}$ für den Deckel (Eingang 1) bzw. für die Gegenelektrode (Eingang 2). Mit diesen Daten ist nun die Frage lösbar, wie groß der Einkoppelfaktor der sensitiven Schicht auf den Transistor ist. Zu diesem Zweck wird das Verhältnis aus der Steilheit des Eingangs 1 mit der Transistorsteilheit gebildet. Es macht für diesen Sensor 1:35 aus und deutet auf eine zu geringe Kapazität C_1 zwischen Deckel und floatenden Gate hin. Das Verhältnis 1:32,56 der Steilheiten der beiden Eingänge zueinander gleicht dem der Kapazitäten C_1 und C_2 . Hiermit kann man unter Vernachlässigung des Siliziumnitrids den Abstand des Luftspaltes zu $2,4 \mu\text{m}$ abschätzen. Das ist etwas mehr als die gewünschten $1,7 \mu\text{m}$. In Zukunft muss der Abstand des Luftspaltes etwas verkleinert werden, um den Einkoppelfaktor zu steigern, denn dieser geht in das Signal-Rausch-Verhältnis ein.

Des Weiteren belegen die in Abbildung 7.10 dargestellten Messungen eine ausgezeichnete Linearität des Sensors im Arbeitspunkt bezogen auf eine Spannungsänderung an der sensitiven Schicht. Die berechneten Werte des Linearitätsfehlers sind in Abbildung 7.11 abgebildet. Die Abweichungen vom linearen Mittelwert befinden sich im Bereich von -5 bis 5 V innerhalb eines Fehlers kleiner als $\pm 1,5 \%$. Die Potenzialänderungen, die an der sensitiven Schicht zu erwarten sind, liegen im Bereich von nur 1 V . In diesem Intervall liegt der Fehler sogar unter 1% . Nichtlineare Kennlinien lassen sich also auf chemische oder physikalische Effekte, z. B. Physio- oder Chemisorption, an der sensitiven Schicht zurückführen, und erleichtern folglich weitere Abschätzungen, wie z. B. beim Betrachten der elektrochemischen Potenziale.

7.2.2 Messung einer Wasserstoffkonzentration

Die ersten Sensoren wurden mit 100 nm Platin als sensitiver Schicht aufgebaut. Darunter befindet sich 30 nm Titan als Haftvermittler. Dieses Material wurde neben Palladium in GASFETs am häufigsten untersucht und ist daher gut bekannt. Eine genauere Untersuchung der Reaktionen an Platin findet sich in [Sch02]. Die Schicht wurde mit der Leybold-Heraeus Z550 Magnetron-Sputteranlage im geerdeten DC-Modus abgeschieden.

Befindet sich Wasserstoff in der Atmosphäre, lagern sich Protonen an der Platinoberfläche an. Die Seite des Spannungsteilers mit der sensitiven Schicht (der Deckel) lädt sich positiv auf. Dadurch wird auch das Potenzial der floatenden Gate-Elektrode ins Positive gezogen. Um das zu kompensieren, sammeln sich daraufhin noch mehr Elektronen im Kanal des Transistors. Dessen Leitfähigkeit und somit der Betrag des Stromes (des n-Kanal-Transistors) nehmen zu.

Der Sensor befand sich während der ersten Versuche in einer hermetisch abgeschlossenen Kammer, der synthetische Luft ohne und mit Wasserstoff in verschiedenen Konzentrationen zugeführt werden kann. Es herrschte Raumtemperatur. Die Messungen wurden mit einem HP Parameteranalyzer 4155 durchgeführt.

An den Anschlüssen 4 und 5 wurde der Arbeitspunkt entsprechend Gleichung (7.9) eingestellt. Die Klemmen des Spannungsteilers 1 und 2 lagen unterdessen auf Erdpotenzial. Im Laufe der Messung wurde in die Kammer alternierend

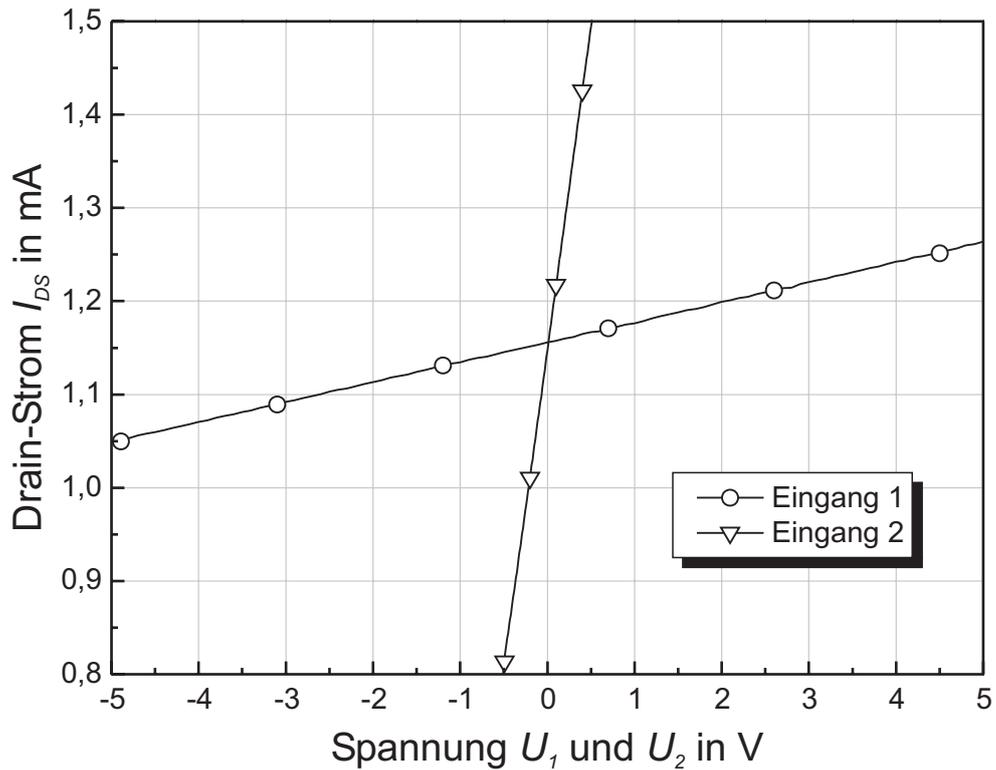


Abbildung 7.10: Transferkennlinien für Eingang 1 und 2 (Deckel und Gegenelektrode) gemessen im Arbeitspunkt

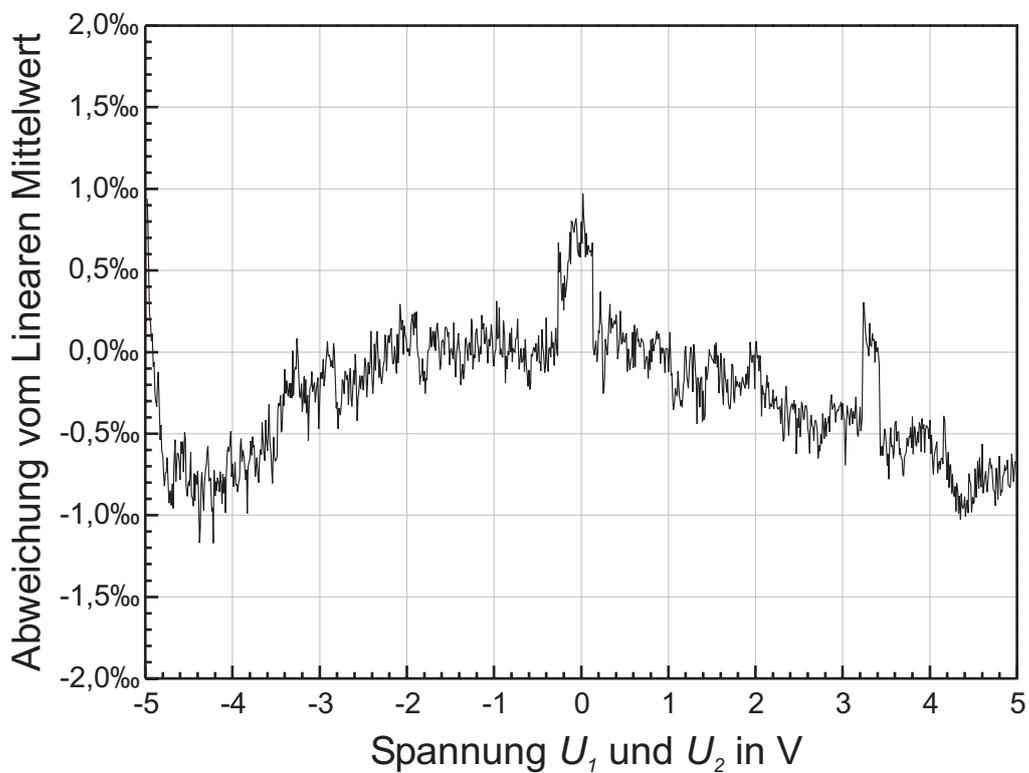


Abbildung 7.11: Linearitätsfehler des Sensors bei Potenzialänderungen an der sensitiven Schicht

Wasserstoff der Konzentrationen 2 % eingeleitet und die Zufuhr wieder gestoppt. Dabei betrug die relative Luftfeuchtigkeit 0 % . Alle 15 s wurde ein Messpunkt aufgenommen. Abbildung 7.12 präsentiert das Ergebnis der ersten Gaskonzentrationsmessung.

Sofort antwortet der Sensor mit einem Sprung um etwa $60 \mu\text{A}$ des Ausgangssignals I_{DS} auf die Konzentrationsänderung des Wasserstoffs auf 2 % . Im Anschluss fällt das Signal kontinuierlich ab. Nach dem Abstellen des Wasserstoffs beobachtet man einen augenblicklichen Abfall um ungefähr $45 \mu\text{A}$ des Stromes I_{DS} . Hierauf steigt er langsam aber konstant an. Diesen Verlauf durchläuft der Sensor mehrere Zyklen. Nur bei längerem Halten der Wasserstoffkonzentration bei 0 % verschiebt sich die Basislinie. Der Anstieg des Stromes zwischen den Reaktionen auf die Konzentrationsänderung bleibt indes unverändert. Das verdeutlicht die Ableitung des Signals, die in Abbildung 7.13 dargestellt ist.

Für einen kurzen Moment verhält sich das Signal wie erwartet. Der Strom nimmt im ersten Augenblick ansteigender Wasserstoffkonzentration zu und fällt im umgekehrten Fall. Die meiste Zeit reagiert der Sensor aber gegenteilig zur Theorie. Das deutet auf einen langsamen Fluss von Ladungen auf die Gate-Elektrode hin. Dieser gleicht den Potenzialunterschied aus, sodass der Drain-Strom I_{DS} wieder abnimmt. Infolge der nun nicht mehr genau definierten Spannungsverhältnisse fängt der Sensor an zu driften.

Dass Ladungen in den sensitiven Bereich hinein oder sogar durch das Siliziumnitrid wandern, ist nicht anzunehmen, denn Luftfeuchte ist nicht vorhanden und somit sind feuchteinduzierte Ströme auf der Oberfläche unwahrscheinlich. Gleichermäßen fraglich ist ein Leckstrom durch das 300 nm Oxid, welches das floatende Gate von der Gegenelektrode trennt. Am ehesten muss ein Leckstrom durch das Gate-Oxid direkt über dem Kanal angenommen werden. Dieses hat lediglich eine Dicke von 16 nm. Es wurde schon im Abschnitt 6.5 darauf hingewiesen, dass es bisher noch nicht gelungen war, die Mesaflanken der Transistoren perfekt zu glätten. An den Spitzen der Rauigkeit entstehen Feldüberhöhungen und vermutlich tunneln an diesen Stellen Ladungen.

Es sollte demnach machbar sein, diesen Leckstrom zu reduzieren, wenn der Spannungsabfall über dem Gate-Oxid herabgesetzt wird. Dies wird mit dem Verkleinern der Spannungen U_{SG} und U_{DG} erreicht. Deswegen wurde ein neuer Arbeitspunkt

$$\begin{aligned} U_1 &= U_2 = 0 \text{ V} \\ U_{SG} &= U_4 = -3,5 \text{ V} \\ U_{DG} &= U_5 = -2,3 \text{ V} \\ U_{DS} &= U_{54} = 1,2 \text{ V} \end{aligned} \tag{7.10}$$

knapp über der Einsatzspannung für eine weitere Messung herangezogen. Die Bedingungen dabei unterschieden sich nicht von den Vorigen und wurden bereits beschrieben.

