

Materialien für die Miniaturisierung

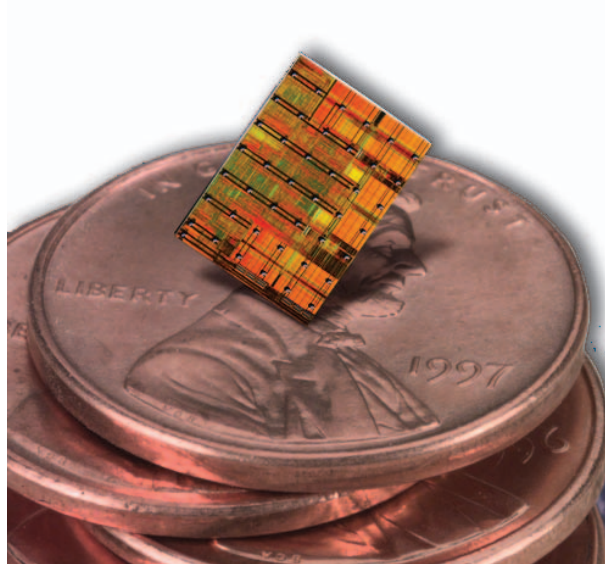
Maßgeschneiderte Materiallösungen für die Mikroelektronik

H. Jörg Osten

Immer komplexeren integrierten Schaltungen verdanken wir leistungsfähigere und schnellere Computer. Über viele Jahre blieben der grundsätzliche Aufbau der Bauelemente und die verwendeten Materialien nahezu unverändert und es genügte im Wesentlichen, die Strukturmaße zu verringern. Für eine Fortschreibung dieser Erfolgsgeschichte der Mikroelektronik bedarf es aber zunehmend neuartiger Materialien.

Die Geschichte der Mikroelektronik begann vor nicht einmal 60 Jahren mit dem ersten Germanium-Transistor. Seit Ende der fünfziger Jahre die Voraussetzungen für integrierte Schaltungen geschaffen wurden, erleben wir eine höchst dynamische Entwicklung. Bereits 1965 sagte Gordon Moore voraus, dass sich die Anzahl von Transistoren pro Chip etwa alle 18 Monate verdoppeln wird [1]. Das „Moore'sche Gesetz“ gilt noch heute; die neuesten Chipgenerationen enthalten mehrere hundert Millionen Transistoren pro Chip. Nehmen wir an, die Größe eines Chips bleibt konstant. Um die Anzahl der sich darauf befindlichen Bauelemente zu verdoppeln, muss ihre Fläche halbiert werden. Das bedeutet, alle Strukturmaße auf dem Chip mit dem Faktor 0,7 zu skalieren ($0,7 \times 0,7 \sim 0,5$). Daraus ergeben sich die allgemein bekannten sog. Technologiegenerationen mit Strukturbreiten von 0,35 μm , 0,25 μm , 0,18 μm , 0,13 μm , 90 nm, 65 nm usw., die bislang die weitere Gültigkeit des Mooreschen Gesetzes sicherstellen.

Die konsequente Anwendung dieser Skalierung führt zu einer exponentiellen Zunahme der Anzahl der Transistoren pro Chip als Funktion der Zeit. Gleichzeitig wächst die Geschwindigkeit der Transistoren und der daraus hergestellten Schaltungen ebenfalls exponentiell: Während der 1973 hergestellte Intel-Prozessor 4004 noch langsamer als 1 MHz war, wurde nur 25 Jahre später bereits die GHz-Grenze durchbrochen (eine Steigerung um mehr als den Faktor 1000!). Damit die neuen Chipgenerationen dennoch bezahlbar bleiben, muss auch der Preis pro Komponente exponentiell sinken. Die Produktionskosten eines Transistors bewe-



Zu den vergleichsweise neuen Materialien in hochintegrierten Computerchips gehört Kupfer, das in den letzten Jahren Aluminium als Material für die Leiterbahnen abgelöst hat. (Quelle: IBM)

gen sich heutzutage im Bereich weniger „Nanodollar“. Zum Vergleich: 1980 dauerte ein Flug von Europa an die Ostküste der USA ca. 7 Stunden und kostete 800 \$. Hätte sich das Flugwesen seit dieser Zeit mit der gleichen Dynamik entwickelt wie die Mikroelektronik, so würden wir heute mit annähernder Lichtgeschwindigkeit in weniger als einer Sekunde für weniger als einen Cent fliegen können.

Bis zum Anfang der 90er-Jahre bestand Skalieren tatsächlich im Wesentlichen darin, die Strukturmaße zu reduzieren. Der grundsätzliche Aufbau der Bauelemente sowie die verwendeten Materialien blieben nahezu unverändert. Um die Miniaturisierung in der Mikroelektronik fortführen zu können, war in den letzten Jahren jedoch der zunehmende Einsatz neuer oder modifizierter Materialien notwendig.

KOMPAKT

- ▶ Neue Materiallösungen sind erforderlich, um mikroelektronische Schaltungen noch weiter miniaturisieren zu können. Dabei gilt es die nötigen Prozess- und Fertigungsschritte parallel mitzuentwickeln.
- ▶ Benötigt werden verstärkt Materialien, deren Eigenschaften sich durch Kombination verschiedener Komponenten genau auf den Anwendungszweck einstellen lassen.
- ▶ Ein Beispiel sind verspannte Silizium-Germanium-Mischschichten, bei denen sich die elektronischen und optischen Eigenschaften durch Wahl der Germaniumkonzentration gezielt einstellen lassen.