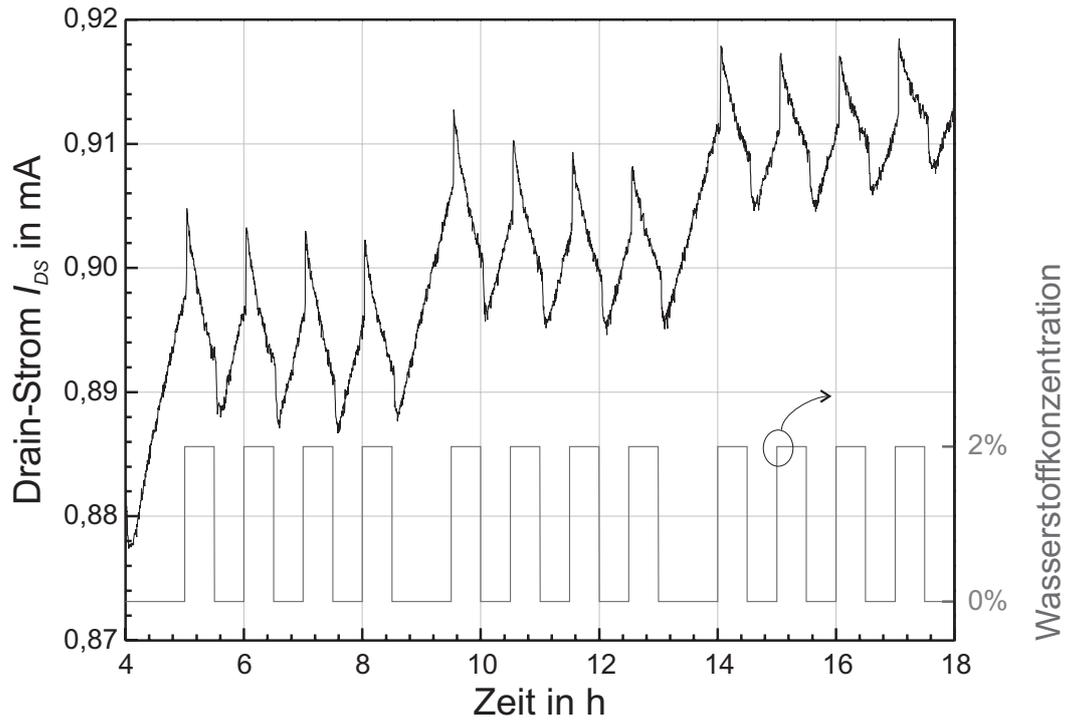


Abbildung 7.12: Reaktion des Sensors bei Wasserstoffbeaufschlagung. Der Arbeitspunkt entspricht (7.9)

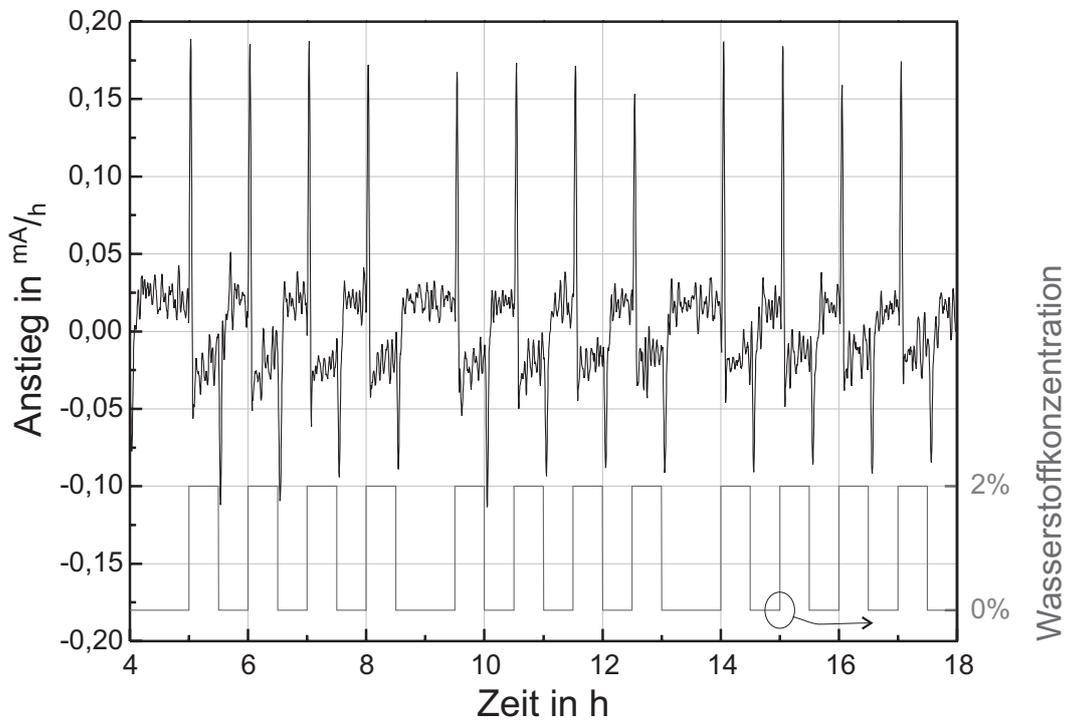


Abbildung 7.13: Ableitung des Signals der Messung in Abbildung 7.12

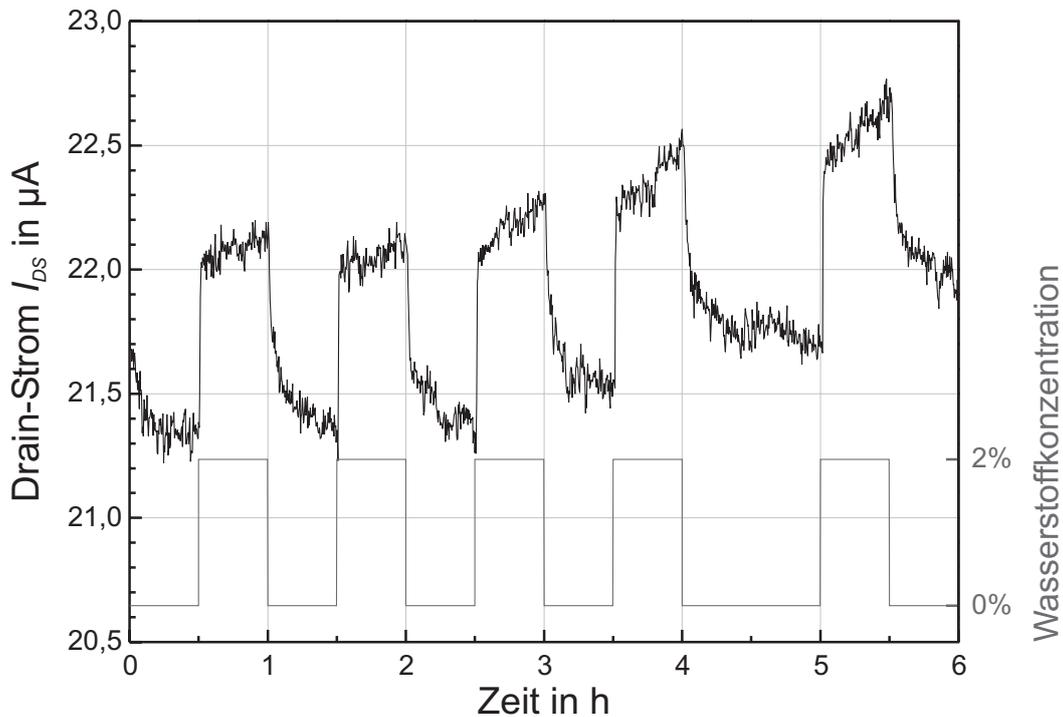


Abbildung 7.14: Reaktion des Sensors bei Wasserstoffbeaufschlagung im neuen Arbeitspunkt (7.10)

Die Antwort des Sensors dieser neuen Messung (Abbildung 7.14) ähnelt schon wesentlich besser den Rechteckimpulsen der Wasserstoffkonzentration. Unmittelbar mit dem Einleiten von 2 % Wasserstoff springt das Signal um $0,65 \mu\text{A}$ nach oben. Nach dem Stopp der Zufuhr reagiert der Sensor zwar prompt, der Ausgangspunkt stellt sich jedoch erst allmählich wieder ein. Ein schnelleres Ansprechen bei einem Anstieg im Gegensatz zu einem langsameren Abklingen bei Abnahme der Konzentration wurde auch schon in früheren Arbeiten beobachtet, die mit der hier verwendeten sensitiven Schicht experimentierten [Sch02]. Der Wasserstoff braucht mehr Zeit, um vollends vom Platin zu desorbieren als zum Sättigen der Oberfläche durch Adsorption. Daher hat das Signal beim Anheben der Konzentration schneller den Endwert erreicht als beim Wegnehmen des Gasflusses. Es ist bemerkenswert, dass die Basislinie ziemlich stabil bleibt.

Diese Messung (Abbildung 7.14) belegt die obigen Überlegungen. Ein Rücklaufen des Signals ist nicht mehr zu sehen, nachdem die Spannungen reduziert wurden. Es lässt sich also schlussfolgern, dass das Gate-Oxid noch nicht dicht genug ist und daher durch dieses ein Leckstrom auf die floatende Gate-Elektrode fließt. Der Sensor arbeitet wie erwartet.

8 Ausblick

In der vorliegenden Arbeit wurde ein Gassensor als Mikrosystem auf Basis der Siliziumtechnologie entworfen. Sie schildert die Umsetzung des Konzeptes und die Entwicklung des Herstellungsprozesses für diesen Sensor. Ein Prototyp wurde komplett aufgebaut und erste Messungen von Gaskonzentrationen durchgeführt. Alle Arbeiten fanden am Institut für Nanotechnologie und Mikrosystemtechnik der Universität der Bundeswehr statt.

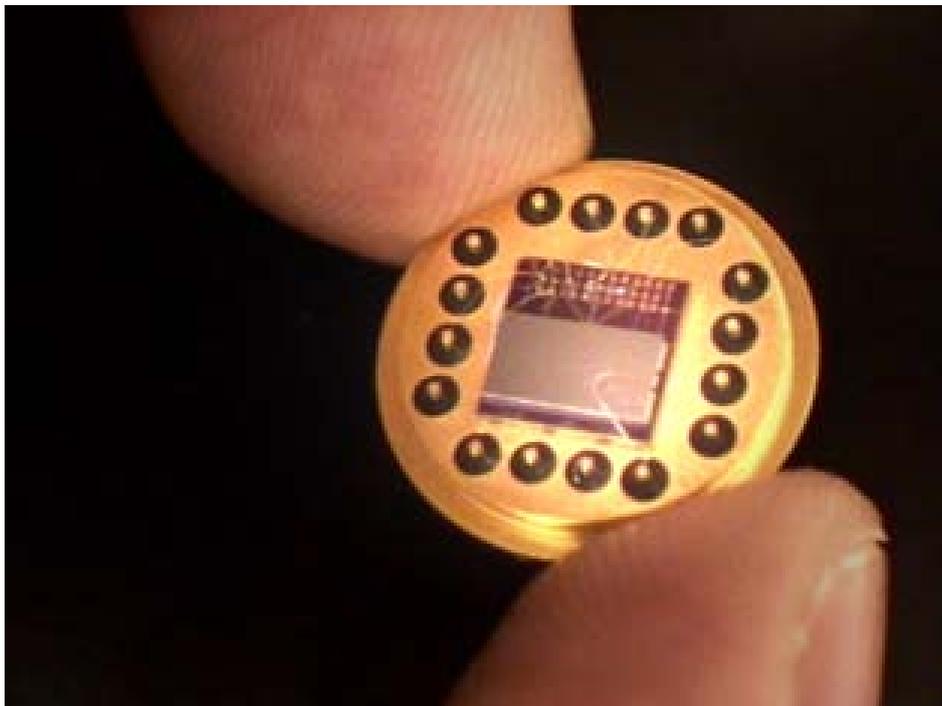


Abbildung 8.1: Ein Gassensor für tragbare Applikationen

Erste Tests, mit dem Gassensor zu messen, demonstrierten reversible Reaktionen und belegten die Funktion des in dieser Arbeit vorgestellten Konzeptes. Aus den Messungen ergibt sich das vorrangige Ziel, das Gate-Dielektrikum des Auslesetransistors zu verbessern und die Leckströme quer durch dieses zu unterbinden.

Es ist bisher noch nicht gelungen absolut glatte Flanken an 2 bis 3 μm hohen Mesen herzustellen, wie sie für die hier verwendeten Transistoren benötigt wurden. Wie sich jedoch in dieser Arbeit herausstellte, sind diese für gut funktionierende Transistoren unabdingbar. Zum einen führen raue Grenzflächen zwischen Kanal und Gate-Oxid zu geringen Beweglichkeiten, weil die Ladungsträger an der Grenzfläche gestreut werden. Zum anderen zeigen die auf unebenen Flächen gewachsenen Oxide hohe Leckströme. Besonders Letztere erwiesen sich bei

Gaskonzentrationsmessungen als störend, denn Sie verändern das Potenzial der floatenden Gate-Elektrode, worauf der Sensor zu driften beginnt. Mehrere Wege könnten zur Lösung dieser Problematik führen.

Mit dem Verwenden anderer SOI-Substrate, die eine dünnere Schichtdicke (etwa um $0,5\ \mu\text{m}$) des Device-Layers besitzen, wären die Mesen von vornherein nicht so hoch. Die Versuche dieser Arbeit zeigten, dass es durchaus machbar ist, ausreichend glatte Mesen mit einer Höhe von ungefähr $1\ \mu\text{m}$ herzustellen. Es war aber nicht möglich solche Substrate für diese Arbeit zu bekommen.

Ein Ausweg wäre, vor dem Herstellungsprozess die vorhandenen SOI-Wafer solange thermisch zu oxidieren, bis der verbleibende Device-Layer nur noch die gewünschte Dicke besitzt. Darüber hinaus hat das den erfreulichen Nebeneffekt, dass die Dotierstoffkonzentration dieses Films durch Segregation [Gro67, Eis01] noch weiter erhöht wird. (Es ist nämlich schwierig sehr hoch dotierte Substrate zu bekommen.) Ferner spezifizieren die Hersteller nur einen Bereich. Im Falle der verwendeten Substrate waren das $0,008$ bis $0,02\ \Omega\text{cm}$, was $3,5 \cdot 10^{19}$ bis $10^{20}\ \text{cm}^{-3}$ entspricht. Erfahrungsgemäß liegt die tatsächliche Konzentration eher am unteren Ende. Es sollte untersucht werden, ob sich der Device-Layer auf eine Stärke von $0,5\ \mu\text{m}$ verdünnen lässt, ob der Dotierstoffanteil danach die gewünschte Konzentration hat und ob darauf dann noch zuverlässig epitaktisch ein kristalliner Film gewachsen werden kann.

Der Ätzprozess selbst muss noch eingehender untersucht werden. Ein weiteres Absenken des Druckes und der Temperatur könnte ebenfalls Verbesserungen bringen. Andere Fotolacke bilden unter diesen Bedingungen eventuell mehr Polymere, was tieferes Ätzen zuließe.

Der nächste Vorschlag wäre, das Gate-Oxid gegen andere Materialien, v. a. High-k-Dielektrika, auszutauschen. Diese Materialien kommen dann zum Einsatz, wenn es darum geht, das Dielektrikum elektrisch dünner und gleichzeitig physikalisch dicker zu gestalten. Für diesen Sensor könnte unter Beibehaltung der äquivalenten Schichtdicke des Siliziumdioxids von $16\ \text{nm}$ ein wesentlich dickeres Dielektrikum auf den Kanal abgeschieden werden. Aufgrund der größeren Dicke sollte das Tunneln nicht so stark ausfallen [Sze69].

Wie in der Arbeit schon darauf hingewiesen, entstehen während zweier Prozessschritte (dem Mesaätzen und dem Freistellen) Flanken, die für den Kanal genutzt werden könnten. Der erste Ätzschritt muss nur bis zu einer Tiefe bis knapp unter die p-Schicht durchgeführt werden. Das waren im Falle des vorgestellten Prozesses lediglich 550 bis $600\ \text{nm}$. Die verschiedenen untersuchten Transistor-Designs sollten die Frage klären, welche der Kanten am besten für den Kanal geeignet sei. Leider kann Transistor Typ 3, der die kleine Stufe nutzt, aufgrund des hohen Zuleitungswiderstandes nicht für weitere Untersuchungen herangezogen werden. In einem neuen Layout wäre dies zu berücksichtigen. Das Problem könnte ein Transistor lösen, der ähnliche Geometrien besitzt wie Transistor Typ 3. Dieser dürfte aber nur eine Kante der Mesa für den Kanal nutzen, demzufolge nicht spiegelbildlich aufgebaut sein. Das Design darf keine zweite Metallisierungsebene erfordern.

Die vorgestellten Prozessschritte wurden so lange optimiert, bis sich der präsentierte Sensor herstellen ließ. Es besteht hierbei noch einiges Optimierungspotenzial.

Es wurde eine temperaturstabile Metallisierung in dieser Arbeit entwickelt, die bis 450 °C kein Spiking zulässt. Trotzdem ist der beste Weg, diese zu erzeugen, noch nicht gefunden. Durch eine Modifikation der Sputter-Anlage lässt sich ein gleichzeitiges DC-Sputtern und RF-Ätzen durchführen. Auf diese Weise ist es denkbar, ein besseres stöchiometrisches Verhältnis der Legierung bei gleichzeitig verringertem Sauerstoffgehalt zu erzielen. Zum Erreichen dieses Zieles ist ebenso ein Rücksputtern zum Entfernen des Wassers aus dem Rezipienten vor der eigentlichen Abscheidung zu untersuchen.

Aber auch der Herstellungsprozess bedarf kleinerer Änderungen, um die Ausbeute zu erhöhen und den Sensor zu verbessern.

Im Verlaufe der elektrischen Messungen des Sensors stellten sich viele Sensoren als unbrauchbar heraus, weil die Gegenelektrode sich nicht ansteuern ließ. Das Kontaktlochöffnen, so wie es in dem vorgestellten Herstellungsprozess vorgeschlagen wurde, führte doch zu erheblichen Ausfällen. Die Idee war, zuerst die Kontaktlöcher der Transistoren mit einem RIE-Ätzprozess durch nur 130 nm Siliziumnitrid zu öffnen und danach mit derselben Lackmaske das Oxid darunter nasschemisch zu ätzen. Die Oxiddicke in den Kontaktlöchern der Transistoren beträgt nur 16 nm, über der Gegenelektrode hingegen 300 nm. Diese Differenz zog aber technologisch erhebliche Schwierigkeiten nach sich. Die Kontaktlöcher der Transistoren müssen sehr genau justiert sein; das dünne Oxid ist dafür einfach zu öffnen. Die Justiergenauigkeit für die Kontaktlöcher der Gegenelektrode muss nur 10 µm genau sein, muss jedoch länger geätzt werden. Einfacher ist es, das Öffnen der verschiedenen Kontaktlöcher auf zwei Maskenschritte zu verteilen, um den unterschiedlichen Anforderungen gerecht zu werden.

In der Zukunft muss der Herstellungsprozess dahingehend erweitert werden, dass die Metallisierung ebenfalls passiviert wird. Die Zuleitungen des untersuchten Sensors sind erheblich länger (bis zu 5 mm) als diejenigen, die sonst für die Untersuchung von einzelnen Bauelementen hergestellt werden. Sie liegen offen und unterliegen insofern selbst Oberflächenreaktionen mit den umgebenden Gasen. Außerdem können bei vorhandener Feuchtigkeit Leckströme zwischen ihnen auftreten, die das Signal verfälschen.

Ein Forschungsziel des Instituts für Nanotechnologie und Mikrosystemtechnik ist die Selektive Epitaxie. Mit Hilfe dieser Technologie wird man im Stande sein, die Schichten für die Auslesetransistoren gezielt lokal abzuscheiden. Dann wäre es denkbar, den vorgestellten Prozess auf einem Wafer durchzuführen, der bereits einen CMOS Prozess durchlaufen hat. Es wird möglich sein, Sensor und Auswerteelektronik zu integrieren.

Anhang A

Die Abhängigkeit der Lage des Fermi-Niveaus von der Temperatur

Die Lage des Fermi-Niveaus W_F hängt von der Temperatur T ab. Das soll in diesem Abschnitt genauer erläutert werden.

Ein homogen dotierter (ebenso ein undotierter) Halbleiter ist im thermodynamischen Gleichgewicht in jedem Volumenelement elektrisch neutral. Für die Raumladungsdichte ρ gilt also, unter der Annahme, dass alle Dotieratome ionisiert sind, was ab etwa 100 K zutrifft [Gro67, Sze69]:

$$\rho = q(p - n + N_D - N_A) = 0 \quad (\text{A.1})$$

wobei,

- ρ Raumladungsdichte
- q Elementarladung
- p Anzahl der Löcher
- n Anzahl der Elektronen im Leitungsband
- N_D Dotierstoffkonzentration V-wertiger Atome
- N_A Dotierstoffkonzentration III-wertiger Atome.

Zusammen mit dem Dichteprodukt

$$n \cdot p = n_i^2 \quad (\text{A.2})$$

erhält man eine quadratische Gleichung. Setzt man deren Lösung für n bzw. p den Beziehungen [Gro67, Sze69]

$$n = n_i e^{(W_F - W_i)/kT} \quad (\text{A.3})$$

$$p = n_i e^{(W_i - W_F)/kT} \quad (\text{A.4})$$

gleich erhält man für n-dotierte Halbleiter

$$W_F - W_i = kT \ln \left(\frac{1}{2} \left(\frac{N_D - N_A}{n_i} + \sqrt{\frac{(N_D - N_A)^2}{n_i^2} + 4} \right) \right) \quad (\text{A.5})$$

bzw. für p-dotierte

$$W_F - W_i = -kT \ln \left(\frac{1}{2} \left(\frac{N_A - N_D}{n_i} + \sqrt{\frac{(N_A - N_D)^2}{n_i^2} + 4} \right) \right) \quad (\text{A.6})$$

wobei k für die Boltzmannkonstante steht. Damit ist die Lage des Fermi-Niveaus W_F bezogen auf das intrinsische Energieniveau W_i bestimmt. Trotz des ersten Anscheins der Gleichungen (A.5) und (A.6), ist dieser Zusammenhang nicht proportional zur Temperatur T . Die Intrinsische Dichte n_i ist selbst streng von Letzterer abhängig. Bei sehr großen Temperaturen nimmt die Intrinsische Dichte schließlich größere Werte als die Dotierkonzentration an. In diesem Falle geht das Argument der \ln -Funktion gegen 1. Folglich verschwindet die Differenz ($W_F - W_i$) zwischen dem Fermi- und dem intrinsischen Energieniveau. Diese Beziehungen sind in Abbildung 5.2 für verschieden Dotierkonzentrationen grafisch veranschaulicht.

Anhang B

Herstellungsverfahren für SOI-Wafer

Um SOI-Substrate herzustellen, wurden verschiedenste Materialien als Dielektrikum untersucht. Dabei entstanden viele Techniken, die auf einem epitaktischen Wachstum beruhen. Die Idee hierbei ist kristallines Silizium auf einen einkristallinen Isolator abzuscheiden. Entweder besteht der gesamte Handle aus diesem Isolator oder auf einen Si-Wafer wird mittels Heteroepitaxie ein Film des Isolators und einer aus Silizium aufgebracht. Vertretbar gute Ergebnisse lassen sich allerdings nur erzielen, wenn die die Gitterabstände der beiden Materialien, hier also Silizium und eines anderen, gut übereinstimmen. Das schon aus historischen Gründen wohl meist untersuchte Material ist hier sicherlich Saphir¹, welches bis Anfang der 80iger Jahre das einzige verfügbare SOI-Material war. Aber auch Wafer mit kubischem Zirkonium² oder Kalziumfluorid³, um nur wenige Beispiele zu nennen, sind erhältlich. Ein Problem all dieser heteroepitaktisch hergestellten Materialstapel liegt neben der hohen Defektdichte in der starken Verspannung der Wafer. Da bei gleichem Gitterabstand zweier Kristalle selten der Temperaturausdehnungskoeffizient gleich ist, und die meisten dieser Schichten bei ungefähr 1000 °C aufgebracht werden, sind diese Substrate bei Raumtemperatur extrem zerbrechlich. Das limitiert die Ausbeute in hohem Maße. Zusätzlich sollte erwähnt sein, dass es nicht wünschenswert ist, Stoffe wie Kalzium oder Zirkonium in eine Silizium-Reinraumumgebung zu bringen. Aus den erwähnten Gründen werden in dieser Arbeit nur Wafer aus Si und SiO₂ berücksichtigt. Einen umfassenderen Überblick über mögliche weitere Materialien gibt [Col02].

Zur Herstellung von kristallinem Silizium auf dielektrisch isolierendem Material wurden schon die unterschiedlichsten Ansätze untersucht und realisiert.

Ein Ansatz ist, eine Isolatorschicht, die ein Si-Wafer trägt, lokal zu öffnen und das freigelegte Silizium als Kristallisationskeim zu nutzen. Das Substrat soll dabei die Orientierung des Kristalls vorgeben, der das Dielektrikum lateral überwächst. Abbildung B.1 zeigt Beispiele für dieses Prinzip.

Es wird eine Poly-Silizium-Schicht über den, wie oben erwähnt präparierten Wafer gelegt. Diese wird im Anschluss daran rekristallisiert. Dies kann mit Hilfe eines Lasers, wie Abbildung B.1a zeigt geschehen, oder man verwandelt amorphes

¹Al₂O₃

²(Y₂O₃)_m · (ZrO₂)_{1-m}

³CaF₂

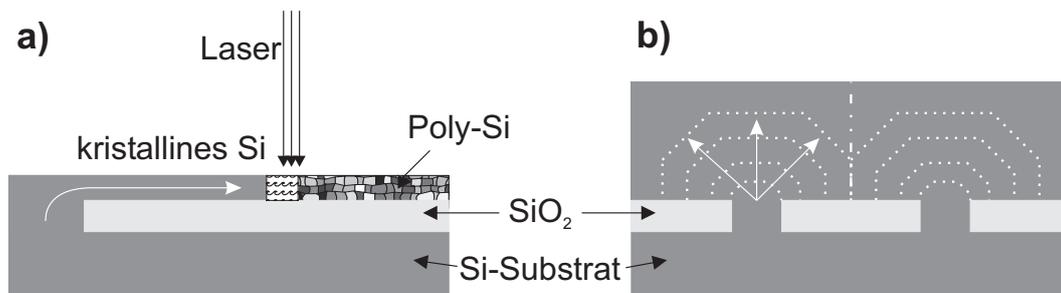


Abbildung B.1: a) laterale Rekrystallisation von Polysilizium über SiO₂,
b) epitaktisches Überwachsen von SiO₂

Silizium mit einem Temperaturschritt von 575 - 600 °Celsius in Kristallines. Man kann aber auch das SiO₂ direkt epitaktisch überwachsen (Abbildung B.1b).

Nachteil der angesprochenen und ähnlicher Verfahren ist die Tatsache, dass nie der gesamte Wafer eine dielektrisch isolierte aktive Schicht trägt, sondern nur Inselchen an SOI erzeugt werden können. Diese nehmen je nach Verfahren Flächen ein, die wesentlich kleiner als 1 mm² sind. Nur mit hohem Aufwand gelingt es, SOI-Filme mit einem Ausmaß von 4 mm² herzustellen [SKI⁺87]. Ein weiteres Problem ist die Entstehung von Korngrenzen, wo zwei Rekrystallisationsfronten aufeinander treffen. Schwierigkeiten zeigen sich auch in der schweren Kontrollierbarkeit der Kristallorientierung. Außerdem werden diese Substrate nicht auf dem Markt als Serienprodukt vertrieben. Aus den genannten Gründen werden in dieser Arbeit diese Substrate nicht weiter betrachtet. Einen tieferen Einblick in diese Verfahren gibt wiederum [Col02].

B.1 Das SIMOX-Verfahren

Mit SIMOX(Separation Implanted Oxygene)-Verfahren ist es möglich ein durchgängiges vergrabenes Oxid zu erzeugen. Dabei ist das Prinzip denkbar einfach. Es ist in Abbildung (B.2) skizziert. Mit einem Implantationsschritt werden Sauerstoffionen unter die Oberfläche eines Silizium-Wafers gebracht und danach bildet sich unterhalb der obersten Schicht Silizium während eines Temperschlittes ein SiO₂-Film aus.