Verspannte Mischschichten

Eine wichtige Gruppe neuer Materialien bilden Heterostrukturen, die aus einer Abfolge von verschiedenen Halbleitermaterialien mit unterschiedlichen Bandlücken bestehen. Die Vielzahl an realisierbaren Verbindungen eröffnet weite Variationsmöglichkeiten, um die wichtigsten elektronischen und optischen Eigenschaften gezielt einzustellen. Für Heterostrukturen auf Siliziumbasis bieten sich in erster Linie andere Elemente der 4. Hauptgruppe an. Da sich die Atomradien von Kohlenstoff und Silizium um

Prof. Dr. H. Jörg Osten, Institut für Materialien und Bauelemente der Elektronik, Universität Hannover, Appelstr. 11A, 30167 Hannover

ca. 40 % unterscheiden, lassen sich diese Elemente im festen Zustand nicht mischen. Silizium-Kohlenstoff-Verbindungen können im Gleichgewicht nur als stöchiometrisches Siliziumkarbid (SiC) existieren, wobei die Unterschiede in den Gitterstrukturen die Herstellung von Heterostrukturen äußerst erschweren.

Germanium hingegen besitzt wie Silizium eine Diamantstruktur mit einer nur um 4,2 % größeren Gitterkonstante. Daher lassen sich beide Materialien unbegrenzt (isomorph) mischen. Die so entstehenden $\text{Si}_{1-x}\text{Ge}_x$ -Verbindungen ($0 < x < 1$) bilden eine völlig neue Klasse von Halbleitermaterialien, bei denen sich durch die Wahl von x die gewünschten elektronischen Eigenschaften wie Bandlücke und Bandanpassung einstellen lassen. Reines Germanium besitzt eine deutlich höhere Elektronen- und Löcherbeweglichkeit als Silizium; ein Vorteil, der in neuerer Zeit wieder verstärkt in den Blickpunkt rückt. Versuche, reine Ge-Schichten auf Si-Substraten zu integrieren, befinden sich allerdings noch im Forschungsstadium. Darüber hinaus ist die Bandlücke von Germanium deutlich geringer und die Eigenleitung um mehr als drei Größenordnungen höher als in Silizium [2]. Schließlich ist der Wärmeausdehnungskoeffizient von Ge doppelt so groß, der Schmelzpunkt aber geringer und liegt unterhalb üblicher Prozesstemperaturen bei der Herstellung von Si-basierenden Chips.

In den letzten Jahren ist es gelungen, verschiedene Bauelemente aus Silizium-Germanium-Schichtsystemen herzustellen. Dazu zählen u. a. Heterobipolartransistoren (HBT), Hetero-MOSFETs und -MODFETs, Tunnelioden und -transistoren, Wellenleiter, IR-Detektoren oder Photodetektoren. SiGe-HBT's finden sich bereits in Massenprodukten; die Hauptanwendungsgebiete sind schnelle und rauscharme Analogschaltungen, wie etwa im Empfänger-/Sendeteil von Handys [3].

Um einkristalline Si-Ge-Schichtsysteme herzustellen, werden sowohl chemische als auch physikalische Schichtabscheidungsverfahren eingesetzt. Die chemischen Verfahren (z. B. Chemische Gasphasenabscheidung) basieren auf der Reaktion von silizium- und germaniumhaltigen Verbindungen (z. B. SiH_4 und GeH_4) direkt auf dem Siliziumsubstrat. Zu den physikalischen Verfahren gehört als wichtigstes die Molekularstrahl-Epitaxie. Bei dieser Kristallwachstumsmethode treffen orientierte

Atom- bzw. Molekülstrahlen im Ultrahochvakuum auf ein passendes Substrat und nehmen dort die vom Substrat vorgegebene kristalline Orientierung auf.

In einer $\text{Si}_{1-x}\text{Ge}_x$ -Legierung verteilen sich die Si- und Ge-Atome zufällig auf die Gitterplätze. Gemäß der sog. Vegardschen Beziehung variiert die Gitterkonstante des Mischkristalls nahezu linear mit x zwischen dem Si- und Ge-Wert: $a_{\text{SiGe}}(x) = x a_{\text{Si}} + (1-x) a_{\text{Ge}}$. Diese lineare Interpolation zwischen dem Si-Wert von 0,357 nm und dem Ge-Wert von 0,357 nm weicht maximal

0,125 % von der tatsächlichen Gitterkonstante ab. Aufgrund dieses Unterschieds muss sich das Gitter der aufwachsenden Schicht dem durch das Substrat vorgegebenen Gitter anpassen; es wird gegenüber seiner Volumenstruktur verspannt. Man spricht von „Strained-Layer-Epitaxy“, kohärentem oder angepassten Wachstum. Ist die Gitterfehlanpassung gering, d. h. unterscheiden sich die Gitterkonstanten wenig, nimmt die wachsende Schicht in den Netzebenen parallel zur Substratoberfläche deren Gitterkonstante an. Senkrecht dazu stellt sich eine andere Bindungslänge ein (Abb. 1). Bei diesem „pseudomorphen“ Wachstum stellt sich in der verspannten Schicht anstelle der ursprünglich kubischen Kristallstruktur eine tetragonale Einheitszelle ein.¹⁾

Pseudomorphes Wachstum verspannter Schichten ist nicht bis zu jeder beliebigen Dicke möglich. Oberhalb einer so genannten kritischen Schichtdicke (h_c) setzt eine Relaxation der Spannung durch plastische Veränderungen ein, die die Materialeigenschaften deutlich verschlechtern können. Dabei unterscheidet man verschiedene Formen zum Abbau der Spannungsenergie: Reines Germanium wächst z. B. pseudomorph nur bis zu einer Dicke von drei Monolagen auf Silizium. Danach bilden sich dreidimensionale Inseln, die durch elastische Verformung Spannungsenergie abbauen können. Alternativ können sog. Fehlanpassungsversetzungen entstehen und sich bewegen. Da Versetzungen entweder geschlossene Schleifen bilden oder aber an einer Kristalloberfläche enden müssen, müsste theoretisch eine Fehlanpassungsversetzung quer über den gesamten Wafer laufen und dort erst enden. Das ist jedoch in einem realen Kristall nicht möglich, da die Versetzungen mit anderen Versetzungen und sonstigen Kristallfehlern wechselwirken. So entstehen kurze Versetzungssegmente, die dann durch sog. Durchstoßversetzungen (Threadings) mit der Kristalloberfläche verbunden werden. Threadings stoßen durch die Schicht nach oben zur Wachstumsebene und beeinflussen daher die elektrischen Eigenschaften der genutzten Schicht negativ. Sie reduzieren z. B. als Streuzentren deutlich die erreichbare Ladungsträgerbeweglichkeiten.