Die eingeschossene Dosis muss extrem hoch sein. Für einen herkömmlichen Dotier-Implantationsschritt sind Ionendosen von 10¹⁶ cm⁻² bereits sehr hoch. Das entspricht ungefähr einer 1%igen Konzentration von Dotieratomen. Hier will man aber ein stöchiometrisches SiO₂ erzeugen, was zwei Sauerstoffatome pro Siliziumatom bedeutet. Die Gesamtzahl der benötigten Ionen erhöht sich zusätzlich, da das Implantationsprofil nicht rechteckig sondern gaussförmig ist. Erst ab der Dosis 1,4 · 10¹⁸ cm⁻² entsteht zuverlässig eine geschlossene vergrabene Oxidschicht. Für die heutzutage gefertigten SIMOX-Wafer wird das 200- bis 500fache der maximalen Dosis einer Dotierimplantation verwendet. Um das zu bewerkstelligen,

wurden spezielle Sauerstoff-Implanter entwickelt [BDD⁺03]. Typischerweise verwendet man heute eine Implantationsdosis von $2 \cdot 10^{18} \text{ cm}^{-2}$ mit einer Energie von 180 keV, wobei das Resultat ein 200 nm dicker Device-Layer und ein 400 nm BOX ist.

Schon ab einer Ionendosis von 10^{15} cm^{-2} wird das Kristallgitter so stark geschädigt, dass das Silizium in einen amorphen Film umgewandelt wird. Herkömmlich werden die Defekte durch einen anschließenden Temperaturschritt ausgeheilt. Da hier jedoch das vergrabene Oxid die obere Siliziumschicht von dem Handle trennt, kann so die Kristallorientierung nicht kontrolliert werden. Deshalb wird die Implantation bei einer Temperatur zwischen 400 und 600 °C durchgeführt. So können die Kristallschäden schon während des Implantierens wieder ausheilen. Der Device-Layer weist zwar im Anschluss eine hohe Defektdichte auf, ist aber genügend intakt, um beim folgenden Temperaturschritt die Kristallinität und Orientierung wieder herzustellen.

Der Annealing-Prozess wird bei 1300 bis 1350 °C in einer Stickstoff- oder Argon-Atmosphäre sechs Stunden lang durchgeführt. Neben dem oben angesprochenen Ausheilen der Kristalldefekte tritt noch ein anderer Effekt ein. Das Implantationsprofil über die Tiefe ist annähernd gaussförmig. Mit der Erhöhung der Temperatur steigt die Beweglich- und die Löslichkeit von Sauerstoff in Silizium. Kleine SiO_2 -Klumpchen lösen sich auf und gleichzeitig fällt SiO_2 in sauerstoffreichem Silizium aus, scheint also an größeren Ausfällungen zu kondensieren. Bei genügend großen Temperaturen bleibt nur das vergrabene Oxid als durchgängige Schicht übrig.

Mit unterschiedlichen Strategien versucht man die Defektdichte, die als Folge des Implantierens auftritt, zu reduzieren. Man kann durch die zwei bis dreifache Wiederholung des Verfahrens die Implantationsdosen auf die einzelnen Durchläufe aufteilen (Multiple Implants). Dadurch wird das Material weniger geschädigt und es lassen sich hoch qualitative Wafer herstellen. Aufgrund des Mehraufwands steigen die Kosten.

Ein anderer Ansatz (Low Energy Implant) setzt die Implantationsenergie auf 35 bis 70 keV herab, wodurch die benötigte Dosis sinkt. Der Device-Layer und das BOX werden dünner. Durch die geringeren Energien verringern sich auch die Verunreinigungen, die aus dem Implanter selbst herrühren.

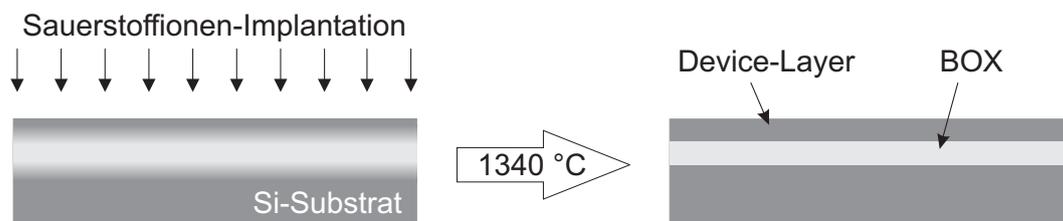


Abbildung B.2: SIMOX-Verfahren zur Herstellung von SOI-Substraten

Ganz ähnlich geht man beim Low-Dose-SIMOX-Verfahren vor. In einem weiteren schmalen Prozessfenster mit Dosen um $4 \cdot 10^{17} \text{ cm}^{-2}$ und einer Energie von 120 keV ist es möglich vergrabene geschlossene Filme zu erzeugen. Auch hier erhält man eine dünne aktive Schicht und ein dünnes vergrabenes Oxid. Ein großer Teil der Kosten kann gesenkt werden, weil sich der Durchsatz der Implantier erhöht. Das vergrabene Oxid mit diesem Verfahren hergestellter SOI-Wafer kann mit dem ITOX-Prozess (Internal Thermal Oxidation) noch nachträglich verstärkt werden. Dazu wird der Wafer noch einmal oxidiert (1350 °C), wobei nicht nur ein Oxid auf der Oberfläche wächst, sondern auch das Innere.

Am Rande soll noch erwähnt sein, dass durch das Implantieren von Stickstoff oder Stickstoff und Sauerstoff vergrabene Nitrid- bzw. Oxinitridschichten erzeugt werden können. Die Akronyme dieser Verfahren heißen SIMNI (Separation by Implanted Nitrogen) und SIMON (Separation by Implanted Oxygen and Nitrogen).

Qualitätsprobleme bei diesem Material rühren v. a. vom Implantieren her. Dadurch leidet die Kristallqualität. Das zeigt sich durch die hohe Anzahl von „HF-Defekten“. Zur Ermittlung dieses Wertes wird das Substrat Flusssäure ausgesetzt. Normalerweise greift diese Silizium nicht an, sind aber Fehlstellen in diesem vorhanden, ätzt die Säure kleine Löcher in den Device-Layer und löst das darunter liegende Siliziumdioxid. Das kann dann leicht mit dem Mikroskop ausgewertet werden. Das Behandeln mit Secco-Ätze⁴ ist ein ähnliches Verfahren und liefert auch als Ergebnis eine relativ hohe Anzahl von Versetzungen und Stapelfehlern. Außerdem sind die Konzentration an Metallen sehr hoch, die während dem Implantieren von den Wänden der Ionen-Implantern abgesputtert werden und bei den Temperschritten in das Material eingebaut werden. Diese Silizide werden von Flusssäure gelöst, sodass sich auch dadurch die „HF-Defektrate“ erhöht. Typischerweise findet man viele Versetzungslinien und Stapelfehler in der Nähe des Dielektrikums.

Das Dielektrikum besitzt nicht die Qualität eines thermischen Oxids, das Durchbruchspannungen von 10 - 16 MV/cm aufweist. Das BOX dieses Verfahrens erreicht Werte von ungefähr 8 MV/cm. Es ist siliziumreich und man findet Siliziumkristallite in ihm. Ein weiteres Problem dieses Materials sind die „Pinhole“-Defekte, welche, wie Fäden aus Silizium, vom Handle durch das Oxid zur aktiven Sicht reichen. Diese entstehen durch Mikromaskierungen aufgrund von Verunreinigungen beim Implantieren. Die Dichte solcher Fehler wird mit $0,1 \text{ cm}^{-2}$ angegeben [ibi03].

B.2 Gebondete SOI-Wafer

Neben dem oben genannten SIMOX-Verfahren wird zur kommerziellen Herstellung von SOI-Substraten hauptsächlich das Bonden verwendet. Zwei Wafer die jeweils eine Oxidschicht tragen werden in einem klebstofffreien Prozess miteinander

⁴1 Teil $\text{K}_2\text{Cr}_2\text{O}_7$: 2 Teile HF (49 %)

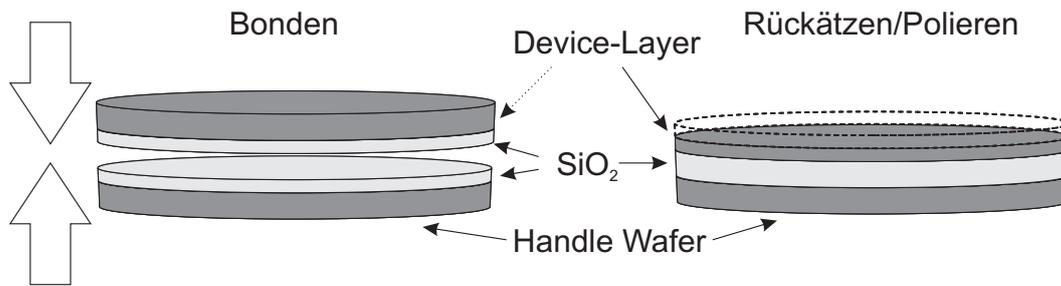


Abbildung B.3: BESOI-Verfahren zur Herstellung von SOI-Substraten

verschweißt oder gebondet. Danach wird einer der beiden Wafer soweit abgetragen, dass nur noch der Device-Layer stehen bleibt. Die Abbildung B.3 zeigt dieses Prinzip.

Der Bondmechanismus dabei ist der gleiche, wie beim Siliziumdirektbunden. Werden zwei hydrophile Oberflächen zusammengedrückt, bilden sich zwischen den auf den Oberflächen adsorbierten $(\text{OH})^-$ und H_2O -Molekülen Van-der-Waals-Kräfte aus. Der Anpressdruck wird nur zum Verdrängen der Luft im Spalt und zum Ausgleich von großflächigen Wölbungen der Oberflächen benötigt. Die extremen Ebenheiten der Wafer ermöglichen dabei das Annähern der Fügepartneroberflächen. An Erhebungen in der Mikrorauigkeit berühren sich die Materialien zunächst und es bilden sich dort Wasserstoffbrücken aus. Die Kraft reicht aus, um die Mikroprofile durch plastische Verformung der Oberfläche völlig auszugleichen und überall Wasserstoffbrücken entstehen zu lassen. Es breitet sich als Folge eine Bondfront mit einer Geschwindigkeit von mehreren cm/s aus. Diese Verbindung ist allerdings noch lösbar. Durch anschließendes Tempern wird diese Verbindung verfestigt. Bei einer Temperatur von etwa 700°C werden die Hydroxyl-Brücken gegen kovalente Sauerstoffbindungen in der Reaktion



ausgetauscht. Das Wasser sammelt sich in kleinen Bläschen zwischen den zusammengefügt Oberflächen. Diese verschwinden bei noch höheren Temperaturen wieder, wenn das Wasser das Silizium oxidiert. Der dabei entstehende Wasserstoff diffundiert heraus. Bei Temperaturen ab 1100°C verfließt das SiO_2 ineinander, die Grenzfläche verschwindet und die Wafer werden vollständig miteinander verschweißt.

Es ist auch möglich nur eine Scheibe der beiden Wafer vor dem Bonden zu oxidieren. Durch Vorbehandeln der Si-Oberfläche wird diese hydrophil.⁵ An deren Oberfläche bilden sich Si-OH -Gruppen, über die eine Bindung zur Fügepartneroberfläche nach dem oben erwähnten Prinzip hergestellt wird [GRMS00].

Probleme machen Partikel beim Bonden. Teilchen mit einer Größe kleiner als $1\ \mu\text{m}$ können Lücken mit Durchmessern von mehreren Millimetern erzeugen.

⁵z. B. Tauchen in ein Gemisch aus 90 Teile H_2O :5 Teile H_2O_2 :10 Teile NH_4OH

Da diese Verunreinigungen meistens organisch sind, lassen sie sich nicht durch Temperschritte entfernen. Um solche Fehlstellen zu vermeiden, wird der Bond-Prozess in einer Reinstraumgebung durchgeführt. Vor dem Kontakt muss der Spalt zwischen den beiden Wafern mit Reinigungskemikalien und deionisiertem Wasser gespült, und im Infrarotlicht oder durch Schleudern getrocknet werden.

Nachdem Bondprozess ist es erforderlich, die obere Siliziumschicht von der ursprünglichen Dicke (d. h. im Falle eines 100-mm-Wafer bis zu 500 μm) auf die gewünschte Schichtstärke zu verdünnen. Dazu wird zunächst meistens das Schleifen eingesetzt, um den Großteil des Materials abzutragen, da diese Methode ziemlich schnell und kostengünstig ist. Die letzten wenigen Mikrometer können dann mittels CMP (Chemico-Mechanical-Polishing) weggenommen werden. Diese Technik ist ebenso einfach, schnell und v. a. billig. Allerdings gibt es hierfür keinen richtigen Ätzstopp. Als Referenz wird hier die Rückseite des Wafers genommen und so lässt sich die Schichtdicke nicht beliebig genau kontrollieren. Deshalb ist dieses Verfahren auf eher dicke SOI-Filme von einigen Mikrometern beschränkt. Außerdem kann dieses Polieren nicht mit einem Mehr-Scheiben-Polierer durchgeführt werden, da für jede einzelne Scheibe genau justiert werden muss. Zur Erzeugung von dünneren Device-Layern sind spezielle Maschinen entwickelt worden, die eine in-situ Messung der Schichtdicke während des Polierens vornehmen. Damit ist es möglich, eine SOI-Schichtdicke von einem Mikrometer und Schichtdickenschwankung besser als 10 % herzustellen.

Die Kristallqualität der gebondeten SOI-Wafer hängt im Wesentlichen von der Qualität der verwendeten Ausgangssubstrate ab. Durch das Bonden lassen sich thermische Oxide mit hohen Qualitäten oder gesputterte Oxide mit geringen mechanischen Verspannungen verwenden. Diese sind frei von Pinhole-Defekten. Nachteilig ist allerdings die relativ schlechte Schichtdickenkontrolle bei dem Schleif- und Polierverfahren. Absolut liegt der Fehler minimal bei 0,5 μm , was bei den minimalen Schichtdicken des SOI von 1 μm zu einem relativen Fehler von 50 % führt.

B.3 BESOI-Wafer

Sehr viel genauer lässt sich eine gewünschte Schichtdicke einstellen, wenn nach dem Schleifen ein Ätzschritt (der sog. Etch-back) folgt, für den man vorher einen Ätzstopp erzeugt. BESOI (Bonded and Etchback Silicon on Insultator) heißt das Akronym für solche SOI-Wafer. Wie in Abbildung B.3 dargestellt, werden zwei oxidierte Silizium-Wafer gebondet. Danach wird eine Seite geschliffen und geätzt, bis die gewünschte Schichtdicke der aktiven Schicht erreicht ist.

Am einfachsten lässt sich der Ätzstopp über ein implantiertes Dotierprofil realisieren. Zum Beispiel wird eine hoch dotierte p-Schicht aus Bor in einen leicht dotierten Wafer eingebaut. Nach dem Schleifen wird mit einer selektiven

Ätzlösung, wie KOH⁶ oder TMAH⁷, das leicht dotierte Material entfernt. Danach wird mit „1-3-8-Ätze“⁸ selektiv die hoch dotierte p-Schicht entfernt. Mit einem anschließenden Wachsen und sofortigem Entfernen eines thermischen Oxides kann die Bor-Konzentration in der aktiven Schicht wieder verringert werden.

Ausgefeiltere Techniken nutzen epitaktisch gewachsene Schichten als Ätzstopp. Auf diese wird der Device-Layer epitaktisch aufgewachsen. Es soll am Rande erwähnt sein, dass auch mit Ätzstopps aus implantiertem Stickstoff oder Kohlenstoff sowie mit epitaktisch gewachsenem Silizium-Germanium und Siliziumcarbit experimentiert wurde. Die Variation der Schichtdicke über den Waferdurchmesser hängt bei diesen Rückätzverfahren im Wesentlichen von dem Ätzstopp ab und kann auf ungefähr 10 nm genau kontrolliert werden.

Die Kristallqualität hängt von der Implantation oder der Epitaxie ab. Der Mehraufwand steigert allerdings den Preis der ohnehin schon teuren Substrate. Das Rückschleifen oder -ätzen der Wafer kostet durch dieses Verfahren zwei Wafer. Einen als Handle und einen für den Device-Layer.

B.4 UNIBOND-Wafer

Um den hohen Verbrauch an Silizium zu verringern sind verschiedene Techniken entwickelt worden, bei denen nur die dünne Oberflächenschicht auf den Handle-Wafer übertragen werden. Ein solcher Prozess ist der Smart-Cut-Prozess bei dem UNIBOND Material hergestellt wird.

Dessen Ablauf illustriert Abbildung B.4. Zuerst wird ein Wafer thermisch oxidiert. Dieses SiO₂ dient später als die vergrabene isolierende Schicht. In diesen Wafer wird dann Wasserstoff oder Wasserstoff mit geringen Anteilen Helium implantiert. Die Dosen hierbei liegen in der Größenordnung von $5 \cdot 10^{16} \text{ cm}^{-2}$. Dadurch werden Mikrobäschen und -löcher in der entsprechenden Implantations-tiefe erzeugt. Während der Implantation verliert das Oxid wegen Verunreinigungen die hydrophile Eigenschaft. Deshalb ist man gezwungen die Oberfläche vor dem Bonden sehr gut zu reinigen. Nun wird ein Bondprozess, wie in Abschnitt B.2 erklärt, durchgeführt.

Darauf folgt ein zweistufiger Ofenprozessschritt, mit dem zum einen die zu übertragende Schicht getrennt und zum anderen das SiO₂ mit dem Handle-Wafer verschweißt wird. Die erste Stufe erfolgt bei 500 °C für zwei Stunden. Dabei nehmen die Mikrobäschen an Größe zu und breiten sich vorzugsweise entlang der $\langle 1-0-0 \rangle$ - Ebenen aus. Dieser Mechanismus wurde auch an anderen Materialien beobachtet, die Wasserstoff oder Helium ausgesetzt waren [Sch02]. Schließlich splitten diese Fehlstellen im Kristall den Wafer parallel zur Oberfläche in zwei Teile. Für weitere

⁶Kaliumhydroxid-Lösung

⁷Tetramethylammoniumhydroxid-Lösung

⁸Fluss-, Salpeter- und Essigsäure: 1 Teil HF (49 %): 3 Teile HNO₃ (65 %): 8 Teile CH₃COOH (100 %)

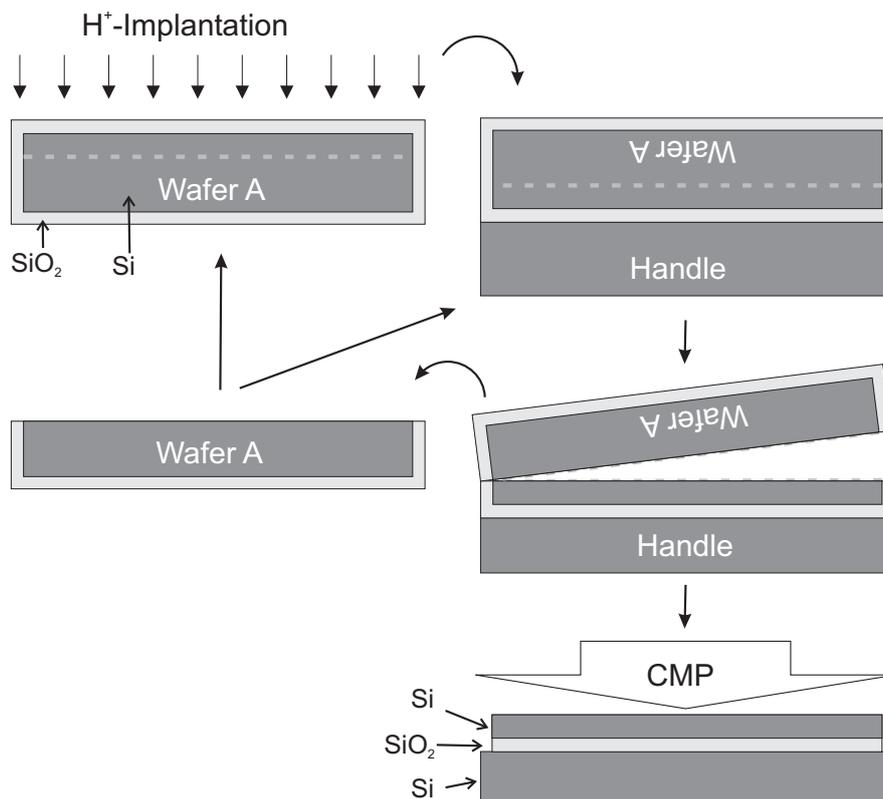


Abbildung B.4: Smart-Cut-Verfahren zur Herstellung von SOI-Substraten

zwei Stunden wird die Temperatur auf 1100 °C erhöht, was die Haftkraft zwischen der Schnittstelle, wie schon beschrieben, verstärkt.

Zum Schluss wird mittels CMP die Oberfläche des Substrates geglättet, da direkt nach dem Trennen eine hohe Rauigkeit zurückbleibt. Bei diesem Schritt werden einige 10 nm abgetragen und Rauigkeiten kleiner als 0,15 nm erzielt.

Der abgetrennte Teil des implantierten Wafers wird wieder verwendet. Er dient für den nächsten Durchlauf als Handle-Wafer, kann aber sogar wieder dem Implantieren zugeführt werden. Das senkt die Kosten enorm, denn theoretisch wird nur ein Wafer mehr benötigt als man insgesamt in der Charge herstellen will.

Die Qualität des Oxides ist vergleichbar mit derjenigen, der oben genannten gebondeten SOI-Wafer. Die Variation der Dicke der aktiven Schicht kann sehr genau kontrolliert werden. Die Wasserstoffimplantation schädigt aber die aktive Schicht, was zu einer erhöhten Zahl an HF-Defekten führt.

B.5 ELTRAN-Wafer

Ähnlich wie beim Smart-Cut-Verfahren wird beim ELTRAN (Epitaxial Layer Transfer)-Prozess nur eine dünne Schicht auf den Handle-Wafer übertragen. Abbildung B.5 veranschaulicht den Ablauf schematisch.

Zuerst wird ein einkristalliner Silizium-Wafer, der Keim-Wafer, elektrochemisch in einem Gemisch aus Flusssäure und Ethanol geätzt. Der Aufbau entspricht dabei einer elektrochemischen Zelle, wobei das Silizium als Anode geschaltet ist. Dabei wird das Silizium nicht auf der gesamten Fläche abgetragen, sondern es wird porös. Die Reaktion findet nur am Ende der Kanälchen statt und diese werden immer länger. Über die Stromdichte wird dabei die Größe der Poren eingestellt. Zunächst bildet sich bei einer Stromdichte von etwa 8 mA/cm^2 eine $13 \mu\text{m}$ dicke Schicht mit Poren einer Dichte von 10^{11} cm^{-2} und eines Durchmessers von ein paar Nanometern. Danach wird durch Erhöhung der Stromdichte auf ungefähr 22 mA/cm^2 darunter eine zweite poröse Schicht der Dicke $3 \mu\text{m}$ mit etwa dreimal

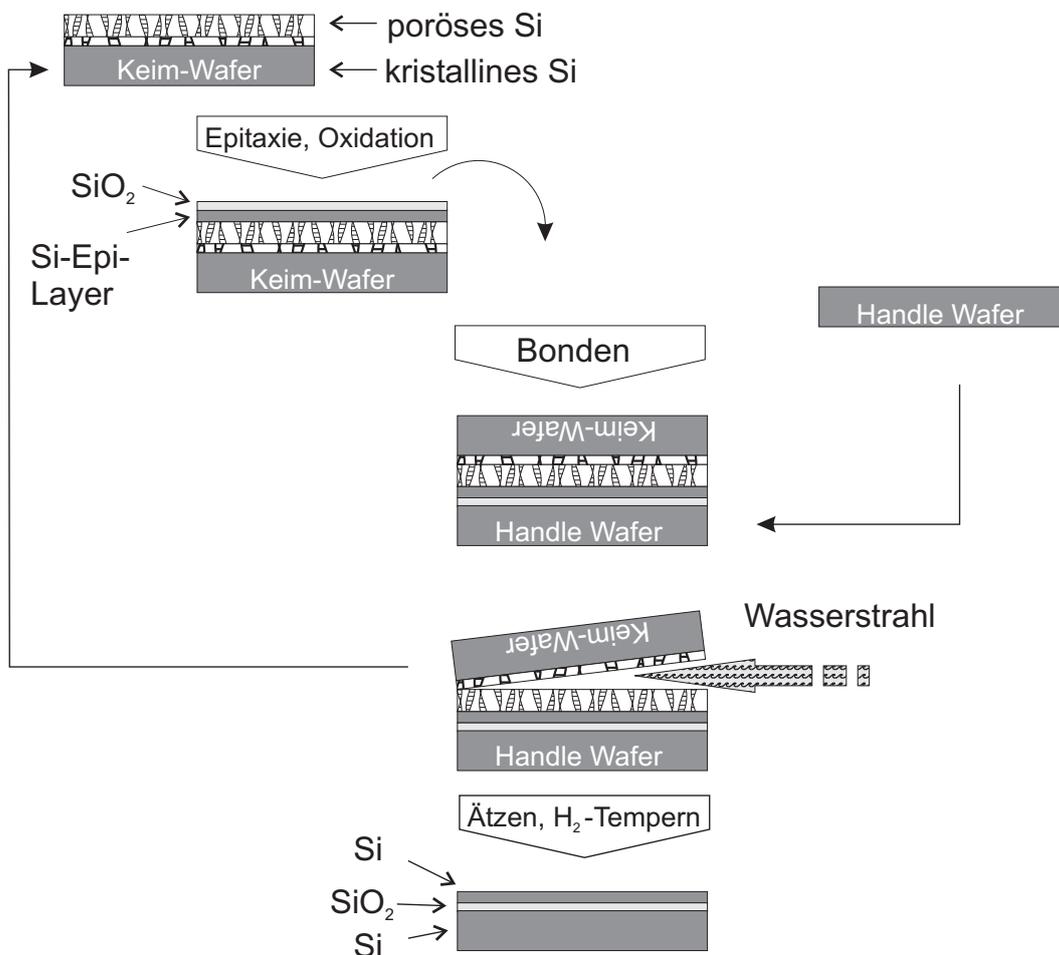


Abbildung B.5: ELTRAN-Verfahren zur Herstellung von SOI-Substraten

so großen Poren erzeugt. Diese Doppelschicht ist notwendig, um später die aktive Schicht abtrennen zu können, ohne sie zu beschädigen.

Ein 3 nm Trockenoxid wird bei etwa 400 °C gewachsen, um die bisher erzeugte Struktur bei den folgenden Prozessschritten zu schützen. Danach wird dieser Wafer in dem CVD-Reaktor bei 1100 °C in einer Wasserstoff-Atmosphäre getempert. Dabei verschließen sich die Poren und es entsteht wieder eine glatte geschlossene Oberfläche. Das Oxid in den Poren verhindert dabei, dass die poröse Struktur verloren geht. Nun wird mittels CVD die zu übertragende Schicht abgeschieden.

Bei einer anschließenden thermischen Oxidation entsteht das SiO₂ über der epitaktisch gewachsenen Schicht, das später das vergrabene Dielektrikum bildet. Dieses wird nun kopfüber auf einen Handle-Wafer in dem oben beschriebenen Verfahren gebondet.

Um den Keim-Wafer wieder zu trennen und die aktive Schicht zu übertragen, wird ein Wasserstrahl mit einem Druck von 200 bis 600 bar mit einem Durchmesser von 100 µm verwendet. Während der Strahl den Wafer-Stapel trifft, rotiert dieser. Der Effekt ist weniger ein Schneiden als ein Schälen. Ein Keil aus Wasser bildet sich zwischen dem zu trennenden Silizium aus, der im Gegensatz zu einem festen Keil die Kraft nicht nur an zwei Punkten einleitet, weil der „flüssige“ Keil seine Form ändert und sich anschmiegt. So konzentrieren sich die Verspannungen nicht auf einen Punkt. Daher wird eine Beschädigung der Wafer vermieden. Die Trennung erfolgt genau an der Grenzfläche zwischen den zwei unterschiedlich porösen Schichten [SYS⁺99].