SiGe-Schichten können bei relativ geringen Temperaturen auch versetzungsfrei oberhalb der kritischen Schichtdicken wachsen, da die Energie nicht für die Generierung/Bewegung der Versetzungen ausreicht. Allerdings sind sie *metastabil*, sodass bei einer nachträglichen Erwärmung, z. B. bei späteren Prozess-Schritten, Versetzungen entstehen und damit letztendlich die Ladungsträgerbeweglichkeiten verringert werden. Abbildung 2 zeigt die mit verschiedenen Modellansätzen be-

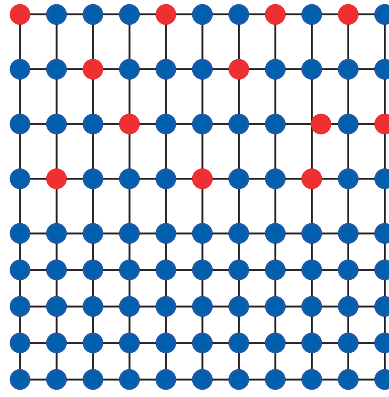


Abb. 1: Pseudomorphes Wachstum von SiGe-Schichten auf Si-Substraten (vgl. Text)

1) Während eine pseudomorph auf einem Si-Substrat gewachsene $\text{Si}_{1-x}\text{Ge}_x$ -Schicht immer unter Druckspannung steht, steht umgekehrt eine pseudomorphe Si-Schicht auf einem Substrat mit größerer Gitterkonstante (z. B. relaxiertes, d. h. nicht-verspanntes $\text{Si}_{1-x}\text{Ge}_x$) unter Zugspannung.

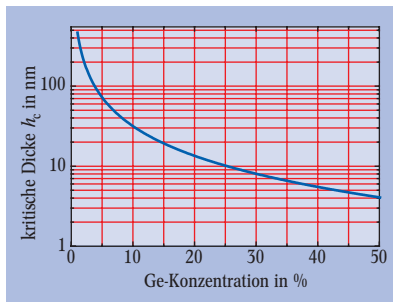


Abb. 2: Mit wachsender Germanium-Konzentration sinkt die kritische Schichtdicke von verspannten $\text{Si}_{1-x}\text{Ge}_x$ -Schichten auf Si(001).

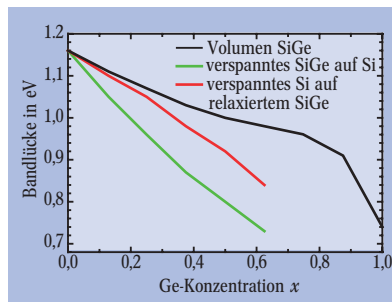


Abb. 3: Bandlücke für $\text{Si}_{1-x}\text{Ge}_x$ -Volumenmaterial (schwarz), verspanntes $\text{Si}_{1-x}\text{Ge}_x$ auf Si(100) (grün) und verspanntes Si auf relaxiertem $\text{Si}_{1-x}\text{Ge}_x$ (rot, alles bei 90 K).

rechnet maximal mögliche Schichtdicke für perfektes pseudomorphes Wachsen von Si-reichen $\text{Si}_{1-x}\text{Ge}_x$ -Schichten auf Si (mit $x < 50\%$). Eine pseudomorphe SiGe-Schicht mit 20 % Germanium ist demnach nur bis zu einer Dicke von 15 nm auf Si stabil. Diese Grenze schränkt den Entwurf von SiGe-Heterostrukturen und darauf basierenden Hetero-Bauelementen stark ein.

Pseudomorphes Wachstum von $\text{Si}_{1-x}\text{Ge}_x$ auf Si begrenzt nicht nur die erreichbare Schichtdicke und hebt die kubische Symmetrie auf, sondern beeinflusst grundsätzlich auch die Bandstruktur sowie die Bandanpassung an das Substrat (*band alignment*) und reduziert die effektiven Massen der Elektronen und Löcher in der Schichtebene. Diese Effekte sollen im Folgenden etwas detaillierter behandelt werden [4].

Elektronische Eigenschaften

Da die tetragonale Verspannung der $\text{Si}_{1-x}\text{Ge}_x$ -Schichten auch deren Bandstruktur beeinflusst, lässt sich die Bandlücke mit Hilfe der Verspannung einstellen („band gap engineering“, Abb. 3). Interessant ist auch die Anpassung der Bänder an den Heterogrenzflächen für die beiden verschiedenen verspannten Systeme (unter Druckspannung stehendes SiGe auf Si sowie unter Zugspannung stehendes Si auf Volumen-SiGe). Nicht nur der Typ des Heteroüberganges unterscheidet sich in den beiden Fällen. Das verspannte SiGe auf Si zeigt einen relativ großen Versatz (offset) im Valenzband (Abb. 4a) und ist damit insbesondere geeignet, um Löcher in sog. Quantentrogstrukturen zu lokalisieren (confinement). Eine solche Quantentrogstruktur besteht aus einer dünnen SiGe-Schicht zwischen dicken Si-Schichten. Die Valenzbandkante bildet dann einen Potentialtopf mit gebundenen, lokalisierten Zuständen für die Löcher, wie aus der eindimensionalen Quantenmechanik bekannt ist. Dagegen ist verspanntes Si auf Volumen-SiGe besser geeignet, um Elektronen zu lokalisieren (Abb. 4b). Neben der Bandstruktur ändert sich auch der Charakter des Leitungs- und des Valenzbandes deutlich. Die im unverspannten Zustand 6-fach entarteten Minipalmen in einen 4-fach und einen 2-fach entarteten Zustand auf. Die Aufhebung der sechsfachen Entartung hat folgende Konsequenzen für die Beweglichkeit der Ladungsträger:

Zum einen können die Löcher (oder Elektronen) in einem mehrfach entarteten Valenzbandmaximum (oder Leitungsbandminimum) zwischen den verschiedenen energiegleichen Zuständen hin und her springen. Diese Streuung (Intervalley Scattering), die direkt die entsprechenden Ladungsträgerbeweglichkeiten reduziert, wird bei ausreichend tetragonaler Gitterdeformation unterdrückt. Zum anderen weisen die verschiedenen entarteten Zustände im unverspannten Silizium unterschiedliche effektive Massen auf. Je flacher z. B. das Leitungsband am Minimum ist, desto größer wird die effektive Masse der Elektronen. Somit kann es „schwere“ und „leichte“ Elektronen geben. Zum Ladungstransport tragen alle Massen gewichtet nach ihren Anteilen (Zu-

standsichten) bei. Analoges gilt auch für die Löcher im Valenzband. Die effektive Masse der „schweren“ Löcher beträgt $0,65 m_0$ (m_0 ist die Masse eines freien Elektrons); die der leichten Löcher nur $0,14 m_0$. Da im tetragonal verspannten Material die Entartung aufgehoben ist, tragen z. B. im verspannten Silizium nur noch die leichten Löcher zum Stromtransport in der Schichtebe-

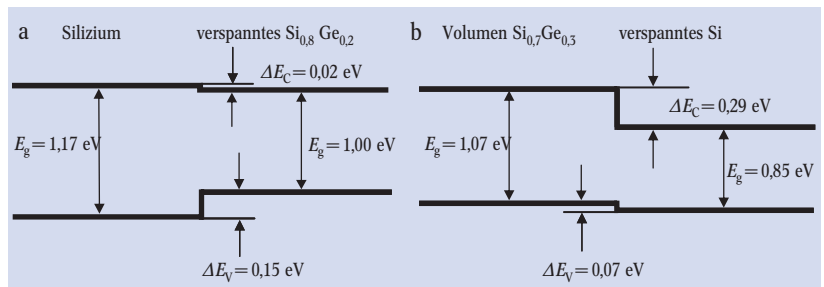


Abb. 4: Bandanpassung für verspanntes $\text{Si}_{0.8}\text{Ge}_{0.2}$ auf Si (a) und für verspanntes Si auf relaxiertem $\text{Si}_{0.7}\text{Ge}_{0.3}$ (b).

ne bei; die zum Stromfluss beitragende effektive Masse wird deutlich kleiner. Dasselbe gilt für Elektronen.

Beide Effekte zusammen erhöhen die Ladungsträgerbeweglichkeiten in verspannten Schichten signifikant. Bei verspanntem Silizium gelang es bereits, eine um mehr als 80 % höhere Elektronenbeweglichkeit nachzuweisen. Andererseits führt die Variation in der atomaren Zusammensetzung in einer SiGe-Schicht zu zusätzlicher Legierungsstreuung, welche die Beweglichkeiten reduzieren kann. Optimal für Hochbeweglichkeitskanäle wären daher verspannte reine Schichten aus Si oder Ge.

SiGe-Heterobipolartransistoren

Zwei gegeneinander geschaltete pn-Übergänge bilden einen Bipolartransistor. Im gebräuchlichen Fall des npn-Transistors tragen Elektronen den für das Bauelement bestimmenden Strom, im Falle des pnp-Transistors sind es Löcher. Die drei Dotierbereiche sind jeweils ohmsch kontaktiert; die Anschlüsse werden Emitter, Basis und Kollektor genannt. Beim konventionellen Silizium-Bipolartransistor lässt sich die Leitfähigkeit der Basis zur Erzielung einer höheren Diffusionsgeschwindigkeit durch weitere Dotierung nicht beliebig steigern, weil der Rückstrom von Löchern von der Basis und dem Emitter dadurch zunimmt und der Emitter-Injektionswirkungsgrad (durch Rekombination) absinkt.

Der Heterobipolartransistor (HBT, Abb. 5) ist eine Weiterentwicklung des konventionellen Bipolartransistors (BJT) [5]. Hier hat die aus $\text{Si}_{1-x}\text{Ge}_x$ bestehende Basis eine kleinere Energielücke als das Emitter- und Kollektormaterial. Durch den Bandkantensprung im Valenzband (vgl. Abb. 4a) haben die Löcher eine größere Energiebarriere zu überwinden als die Elektronen, was die Emittereffizienz erhöht. Bei sonst gleichem Design kann die Stromverstärkung eines HBT damit um Größenordnungen höher sein als die des BJT. Während darüber



Abb. 5: Querschnitt durch einen SiGe-Heterobipolartransistor mit Trench Isolation, bei der man schmale Gräben im Halbleitermaterial ätzt und diese mit einem Isolator füllt. Dadurch lassen sich die einzelnen Bauelemente dichter aneinander bringen, ohne sich gegenseitig zu beeinflussen. Die Kontakte sind (v. l.) Basis, Emitter und Kollektor. (Quelle: IBM)

hinaus im normalen BJT die Basis um einen Faktor 100 niedriger als der Emitter dotiert werden muss, kann diese Beschränkung im HBT fallen. Damit ist ein geringerer Basiswiderstand möglich und außerdem eine kleinere Basisbreite, was zu geringeren Transitzeiten der Ladungsträger durch die Basis und somit zu schnelleren Transistoren führt. Durch den Bandkantensprung lässt sich zusätzlich ein geringeres Rauschen sowie eine bessere Linearität erreichen.