Das verbleibende poröse Silizium auf dem SOI-Substrat wird mittels eines Gemischs aus Flusssäure, Wasserstoffperoxid und Wasser entfernt. Wegen der extrem großen Oberfläche des porösen Materials lässt sich dieses mit Selektivitäten bis zu 100000 gegenüber dem kristallinen Silizium bewerkstelligen. Aufgrund des Kapillareffektes dringt die Säure in die Poren und vergrößert diese in alle Richtungen. Sobald sich die Hohlräume so weit ausgedehnt haben, dass sie sich gegenseitig berühren, kollabiert die poröse Struktur.

Zum Schluss wird durch eine Temperung in Wasserstoff-Atmosphäre bei 1100 °C die Oberfläche geglättet. Hierbei werden Rauigkeiten kleiner als 0,1 nm erzielt [SY94]. Dieses Resultat ist mit sehr glatt poliertem Silizium vergleichbar.

Die andere abgetrennte Siliziumscheibe kann wieder aufbereitet und neu verwendet werden. So kann der Keim-Wafer drei bis vier Zyklen durchlaufen. Da das gesamte Material, welches für das SOI-Substrat gebraucht wird, wie Epi-Schicht, BOX oder Handle-Wafer neu hergestellt oder eingebracht werden, verschlechtert sich auch nicht durch die Wiederverwendung die Qualität.

Da bei diesem Verfahren die aktive Schicht epitaktisch aufgewachsen wird, enthält sie keine COPs (Crystal Originated Partikel). Das sind Löcher, die regelmäßige Oktaeder aus den $\langle 1-1-1 \rangle$ -Flächen bilden und mit SiO₂ gefüllt sind. Diese entstehen beim Ziehen von Silizium aus der Schmelze durch den anwesenden Sauerstoff. Epitaktisch gewachsenes Silizium enthält diese Fehlstellen nicht. Ein weiterer Vorteil

der ELTRAN-Wafer ist die geringe Rauigkeit aufgrund des Wasserstofftempers. Außerdem ergeben sich sehr einheitliche Dicken des SOI.

B.6 Vergleich der Herstellungsverfahren

Die Vor- und Nachteile der einzelnen Herstellungsverfahren und der damit hergestellten Substrate wurden schon bei deren jeweiligen obigen Erörterungen diskutiert. Tabelle B.1 vergleicht die Hersteller-Spezifikationen verschiedener SOI-Substrate. Diese Werte stellen nicht das maximal bzw. minimal Mögliche dar, sondern geben das wieder, was der jeweilige Hersteller garantiert. Wie aus der Tabelle B.1 ersichtlich, ähneln sich die Qualitätsmerkmale sehr. Das mag an Konkurrenzgründe liegen. Deshalb geben die einzelnen Firmen z. B. die Partikeldichten für verschieden Größen an. So erwähnt Ibis für ihre SIMOX-Produkte wohl lieber die Metallverunreinigung mit jeweils $5 \cdot 10^{10}$ Atome/cm² für die gesonderten Metalle [ibi03], statt die gesamten $1,5 \cdot 10^{11}$ Atome/cm² für alle Metalle.

Nicht die Qualitätsmerkmale alleine sind für die Wahl des Substrates entscheidend. Vielmehr erfüllen die Herstellungsverfahren unterschiedliche Nachfragen und Anforderungen aus verschiedenen Bereichen, wie z. B. CMOS- oder MEMS-Fertigung. Zuerst muss die Frage gestellt werden, wie die aktive Schicht und das vergrabene Dielektrikum für die spezifische Anwendung aussehen soll.

Weil die sensitiven Flächen und die Transistoren elektrisch voneinander getrennt sein sollen und Leckströme verhindert werden müssen, werden diese Bauteile durch Mesaätzen freigestellt. Dazu muss ein Graben in das SOI bis herunter auf das BOX geätzt werden. Der Aufwand dabei steigt mit der Tiefe des Grabens. Anschließend werden die Elemente kontaktiert und verschalten. Es müssen also Zuleitungen über die entstehenden Mesastufen geführt werden. Je höher eine Stufenkante ist, desto schwieriger gestaltet sich die Strukturierung auf ihr.

Für den Sensor, der in dieser Arbeit gefertigt werden soll, dient das Dielektrikum nicht alleinig zur elektrischen Isolation. Gegebenfalls soll dieser Sensor noch um eine Heizung erweitert werden. Dazu soll der Chip von der Rückseite her verdünnt werden, um nur wenig Material stehen zu lassen, sodass die sensitive Fläche von hinten her ohne große Energiezufuhr zu beheizen ist. Dafür soll das vergrabene SiO₂ als Ätzstopp und als tragende Membran dienen. Da diese eine genügende mechanische Festigkeit aufweisen muss, kann die Stärke dessen nicht beliebig dünn gewählt werden.

Weiter könnte man fordern, dass die Siliziumschicht eine möglichst homogene Dicke besitzt. Wenn Bauelemente auf einem SOI-Substrat freigestellt werden sollen, d. h. wenn diese frei auf dem Dielektrikum stehen sollen, ist es wichtig, dass überall auf dem Wafer bis auf das vergrabene Siliziumdioxid herunter geätzt wird, sonst leidet darunter die Ausbeute. Ist die Dicke des Device-Layers nicht gleichmäßig über dem Wafer verteilt, müssen an den dünneren Stellen lokal die Gräben und Flächen, die bis zum BOX reichen sollen, überätzt werden. Das führt

	ELTRAN	SIMOX	UNIBOND	Bonded SOI
Variation der SOI-Dicke ^a	±5 % (±2,7 nm)	±2,5 % (±3 nm)	±2,5 % (±5 nm)	±25 % (±500 nm)
Variation der BOX-Dicke	±10 nm	±10 nm	±10 nm	±10 nm
Rauigkeit	0,107 nm	0,7 nm	0,15 nm	0,7 nm
HF-Defekte	0,05 cm ⁻²	0,5 cm ⁻²	0,1 cm ⁻²	
Secco-Defekte	764 cm ⁻²	1000 cm ⁻²	1000 cm ⁻²	
Partikel pro Wafer	13 ^b	8 ^c		10 ^d
Metall-kontamination	5 · 10 ¹⁰ cm ⁻²	1,5 · 10 ¹¹ cm ⁻²	1 · 10 ¹⁰ cm ⁻²	
Literatur	[Elt03][YS01]	[ibi03][BDD ⁺ 03]	[soi03]	[Dan03]

^ain Klammern werden Absolutwerte für dünnste Schichten wiedergegeben

^bPartikel > 0,15 μm

^cPartikel > 0,3 μm

^dPartikel > 0,2 μm

Tabelle B.1: Spezifikationen verschiedener SOI-Herstellungsverfahren im Vergleich

zu Unterätzungen, sodass die Maßhaltigkeit eventuell nicht eingehalten werden kann.

Zusammenfassend hat das Substrat, das für den hier zu entwickelnden Sensor Verwendung geeignet ist, einen möglichst dünnen Device-Layer mit einer Schichtdicke nicht größer als 2 μm . Günstiger wären allerdings nur einige hundert Nanometer. Das BOX ist mindestens 2 μm . Soll der Sensor von hinten her geätzt werden, sodass eine Membran entsteht, muss dieses sogar noch um einiges dicker sein. Von Vorteil wäre dazu ein Sputter-Oxid, das durch die niedrigeren Temperaturen bei der Herstellung nicht so verspannt ist. In Abbildung B.6 ist dargestellt, welche SOI- und BOX-Schichtdicken Substrate der verschiedenen SOI-Herstellungsverfahren abdecken. Wie daraus hervorgeht, ist für die Fertigung des Gassensors ELTRAN- und UNIBOND-Material am besten geeignet. Da während der Prozessierung epitaktische Schichten gewachsen werden, sind die ELTRAN-Wafer dabei den anderen, aufgrund ihrer geringeren Rauigkeit und Defektdichte, vorzuziehen.

Leider werden derzeit ELTRAN-Substrate nur in Größen ab 125 mm gefertigt. Sie scheiden daher von vornherein aus, weil die Anlagen am Institut für Nanotechnologie und Mikrosystemtechnik der Universität der Bundeswehr auf die Bearbeitung von 100-mm-Wafern ausgelegt sind. UNIBOND-Substrate werden zwar mit passendem Durchmesser verkauft, allerdings nur in größeren Chargen. Die Abnahme einer solchen Menge würde den Bedarf zur Forschung und die finanziellen Möglichkeiten im Rahmen dieser Arbeit bei weitem überschreiten. Als Restposten oder in kleinen Stückzahlen sind v. a. gebondete Substrate auf dem Markt vorhanden. Für die weiteren Untersuchungen werden deshalb diese Substrate herangezogen und verwendet.

Für die weiteren Untersuchungen werden deshalb SOI-Wafer der Firma ULTRASIL CORPORATION herangezogen. Diese Substrate werden wie oben beschrieben aus zwei Siliziumscheiben gebondet und anschließend auf die gewünschte Dicke geschliffen. Das geschieht mittels einer in situ Messung während des Schleifens. Daher ist die angegebene Schichtdicke sehr ungenau ($\pm 25\%$). Deutlich ist dies beim Betrachten des Produktes anhand der farbigen Schlieren sowie den parallelen aber unförmigen Linien im Silizium des Device-Layers zu sehen. Auch zwischen BOX und SOI eingeschlossene Partikel sind so zu erkennen. Um sie herum zeigen sich konzentrische Ringe. Diese haben einen Durchmesser von etwa einem Zentimeter. Das Oxid wird thermisch gewachsen und besitzt daher eine sehr homogene Dicke über den Durchmesser. In Tabelle 5.1 werden die Parameter der in dieser Arbeit verwendeten Substrate aufgeführt.

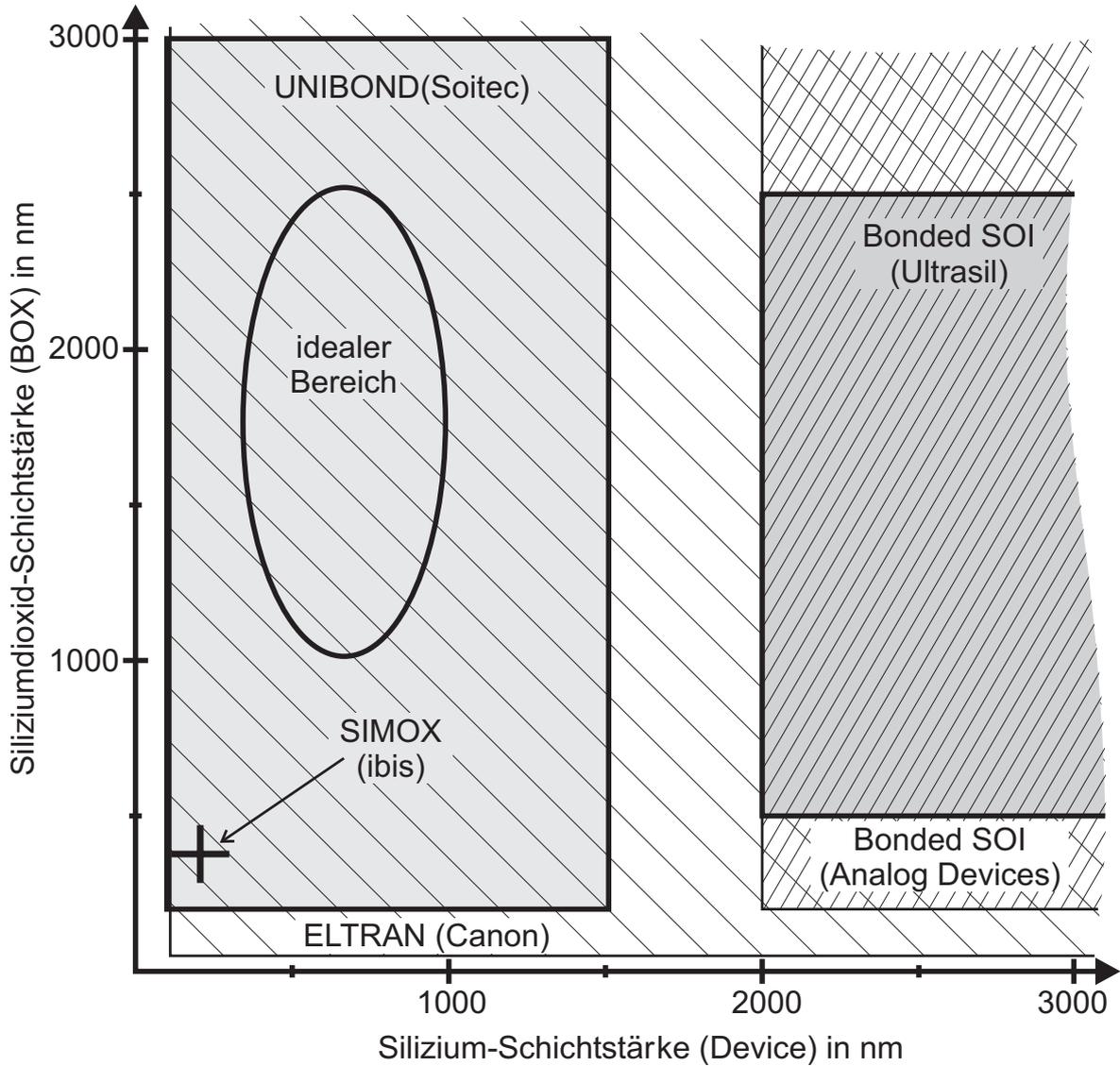


Abbildung B.6: Verfügbarkeit der verschiedenen Schichtdicken-Kombinationen aus SOI und BOX nach Herstellungsverfahren (und Herstellern)

Anhang C

Prozessablauf der Fotolithografie

Prozessschritt	Anlage	Bemerkung
1. Spülen	Rinser	
2. Trocken	Trockenschleuder	
3. Haftvermittler	HMDS-Becken	Dauer: 5 min
4. Belackung	Lackschleuder	AR-P 3840/2 1. 400 rpm; Rampe 0; 15 s 2. 8000 rpm; Rampe 9; 25 s
6. Prebake	Hotplate	Temperatur: 100 °C Zeit: 2 min
5. Abkühlen		
7. Belichten	Karl Süss MA4	Modus: Vacuum Contact Zeit: 18 s
8. Postexposure Bake	Hotplate	Temperatur: 110 °C Zeit: 4 min
9. Abkühlen		
10. Entwickeln	Becherglas	AR-300-47 unverdünnt Zeit: 55 s
11. Entwicklungsstopp	Becherglas	DI-Wasser
12. Spülen	Rinser	
13. Trocken	Trockenschleuder	

Tabelle C.1: Standard-Prozess für die Fotolithografie mit AR-P 3840/2

Prozessschritt	Anlage	Bemerkung
1. Spülen	Rinser	
2. Trocken	Trockenschleuder	
3. Haftvermittler	HMDS-Becken	Dauer: 5 min
4. Belackung	Lackschleuder	AR-U 4030 1. 400 rpm; Rampe 0; 15 s 2. 4000 rpm; Rampe 9; 25 s
6. Prebake	Hotplate	Temperatur: 95 °C Zeit: 2 min
5. Abkühlen		
7. Belichten	Karl Süss MA4	Modus: Vacuum Contact Zeit: 35 s
8. Entwickeln	Becherglas	AR-300-49 : H ₂ O Verhältnis 3:1 Zeit: 3 min
9. Entwicklungsstopp	Becherglas	DI-Wasser
10. Spülen	Rinser	
11. Trocken	Trockenschleuder	
12. Postbake	Hotplate	Temperatur: 110 °C Zeit: 4 min

Tabelle C.2: Standard-Prozess für die Fotolithografie mit AR-U 4030

Anhang D

Blende für das Alusputtern

Ohne eine Blende ist die Schichtdickenverteilung auf dem Wafer nach dem Sputtern nicht homogen. Das hat hauptsächlich drei Gründe. Da der Sputterfleck rund ist, verbleiben nicht alle Stellen, eines sich darunter hinweg bewegenden Substrates, gleich lange unter dem Abscheidebereich. Aufgrund der exzentrischen Rotation des Substrattellers haben verschiedene Punkte auf dem Wafer unterschiedliche Bahngeschwindigkeiten. Schließlich ist die Abscheiderate über den Sputterfleck nicht konstant. Man kann mit Hilfe dieser Fakten eine gute quantitative Näherung für die resultierende Schichtdickenverteilung über den Wafer angeben. Numerisch lässt sich mit dieser Näherung errechnen, welche Zeitspanne sich die einzelnen Stellen auf dem Wafer kürzer unter dem Abscheidebereich befinden müssten, um eine bestimmte homogene Schichtdicke zu erzielen. Durch teilweises Abschatten des Sputterfleckes wird die Gesamtzeit unter diesem verkürzt. Eine wie in Abbildung D.1 auf Seite 166 geformte Blende gleicht diese unterschiedlichen Verweildauern und Sputterraten aus. Dadurch wird die Schichtdicke homogenisiert.

Literaturverzeichnis

- [ACG⁺82] ARMIGLIATO, A. ; CELOTTI, G. ; GARULLI, A. ; GUERRI, S. ; OSTOJO, P. ; ROSA, R. ; MARTINELLI, G.: Characterization of titanium nitride films deposited onto silicon. In: *Thin Solid Films* 92 (1982), S. 341–346
- [ÅCS⁺02] ÅBOM, A. Elisabeth ; COMINI, Elisabetta ; SBERVEGLIERI, Giorgio ; HULMAN, Lars ; ERIKSSON, Mats: Thin oxide films as surface modifiers of MIS field effect gas sensors. In: *Sensors and Actuators, B Chemical* (2002), Nr. B 85, S. 109–119
- [Ada83] ADAMS, A. C.: Dielectric and Polysilicon Film Deposition. In: SZE, S. M. (Hrsg.): *VLSI Technology*. New York St. Louis : McGraw-Hill Book Company, 1983, Kapitel 3, S. 93–129
- [AMR03] ASCHKE, Lutz ; MENGEL, Stefan ; RAU, Jenspeter: Lithographie am Limit. In: *c't* 13 (2003), Jun., S. 198–207
- [BDD⁺03] BLAKE, J. ; DEMPSEY, K. ; DOLAN, R. ; EROKHIN, Y. ; POWELL, P. ; RICHARDS, S. ; IBIS TECHNOLOGY CORPORATION (Hrsg.): 300 mm Ultra-Thin Advantox MLD SIMOX Wafers Manufactured Using i2000 Oxygen Implanter / Ibis Technology Corporation. 32 Cherry Hill Drive, Danvers, MA 01923, 2003. – Produktpräsentation. – 2 S.
<http://www.ibis.com/assets/pdf/intro-simox-100.pdf>
- [BFK⁺03] BASELT, D. R. ; FRUHBERGER, B. ; KLASSEN, E. ; CEMALOVIC, S. ; JR., C. L. B. ; PATEL, S. V. ; MLSNA, T. E. ; MCCORKLE, D. ; WARMACK, B.: Design and performance of a microcantilever-based hydrogen sensor. In: *Sensors and Actuators, B Chemical* (2003), Nr. B 88, S. 120–131
- [Bög98] BÖGNER, Martin: *Low Power-Halbleiter-Gassensoren auf Metalloxidbasis*. Neubiberg, Universität der Bundeswehr München, Diss., 1998
- [BTBv⁺99] BECKER, Th. ; TOMASI, L. ; BOSCH-V.BRAUNMÜHL, Chr. ; MÜLLER, G. ; SBERVEGLIERI, G. ; FAGLI, G. ; COMINI, E.: Ozone detection using low-power-consumption metal-oxide gas sensors. In: *Sensors and Actuators, A Physical* (1999), Nr. A 74, S. 229–232

- [BTV⁺00] BÉVENOT, X. ; TROUILLET, A. ; VEILLAS, C. ; GAGNAIRE, H. ; CLÉMENT, M.: Hydrogen leak detection using an optical fibre sensor for aerospace applications. In: *Sensors and Actuators, B Chemical* (2000), Nr. B 67, S. 57–67
- [Bur03] BURGMAIR, Markus M.: *Einsatz von Metalloxiden in Gas-Sensoren*. Berlin, Universität der Bundeswehr München, Diss., 2003. – 150 S
- [CBL00] CAMPOS, M. ; BULHÕES, L. O. S. ; LINDINO, Cleber A.: Gas-sensitive characteristics of metal/semiconductor polymer Schottky device. In: *Sensors and Actuators, A Physical* (2000), Nr. A 87, S. 67–71
- [CMP⁺04] CHAPPÉ, J.-M. ; MARTIN, N. ; PIERSON, J.F. ; TERWAGNE, G. ; LINTYMER, J. ; GAVOILLE, J. ; TAKADOUM, J.: Influence of substrate temperature on titanium oxynitride thin films prepared by reactive sputtering. In: *Applied Surface Science* 225 (2004), S. 29–38
- [Col02] COLINGE, Jean-Pierre: *Silicon-On-Insulator Technologie : Materials to VLSI, 2nd Edition*. 2. Boston, Dordrecht, London : Kluwer Academic Publisher, 2002
- [CPG99] CHUN, J.-S. ; PETROV, I. ; GREENE, J.E.: Dense fully 111-textured TiN diffusion barriers: Enhanced lifetime through microstructure control during layer growth. In: *Journal of Applied Physics* 86 (1999), Oktober, Nr. 7, S. 3633–3641
- [Dan03] DANCOVICH, John. -. E-mail, gesendet am 03.07.03, Ultrasil Corporation, Silicon Valley, California. 2003
- [DeJ98] DEJULE, Ruth: Trends in Wafer Cleaning. In: *Semiconductor International* (1998), Aug., S. 64–68
- [Dub03] DUBBE, Andreas: Fundamentals of solid state ionic micro gas sensors. In: *Sensors and Actuators, B Chemical* (2003), Nr. B 88, S. 138–148
- [DV03] DESPONT, Michel ; VETTINGER, Peter: Micro/Nanosystem Technology for Probe-based Data Storage. In: *mstnews* 19 (2003), Juni, Nr. 3/03, S. 10–12
- [EDH94] ENDRES, H.-E. ; DROST, S. ; HUTTER, F.: Impedance spectroscopy on dielectric gas sensors. In: *Sensors and Actuators, B Chemical* (1994), Nr. B 22, S. 7–11
- [EE97] ERIKSSON, M. ; EKEDAHL, L.-G.: The influence of CO on the response of hydrogen sensitive Pd-MOS devices. In: *Sensors and Actuators, B Chemical* (1997), Nr. B 42, S. 217–223
- [Eis01] EISELE, Ignaz: *Grundlagen der Silizium-Halbleitertechnologie*. Überarbeitete Fassung 2001. Neubiberg, 2001

- [EKA⁺00] EVANGELOU, E.K. ; KONOFOAS, N. ; ASLANOGLU, X. A. ; DIMITRIADIS, C.A. ; PATSALAS, P. ; LOGOTHETIDIS, S. ; M. KOKKORIS ; KOSSIONIDES, E. ; VLASTOU, R. ; GROETSCHER, R.: Characterization of magnetron sputtering deposited thin films of TiN for use as a metal electrode on TiN/SiO₂/Si metal-oxide-semiconductor devices. In: *Journal of Applied Physics* 88 (2000), Dezember, Nr. 12, S. 7192–7196
- [Elt03] CANON INC. (Hrsg.). *SOI-Epi wafers ELTRAN*. http://www.canon.co.jp/Eltran/elt_sample-e.html. - Abgerufen: 01.07.03. - E-Mail: eltran@chcc.canon.co.jp. - 6770 Tamura, Hiratsuka-shi, Kanagawa-pref., 254 Japan. 2003
- [EMD92] ENDRES, H.-E. ; MICKLE, L. D. ; DROST, S.: A gas sensor system with dielectric and mass sensors. In: *Sensors and Actuators, B Chemical* (1992), Nr. B 6, S. 285–288
- [Feh97] FEHLAUER, Gerd: *Gitterangepaßtes ternäres Silizid für die Silizium-Technologie*. Neubiberg, Universität der Bundeswehr München, Diss., Okt. 1997
- [Fin00] FINK, Christoph: *Vertikale Leistungs-MOSFETs mit Delta Dotierung*. Neubiberg, Universität der Bundeswehr München, Diss., Nov. 2000
- [Fli93] FLIETNER, Bertrand: *Gassensoren auf der Basis von Feldeffekttransistoren mit Luftspalt*. Neubiberg, Universität der Bundeswehr München, Diss., 1993
- [Fra83] FRASER, D. B.: Metallization. In: SZE, S. M. (Hrsg.): *VLSI Technology*. New York St. Louis : McGraw-Hill Book Company, 1983, Kapitel 9, S. 347–384
- [GKV86] GERTHSEN, Christian ; KNESER, H.O. ; VOGEL, Helmut: *Physik*. 15. Auflage. Berlin Heidelberg New York Tokyo : Springer-Verlag, Feb. 1986. – 920 S
- [gme71] Titan. In: GMELIN-INSTITUT (Hrsg.): *Gmelins Handbuch der Anorganischen Chemie*. 8. Auflage. Weinheim/Bergstasse : Verlag Chemie GmbH, 1971
- [GOT86] GIBSON, J. M. (Hrsg.) ; OSBOURN, G. C. (Hrsg.) ; TROMP, R. M. (Hrsg.): *Layered Structures and Epitaxy*. Bd. 56. Materials Research Society, Jun. 1986 (Materials Research Society Symposium Pro)
- [GRMS00] GRACIAS, A. C. ; RIOS, A. N. ; MAIA, I. A. ; SENNA, J. R.: Experiments on Silicon-To-Silicon Direct Bonding. In: *Revista Brasileira de Aplicações de Vácuo* 19 (2000), Nr. 19, S. 19–22

- [Gro67] GROVE, Andrew S.: *Physics and Technology of Semiconductor Devices*. 1. New York Chichester Brisbane Toronto Singapore : John Wiley & Sons, Inc., 1967. – 366 S
- [HEÅL00] HEDBORG, E. ; ERIKSSON, M. ; ÅBOM, L. ; LUNDSTRÖM, I.: Modification of the Metal Gate to Improve the Stability of Gas Sensitive MIS Devices. In: *Proceedings Euroensors XIV*. Kopenhagen, 2000, S. 193–195
- [HOBV04] HERBER, S. ; OLTHUIS, W. ; BERGVELD, P. ; VAN DER BERG, A.: Hydrogel-based sensor for CO₂ measurements. In: *Euroensors XVIII*. Rom, Italien, 12.-15. September 2004, S. 227–228
- [Hof90] HOFFMANN, Kurt ; GOTTWALT, Alfons (Hrsg.): *VLSI-Entwurf : Modelle und Schaltungen*. 1. München, Wien, Oldenbourg : Oldenbourg Verlag GmbH, 1990 (Einführung in die Nachrichtentechnik). – 456 S. – ISBN 3-486-21206-0
- [ibi03] IBIS TECHNOLOGY CORPORATION (Hrsg.): *SIMOX and Advantox : Comparative Parameters / Ibis Technology Corporation*. 32 Cherry Hill Drive, Danvers, MA 01923, 2003 (6/17/02). – Produktpräsentation. http://www.ibis.com/assets/pdf/Ibis_Simox-400.pdf
- [Jan83] JANATA, J.: SCHUTZRECHT: Apparatus and method for measuring the concentration of components in fluids. Oktober 1983. – United States Patent, Nr 4.411.741, (25.10.1983)
- [JP92] JOSWIG, H. ; PAMLER, W.: Stoichiometry effects in TiN diffusion barriers. In: *Thin Solid Films* 221 (1992), S. 228–232
- [KA60] KAHNG, D. ; ATALLA, M. M.: Silicon - Silicon Dioxide Field Induced Surface Devices. In: *IRE-AIEE Solid-State Device Research Conference*. Pittsburgh : Carnegie Institute of Technology, 1960
- [Kat83] KATZ, L.E.: Oxidation. In: SZE, S. M. (Hrsg.): *VLSI Technology*. New York St. Louis : McGraw-Hill Book Company, 1983, Kapitel 4, S. 131–167
- [KBF⁺03] KINTTEL, Thorsten ; BURGMAYER, Markus ; FREITAG, Gunter ; ZIMMER, Martin ; EISELE, Ignaz: Combined Ammonia and Hydrogen Gas Sensor. In: *Sensors, 2003. Proceedings of IEEE* Bd. 1, IEEE, Oktober 2003, S. 191–194
- [KFP⁺87] KUMAR, N. ; FISSEL, M.G. ; POURREZAEI, K. ; LEE, B. ; DOUGLAS, E.C.: Growth and properties of TiN and TiO_xN_y diffusion barriers in silicon on sapphire integrated circuits. In: *Thin Solid Films* 153 (1987), S. 287–301