Problematisch ist jedoch, dass in einem modernen HBT das zur Dotierung der äußerst dünnen Basis-schicht (< 50 nm) verwendete Bor bei späteren Prozessschritten aus der Basis herausdiffundieren kann. Dadurch wird die Basisbreite erhöht und der pn-Übergang in den Silizium-Emitter verschoben, was zu einer parasitären Injektionsbarriere führt. Um das zu vermeiden, können jene Basis-Bereiche, die an Emitter und Kollektor grenzen, undotiert bleiben, wodurch man aber die vorteilhafte hohe Basisdotierung des HBT zum Teil wieder verliert. Stattdessen lässt sich die Bor-Diffusion verhindern, indem man Kohlenstoff mit einer Konzentration ähnlich wie die Dotanten (< 10²⁰ cm⁻³) mit in die Basis einbaut. Solche SiGe:C-Bipolartransistoren für Hochfrequenzanwendungen wurden 1997 erstmalig

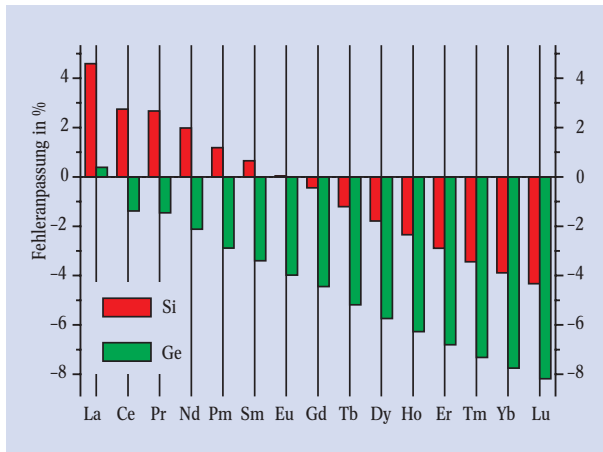


Abb. 6: Gitterfehl-anpassung von kubischen Seltenerd-Oxiden des Typs Me₂O₃ auf Si(001) und Ge(001)

mittels MBE realisiert [6]. Mittlerweile erreichen CVD-gewachsene SiGe:C-HBTs Frequenzen oberhalb von 200 GHz [7]; erste Schaltungen für 60 GHz wurden vorgestellt [8]. SiGe- bzw. SiGe:C-Heterobipolartransistoren für Hochfrequenzanwendungen, insbesondere im Bereich der drahtlosen Kommunikation, werden mittlerweile von mehreren Firmen produziert.

Alternatives Gate-Dielektrikum

Seit den fünfziger Jahren wird Siliziumdioxid (SiO₂) als Gate-Isolator in Feldeffekttransistoren (FETs) eingesetzt. Dabei verringerte sich die Dicke dieser Isolationsschicht bei jeder neuen Technologiegeneration. Während quantenmechanisches Tunneln durch diese Schicht bislang faktisch nicht stattfand, ist der exponentielle Anstieg des Tunnelstroms bei abnehmender Schichtdicke inzwischen zum fundamentalen Problem geworden [9]. So beträgt dieser Leckstrom bei einer Spannung von 1 Volt und einer 3 nm dicken SiO₂-Schicht rund 10⁻⁶ A/cm², steigt aber um sechs Größenordnungen auf rund 0,5 A/cm², wenn man die Oxiddicke um den Faktor 2 auf 1,5 nm reduziert. Gate-Isolatoren aus thermisch oxidiertem Silizium, wie sie in

der klassischen CMOS-Technologie eingesetzt werden, stoßen daher gegenwärtig an ihre physikalischen Grenzen. Da die Schichtdicken bereits wenige Atomlagen erreicht haben und SiO₂ als Gate-Isolator ausgereizt ist, wird gegenwärtig weltweit nach alternativen Gate-Isolatoren gesucht. Nach dem derzeitigen Zeitplan, der „International Technology Roadmap for Semiconductors“, werden neue Isolatoren frühestens in sechs Jahren zwingend erforderlich sein.²⁾ Da die Gate-Isolatoren recht früh im Gesamtprozess der Chipherstellung eingebracht werden (sog. Front-End-Prozess, im Unterschied zu den später besprochenen Low- κ -Polymeren oder der Cu-Metallisierung) kann nicht zwischen Material- und Prozessentwicklung getrennt werden, d. h. dass parallel zu den neuen Materialien auch die Prozesstechnologie entsprechend entwickelt werden muss.

Für eine gegebene Technologiegeneration ergibt sich die notwendige Gate-Kapazität der Feldeffekttransistoren (FETs) aus dem Gesamt-design. Um die Leckströme durch den Gate-Isolator zu verringern und gleichzeitig die Kapazität nicht zu verändern, kann man in einfachster Näherung die Beziehung $C \sim \kappa A/t$ ausnutzen, wobei C die Kapazität, κ die (relative) dielektrische Konstante des Materials, A die Fläche und t die Schichtdicke ist. Bei gleicher Kapazität skaliert die Schichtdicke linear mit der dielektrischen Konstanten κ , d. h. Materialien mit großem κ erlauben es, die Schichtdicke t zu vergrößern und somit den Leckstrom zu verringern. Der Begriff der äquivalenten Oxiddicke (Equivalent Oxide Thickness, EOT) eines beliebigen Dielektrikums mit κ und der Dicke t entspricht dabei der Dicke einer SiO₂-Schicht ($\kappa = 3,9$), die erforderlich wäre, um die gleiche Kapazität zu realisieren: $d_{EOT} = t(3,9/\kappa)$.

Eingebettet in die Randbedingungen der CMOS-Integration werden generell folgende Anforderungen an High- κ -Materialien gestellt:

► **Dielektrische Konstante:** Durch die Beimischung von Stickstoff in SiO₂ lässt sich κ in SiO(N)-Schichten maximal verdoppeln. An derartigen Lösungen wird bereits in der Industrie gearbeitet; sie sind nicht direkt als alternative High- κ -Materialien zu betrachten. Das bekannte Al₂O₃ mit $\kappa = 9$ erlaubt ebenfalls nur eine um den Faktor 2,3 höhere Isolator-dicke und scheidet damit aus längerfristigen Betrachtungen aus. Auf der anderen Seite besitzen insbesondere die verschiedenen Perovskit-Strukturen dielektrische Konstanten größer 100, sodass zur Realisierung eines d_{EOT} von 1 nm Schichtdicken von größer 25 nm notwendig wären. Dies erscheint wenig sinnvoll, da dann die Schichtdicke im Bereich der Kanallänge läge und zu nicht kontrollierbaren Gate-Feldern führen würde. Somit erscheint ein High- κ -Material mit κ zwischen 20 und 40 am aussichtsreichsten. Für einen Wert d_{EOT} von 1 nm müsste das Gatedielektrikum dann zwischen 5 nm (für $\kappa = 20$) und 10 nm ($\kappa = 40$) dick sein.

► **Bandstruktur:** Der Tunnelstrom durch ein Dielektrikum hängt neben der Dicke auch exponentiell von den Barrierehöhen zum Silizium ab. Somit wäre *a priori* ein Material mit möglichst großer Bandlücke E_g wünschenswert. Allerdings skaliert im Allgemeinen die dielektrische Konstante invers mit der Bandlücke ($E_g \sim 1/\kappa$). Somit muss ein Optimum zwischen E_g und κ gesucht werden. Anwendungen in CMOS erfordern sowohl entsprechende Bandoffsets im Leitungsband (für NMOS) als auch im Valenzband (für PMOS). Neben einer ausreichend großen Bandlücke sind somit etwa gleich große Offsets im Valenz- und im Leitungs-

2) Die Roadmap findet sich unter <http://public.itrs.net>.

band notwendig. Eine Reihe von potenziellen High- κ -Materialien (z. B. die Perovskite) erfüllt diese Symmetriebedingung nicht.

► **Stabilität auf Silizium:** Die High- κ -Materialien werden auf einkristallinem Silizium bei geeigneten Temperaturen abgeschieden. Eine Reihe von Metalloxiden, die bereits bei moderaten Temperaturen mit Silizium reagieren und z. B. Metallsilizide oder Silikate bilden, sind ebenfalls als Gate-Dielektrikum ungeeignet [10]. Die Qualität der Grenzfläche zwischen dem Gate-Isolator und dem Si-Kanal bestimmt entscheidend die Bauelemente-Tauglichkeit der Materialien. Dabei ist die Bildung von Grenzflächen-Zwischenschichten mit geringerem κ zu vermeiden, da derartige Zwischenschichten die erreichbaren minimalen äquivalenten Oxiddicken begrenzen.

► **Elektrische Parameter:** Neben den bisher genannten Materialparametern sind weitere Parameter wie die Anzahl von geladenen Störzentren, Grenzflächenzuständen, eventuelle Dipole an der Grenzfläche zum Silizium usw. für die Anwendung von High- κ -Materialien als Gate-Dielektrikum von wesentlicher Bedeutung. Derartige Schäden führen in MOS-Bauelementen u. a. zu erhöhten Gate-Leckströmen, zu reduzierten Drain-Strömen sowie erhöhten Off-Strömen, zur Abnahme der Steilheit und zu unkontrollierbaren Drifts in der Einsatzspannung. Es sei hier erwähnt, dass gegenwärtig kein Material in Sicht ist, welches in all diesen Parametern äquivalent zum SiO_2 ist [11]. Das überrascht auch nicht, wenn man die 40-jährige Entwicklung berücksichtigt, die zum jetzigen Status des SiO_2 führte. Notgedrungen werden beim Einsatz von High- κ -Gate-Isolatoren zuerst bestimmte Abstriche an einigen Leistungsparametern der Transistoren gemacht werden müssen.

► **Prozesskompatibilität:** Die Integration eines neuen Gate-Dielektrikums in einen CMOS-Frontend-Prozess stellt auch erhebliche Anforderungen hinsichtlich der Integration in existierende Fertigungsprozesse. Dazu gehören u. a. die strukturelle und elektrische Stabilität gegenüber CMOS-typischen Temperaturen in verschiedenen Atmosphären, die selektive Strukturierbarkeit (nass und/oder trocken), die Vermeidung von Cross-Kontaminationen der Prozesslinie sowie die Resistenz gegenüber möglichen prozessinduzierten Schädigungen (Implantation, RIE-Ätzen, Fotolack).

Amorphe oder epitaktische Materialien?

Weltweit werden gegenwärtig überwiegend amorphe High- κ -Materialien untersucht. Diese sehr ionischen Metalloxide besitzen relativ niedrige Kristallisationstemperaturen, bei denen sich polykristalline Phasen ausbilden. Diese Phasen sind aufgrund der hohen Leckströme entlang der Korngrenzen sowie der verstärkten Oberflächenrauigkeit als Gate-Isolator ungeeignet. Für die bekanntesten High- κ -Materialien, wie HfO_2 oder ZrO_2 , werden diese Temperaturen in einem CMOS-Prozess deutlich überschritten. Spezielle Maßnahmen zur Erhöhung der Kristallisationstemperatur (z. B. die Beimischung von Si oder Al) sind immer nur als Kompromiss zu bewerten, da sie mit einer Reduzierung von

κ und damit der möglichen Schichtdicke verbunden sind. Eine Alternative wäre der Einsatz von epitaktisch gewachsenen, kristallinen High- κ -Materialien, bei denen keine Gefahr der Rekristallisation besteht. Damit ergibt sich eine deutlich verbesserte thermische Prozessstabilität. Weiterhin ist die Grenzfläche zum Si wohldefiniert, sodass sich bessere Möglichkeiten zum Grenzflächen-Engineering ergeben. Das Wachstum epitaktischer High- κ -Materialien auf Si erfordert eine gewisse Übereinstimmung in den Atomabständen und der Gittersymmetrie und reduziert somit die Anzahl geeigneter Materialien. Neben verschiedenen epitaktischen Perovskit-Strukturen kommen außerdem einige binäre Metalloxide (insbesondere epitaktisch gewachsene binäre Selten-Erd-Oxide, Abb. 6) infrage [12]. Erste erfolgreiche Versuche zur Integration von kristallinen Selten-Erd-Oxide als Gate-Dielektrikum wurden vor kurzem vorgestellt [13].