- [KLG⁺01] KIM, C. K. ; LEE, J. H. ; KANG, W. P. ; YOO, K. S. ; JANG, G. E. ; CHO, N. I.: NO₂ gas sensing characteristics of Pt – WO₃ – Si₃N₄ – SiO₂ – Si – Al capacitor. In: *Sensors and Actuators, B Chemical* (2001), Nr. B 77, S. 67–71
- [KLL⁺00] KIM, C. K. ; LEE, J. H. ; LEE, Y. H. ; CHO, N. I. ; KIM, D. J.: A study on a platinum-silicon carbide Schottky diode as a hydrogen gas sensor. In: *Sensors and Actuators, B Chemical* (2000), Nr. B 66, S. 116–118
- [KP70] KERN, Werner ; POUTINEN, David A.: Cleaning solutions based on hydrogen peroxide for use in silicon semiconductor technologie. In: *RCA Review* 30 (1970), Jun., Nr. 2, S. 187–206
- [Kuh98] KUHLMANN, Ulrike: Isomatte im Chip. In: *c't* 17 (1998), Aug., S. 28
- [LFL⁺00] LEE, K. H. ; FANG, Y. K. ; LEE, W. J. ; HO, J. J. ; CHEN, K. H. ; LIAO, K. C.: Novel electrochromic devices (ECD) of tungsten oxide (WO₃) thin film integrated with amorphous silicon germanium photo detector for hydrogen sensor. In: *Sensors and Actuators, B Chemical* (2000), Nr. B 69, S. 96–99
- [LGE⁺99] LITOVCHENKO, V. G. ; GORBANYUK, T. I. ; EFREMOV, A. A. ; EVTUKH, A. A. ; SCHIPANSKI, Dagmar: Investigation of MIS gas sensitive structures with Pd and Pd/Cu metal layers. In: *Sensors and Actuators, A Physical* (1999), Nr. A 74, S. 233–236
- [Lil30] LILIENFELD, Julius E.: SCHUTZRECHT: Methods and Apparatus for Controlling Electric Currents. Jan. 1930. – United States Patent, Nr. 1.745.175, (28.1.1930)
- [LLYL87] LEI, T.F. ; L. Y. LEU, C. L. L.: Specific Contact Resistivity Measurement by a Vertical Kelvin Test Structure. In: *IEEE Transactions on Electron Devices* 34 (1987), Nr. 6, S. 1390–1395
- [LP84] LEMPERIÈRE, G. ; POITEVIN, J. M.: Influence of the nitrogen partial pressure on the properties of D.C.-sputtered titanium and titanium nitride films. In: *Thin Solid Films* 111 (1984), S. 339–349
- [LSSL75] LUNDSTRÖM, I. ; SHIVARAMANN, M. S. ; SVENSSON, C. ; LUNDKVIST, L.: A hydrogen-sensitive MOS field-effect transistor. In: *Applied Physics Letters* (1975), Januar, Nr. 26, S. 55–57
- [MA96] MIWA, Shozo ; ARAKAWA, Tsuyoshi: Selective gas detection by means of surface plasmon resonance sensors. In: *Thin Solid Films* (1996), November, Nr. 281-282, S. 466–468
- [Mai83] MAIDHOF: *Allgemeine Hinweise für das Kathodenzerstäubungsverfahren.* : Leybold-Heraeus GmbH, 1983. – Betriebsanleitung Sputteranlage Z550

- [MM54] MORIN, F. J. ; MAITA, J.P.: Electrical Properties of Silicon Containing Arsenic and Boron. In: *Physical Review* 96 (1954), Nr. 28, S. 28–35
- [Mog83] MOGAB, C. J.: Dry Etching. In: SZE, S. M. (Hrsg.): *VLSI Technology*. New York St. Louis : McGraw-Hill Book Company, 1983, Kapitel 8, S. 303–345
- [PDO93] POLLAK-DIENER, Gerhard ; OBERMEIER, Ernst: Heat-conduction microsensor based on silicon technology for the analysis of two- and three-component gas mixtures. In: *Sensors and Actuators, B Chemical* (1993), Nr. B 13-14, S. 345–347
- [PVC⁺02] PUIGCORBÉ, J. ; VILÀ, A. ; CERDÀ, J. ; CIRERA, A. ; GRÀCIA, I. ; MORANTE, C. Cané J. R.: Thermo-mechanical analysis of micro-drop coated gas sensors. In: *Sensors and Actuators, A Physical* (2002), Nr. A 97-98, S. 379–385
- [RBB91] VAN ROOSMALEN, A. J. ; BAGGERMAN, J. A. G. ; BRADER, S. J. H.: *Dry Etching for VLSI*. 1. New York : Kluwer Academic/Plenum Publishers, März 1991 (Updates In Applied Physics And Electrical Technology). – 254 S. – ISBN 0–306–43835–6
- [RS60] ROBBINS, H. ; SCHWARTZ, B.: Chemical Etching of Silicon. In: *Journal of the Electrochemical Society* 107 (1960), Februar, Nr. 2, S. 108–111
- [SA02] SIMON, Isolde ; ARNDT, Michael: Thermal and gas-sensing properties of a micro machined thermal conductivity sensor for the detection of hydrogen in automotive applications. In: *Sensors and Actuators, A Physical* (2002), Nr. A 97-98, S. 104–108
- [Sch83] SCHUTZ, R. J.: TiN as a diffusion barrier between CoSi₂ or PtSi and Aluminium. In: *Thin Solid Films* 104 (1983), S. 89–99
- [Sch02] SCHARNAGL, Klaus: *Feldeffekttransistoren mit Luftspalt für den Nachweis von Wasserstoff*. Neubiberg, Universität der Bundeswehr München, Diss., 2002
- [SCMN80] v. SEEFELD, Hermann ; CHEUNG, Nathan W. ; MÄENPÄÄ, Martti ; NICOLET, Marc-A.: Investigation of Titanium-Nitride Layers for Solar-Cell Contacts. In: *IEEE Transactions on Electron Devices* 27 (1980), April, Nr. 4, S. 227–229
- [Sed03] SEDLMAIER, Stefan: *Vertikaler Tunnel-Feldeffekttransistor auf Silizium*. Neubiberg, Universität der Bundeswehr München, Diss., 2003
- [SFS85] SINKE, W. ; FRIJLINK, G.P.A. ; SARIS, F.W.: Oxygen in titanium nitride diffusion barriers. In: *Applied Physics Letters* 47 (1985), September, Nr. 5, S. 471–473

- [Sha02] SHAHIDI, G. G.: SOI technology for the GHz era. In: *IBM journal RES. & DEV.* (2002), march/may, Nr. 2/3, S. 121–131
- [SKI⁺87] SUGAHARA, K. ; KUSUNOKI, S. ; INOUE, Y. ; NISHIMURA, T. ; AKASAKA, Y.: Orientation control of the silicon film on insulator by recrystallization. In: *Journal of Applied Physics* 62 (1987), November, Nr. 10, S. 4178–4181
- [soi03] SOITEC SA (Hrsg.). *Unibond presentation*. http://www.soitec.com/products/p_1.html. - Abgerufen: 01.07.03. - E-Mail: sales@soitec.fr. - Parc Technologique des Fontaines, 38190 Bernin, Frankreich. 2003
- [SPL⁺04] SIMON, E. ; POHLE, R. ; LAMPE, U. ; FLEISCHER, M. ; MEIXNER, H. ; KNITTEL, T. ; EISELE, I. ; FRERICHS, H.-P. ; WILBERTZ, C. ; LEHMANN, M.: GasFET for the Detection of Ammonia. In: *Euroensors XVIII*. Rom, Italien, 12.-15. September 2004, S. 12–13
- [SY94] SATO, Nobuhiko ; YONEHARA, Takao: Hydrogen annealed silicon-on-insulator. In: *Applied Physics Letters* 65 (1994), Oktober, Nr. 15, S. 1924–1926
- [SYS⁺99] SAKAGUCHI, K. ; YANAGITA, K. ; SUZUKI, H. ; OHMI, K. ; YONEHARA, T.: ELTRAN by Splitting Porous Si Layers. In: HEMMENT, F. (Hrsg.): *Silicon-on-Insulator Technology and Devices IX*, 1999 (The Electrochemical Society Proceedings Series, PV 99- 3), S. 117–121
- [Sze69] SZE, S.M.: *Physics of Semiconductor Devices*. 1. New York London Sydney Singapore : John Wiley & Sons, Inc., 1969. – 812 S
- [TP91] TANDON, U. S. ; PANT, B. D.: The Characteristics of Reactive Ion Etching of Polysilicon Using SF₆/O₂ and Their Interdependence. In: *Vacuum* 42 (1991), Nr. 13, S. 837–843
- [WMF88] WIDMANN, Dietrich ; MADER, Hermann ; FRIEDRICH, Hans ; HEYWANG, W. (Hrsg.) ; MÜLLER, R. (Hrsg.): *Halbleiter-Elektronik*. Bd. 19: *Technologie hochintegrierter Schaltungen*. 1. Berlin Heidelberg : Springer-Verlag, 1988. – 320 S
- [YLR⁺01] YEO, Yee-Chia ; LU, Qiang ; RANADE, Pushkar ; TAKEUCHI, Hideki ; YANG, Kevin J. ; POLISHUK, Igor: Dual-Metal Gate CMOS Technology with Ultrathin Silicon Nitride Gate Dielectric. In: *IEEE Electron Device Letters* 22 (2001), Mai, Nr. 5, S. 227–229
- [YS01] YONEHARA, Takao ; SAKAGUCHI, Kiyofumi: Cutting Edge 2: ELTRAN; Novel SOI Wafer Technology. In: *JSAP International* (2001), Juli, Nr. 4, S. 10–16

Danksagung

Diese Arbeit entstand während meiner Tätigkeit als wissenschaftlicher Mitarbeiter am Institut für Physik der Fakultät der Elektrotechnik der Universität der Bundeswehr München. An dieser Stelle möchte ich mich bei allen bedanken die zum Gelingen dieser Arbeit beigetragen haben.

Ich bedanke mich besonders bei meinem Doktorvater Herrn Prof. Dr. Ignaz Eisele für die Möglichkeit an seinem Institut als wissenschaftlicher Mitarbeiter diese Arbeit anzufertigen. Sein Vertrauen, seine stets hilfreichen Diskussionen, sowie sein unermüdlicher Zuspruch ließen diese Arbeit erst zustande kommen.

Herzlich möchte ich mich bei Herrn Dipl.-Phys. Thorsten Knittel bedanken. Ohne ihn wären die Messungen für die Charakterisierung zum Ende der Arbeit hin wohl nicht so schnell geglückt. Die Zusammenarbeit machte eine Menge Spaß und brachte eine Fülle an Ideen hervor.

Spezieller Dank gilt Herrn Dipl.-Ing. Mathias Born für die Hilfe bei der Entwicklung der „glatten Flanken“. Auch das freundschaftliche Klima im gemeinsamen Büro mit den ausgiebigen Diskussionen war äußerst angenehm und inspirierte immer wieder zu neuen Einfällen und Ansätzen.

Bei Dr. Torsten Sulima und Dr. habil. Jörg Schulze möchte ich mich für das Korrekturlesen dieser Arbeit und die anschließende Diskussion bedanken.

Bei all meinen weitem Kollegen möchte ich mich für das Herstellen von Schichten, die Durchführung von Messungen sowie sonstige Hilfestellungen oder Gefallen in meiner täglichen Arbeit bedanken. Sie sorgten immer wieder unbeirrt dafür, dass ich ein Krümelchen in meinen Proben gefunden habe oder mir neue Chemikalien besorgen konnte. Namentlich sind dies: Ulrich Abelein, Dr. K. G. Anil, Kwanchai Anothainart, Prof. Dr. Hermann Baumgärtner, Anton Bayerstadler, Krishna Kumar Bhuwalka, Dr. Markus Burgmair, Vanessa Capodieci, Peter Ciecierski, Thomas Galonska, Dr. Alexandra Ludsteck, Hans Messarosch, Dr. Herbert Pollak, Andreas Rippler, Dr. Klaus Scharnagl, Markus Schindler, Matthias Schmidt, Dr. Stefan Sedlmaier, Oliver Senftleben, Dr. Tanja Stimpel-Lindner, Carolin Tolksdorf, Wahyuwidanarto, Dr. Florian Wiest, Dr. Martin Zimmer. Die mittäglichen Gespräche zu allen Fragen des täglichen Lebens regten immer wieder zu neuen Gedanken an.

Großen Anteil am Gelingen dieser Arbeit trägt zweifelsohne meine Frau Selma, die mir viel Verständnis entgegenbrachte und mir einen beträchtlichen Teil meiner Aufgaben im privaten Bereich abnahm. Dafür herzlichen Dank.