Neue Materiallösungen im Backend-Bereich

Mehrere Kilometer schmaler Dünnschicht-Leiterbahnen verbinden derzeit die einzelnen Komponenten

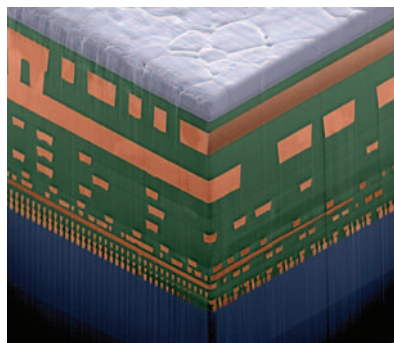


Abb. 7: Im Querschnitt eines Hochleistungs-Mikroprozessors erkennt man die Verdichtung, welche die aktiven Gebiete des integrierten Schaltkreises miteinander verbindet. (Quelle: IBM)

in einem modernen IC. Diese enorm langen Leiterbahnen verbinden in mehreren Lagen übereinander die aktiven Gebiete des ICs (Abb. 7). Trotz der geringen elektrischen Spannungen innerhalb eines ICs werden aufgrund der kleinen Querschnitte der Bahnen extrem hohe Stromdichten erreicht, bei denen herkömmliche massive Leitungsdrähte unmittelbar schmelzen würden. Leiterbahnen der Mikroelektronik hingegen besitzen eine wesentlich bessere Wärmeableitung. Ihre Strombelastung führt in der Regel nicht zu einem unmittelbaren Ausfall, sondern zu einer Degradation, der Elektromigration, die häufig erst nach längerer Betriebszeit zu Funk-

tionsstörungen führen kann. Unter Elektromigration versteht man einen gerichteten Materialtransport, der in erster Linie durch eine Impulsübertragung von Leitungselektronen auf die Atomrümpfe des metallischen Leiters entsteht. Elektromigration kann sowohl zur Widerstandserhöhung bzw. zur vollständigen Unterbrechung elektrischer Leiterbahnen als auch zu Kurzschlüssen benachbarter Leiter führen. Verschiedene Materiallösungen zur Erhöhung der Migrationshärte sind möglich.

Neben den migrationsbedingten Ausfällen tritt ein zusätzlicher Effekt immer stärker in den Vordergrund, der kurz als RC-Verzögerung bezeichnet wird. Betrachten wir zwei parallel übereinander liegende Leiterbahnen als Kondensator, so ergibt sich als Signalverzögerung für die gesamte Schaltung $RC \sim \rho_{\text{met}} \kappa_{\text{ILD}} L^2 / t_{\text{ILD}} t_{\text{met}}$, wobei ρ_{met} der spezifische Widerstand des Leiterbahn mit der Dicke t_{met} und der Länge L ist. κ_{ILD} steht für die dielektrische Konstante und t_{ILD} für die Dicke der isolierenden Zwischenschicht (Interlayer Dielectrics, ILD) zwischen den beiden Leiterbahnen. Grundsätzlich sollte RC nicht größer als die Signalverzögerung am Transistor sein, da sonst die Geschwindigkeit des gesamten Schaltkreises nicht mehr durch die Geschwindigkeit der aktiven Bauelemente bestimmt wäre. Die RC-Verzögerung lässt sich nun reduzieren, indem man den spezifischen Widerstand der Leiterbahn, die Dielektrizitätskonstante des ILD bzw. die Leiterbahnlängen reduziert.

Kupfer statt Aluminium

Verglichen mit dem traditionellen Leiterbahnmaterial Aluminium hat Kupfer mehrere für die Mikroelektronik wesentliche positiven Eigenschaften: So ist sein spezifischer Widerstand geringer ($1,7 \mu\Omega\text{cm}$ gegenüber $2,7 \mu\Omega\text{cm}$ für Al) und seine Resistenz gegen Elektromigration und damit seine Lebensdauer höher. Außerdem lassen sich aufgrund der bei Kupfer-Metallisierung verwendeten *Dual Damascene*-Architektur (Abschei-

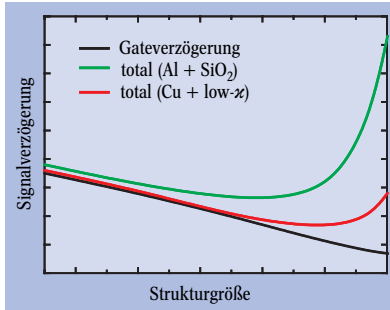


Abb. 8: Modellrechnung für die Gesamtsignalverzögerung auf einem modernen Chip mit Al+SiO₂ (grün) bzw. Cu+Low- κ -Metallisierung (rot) als Funktion der Strukturgröße. Zum Vergleich ist die sich kontinuierlich verringere Gateverzögerungszeit (schwarz) gezeigt. Die Breite der Leiterbahnen skaliert mit der Strukturgröße, ihre Dicke und Länge sind als konstant angenommen.

dung der Metallisierung auf das strukturierte Oxid) Kosten während der Prozessführung einsparen. Nachteilig ist hingegen die schwierige Prozessierbarkeit, da Kupfer die Eigenschaft hat, durch Oxide und ähnlich aufgebaute Dielektrika zu diffundieren. Wird es nicht vollständig eingekapselt, kann es so in die aktiven Gebiete der Schaltung gelangen, diese kontaminieren und unbrauchbar machen. Anders als bei Aluminium bildet sich auf Kupfer bei Kontakt mit der Luft außerdem keine natürliche Passivierung, die das gesamte Volumen vor der Oxidation schützt. Trotz dieser Probleme hat sich in den letzten Jahren die Kupfermetallisierung in modernen hochintegrierten Schaltkreisen durchgesetzt.

Low- κ Zwischenebenen-Isolatoren

Zwischen den einzelnen Leiterbahnen befinden sich Isolatorschichten, die neben ihren dielektrischen Eigenschaften auch eine entsprechende mechanische Stabilität aufweisen müssen, da sie das gesamte Leiterbahnsystem tragen. Wie weiter oben gezeigt, sollte die dielektrische Konstante dieser Schichten so klein wie möglich sein, um möglichst geringe Signalverzögerungen zu erhalten. Die kleinste theoretisch erreichbare Dielektrizitätskonstante von 1 würde durch Vakuum erreicht werden. Da die Leitungen aber nicht in der Luft hängen können, werden stabile Materialien mit möglichst kleinem κ untersucht (Low- κ -Materialien) [14]. Grundsätzlich gibt es zwei Wege zur Reduzierung von κ , nämlich a) die Reduzierung der Dichte der Materialien und b) die Reduzierung ihrer Ionizität. Gegenwärtig untersuchte Low- κ -Materialien lassen sich in folgende Gruppen einteilen: dotierte Oxide, organische Materialien, hochfluorinierte Materialien und poröse Materialien. Die Gruppe der porösen Materialien kombiniert Mikroporen mit den anderen aufgezählten Materialien. Poröse Materialien können daher generell niedrigere Dielektrizitätskonstanten erreichen als die entsprechenden massiven Stoffe. Heutzutage werden bereits Dielektrizitätskonstanten von 2–3 realisiert. Oxide, dotierte Oxide, organische Materialien sowie hochfluorinierte Materialien werden künftig alle mit Mikroporen kombiniert werden. Grundsätzlich sind aber auch für die neuen Materiallösungen verschiedene Anforderungen hinsichtlich ihrer Prozesskompatibilität zu berücksichtigen.

Der Übergang zu einer Kupfer-Metallisierung kombiniert mit geeigneten Low- κ -Materialien wird nicht

allein genügen, um die Anforderungen an zukünftige Metallisierungssysteme zu erfüllen (Abb. 8). Vielmehr muss künftig das verstärkte Ausnutzen von physikalischen Geometrie-Effekten (z. B. Breiten- und Längenabhängigkeiten von Leiterbahnen, Vermeidung von langen Leiterbahnen, eventuelle 3-dimensionale Integration usw.) bereits im Layout des Mikrochips stärker berücksichtigt werden.

Fazit

Wir durchleben gegenwärtig einen grundlegenden Wechsel in der mikroelektronischen Entwicklung/Forschung, bei dem Innovationen auf allen Ebenen notwendig sind. Das betrifft neue Materialien genauso wie neue Bauelementekonzepte und Schaltkreisarchitekturen. Dabei wird die Anzahl neuer Materiallösungen in den nächsten Technologiegenerationen exponentiell steigen. Benötigt werden verstärkt „man-made“-Materialien, deren Eigenschaften sich durch Kombination verschiedener Komponenten genau auf den Anwendungszweck zuschneiden lassen, wie hier an einigen ausgewählte Beispiele gezeigt. Trotz dieser Materialien wird sich die „klassische“ Mikroelektronik (CMOS) physikalischen, technischen wie ökonomischen Grenzen nähern. Vollständig neuartige Lösungsansätze einer zukünftigen Nanoelektronik werden bereits in einigen Jahren benötigt.

Literatur

- [1] G. Moore, Electronics **38**, Nr. 8 (1965)
- [2] R. Dittmeyer et al. (Hrsg.), Winnacker-Küchler: Chemische Technik, Bd. 7, Wiley-VCH, Weinheim (2004)
- [3] J. G. Cressler (Hrsg.), Silicon Heterostructure Handbook, Dekker (2005)
- [4] Properties of Silicon Germanium and SiGe:Carbon, E. Kasper und K. Lyutovich (Hrsg.), EMIS Datareviews Series, No 24 (2000)
- [5] P. Ashburn, Heterojunction Bipolar Transistors, Wiley, New York (2003)
- [6] H. J. Osten et al., Techn. Digest IEDM **97**, 803 (1997)
- [7] B. Heinemann et al., Techn Digest IEDM **04**, 239 (2004)
- [8] W. Winkler et al., ISSCC Febr. 2005
- [9] M. L. Green et al., J. Appl. Phys. **90**, 2057 (2001)
- [10] K. J. Hubbard und D. G. Schlom, J. Material Research **11**, 2757 (1996)
- [11] G. D. Wilk et al., J. Appl. Phys. **89**, 5243 (2001)
- [12] J. Mannhart und D. G. Schlom, Physik Journal, Juni 2005, S. 45
- [13] H. J. Osten et al., J. Crystal Growth **278**, 18 (2005)
- [14] M. Morgen et al., Ann. Rev. Mat. Sci. **30**, 645 (2000)

Der Autor

H. Jörg Osten studierte Physik in Polen an den Universitäten Breslau und Posen. Nach Auslandsaufenthalten in Sofia (Bulgarien), Chicago (USA) und im britischen Cambridge habilitierte er 1986 und wurde im folgenden Jahr Professor für Physikalische Chemie an der Akademie der Wissenschaften der DDR. Von 1988 bis 2002 forschte H. Jörg Osten in verschiedenen leitenden Funktionen am Institut für Halbleiterphysik in Frankfurt (Oder). Seit März 2002 ist er Professor für Materialien und Technologien der Elektronik am Fachbereich Elektrotechnik der Universität Hannover.

