

Fragenausarbeitung bzw. Zusammenfassung des Skriptums

Frage 1: --Kap1: Einleitung--

Nennen Sie Anwendungsgebiete für Modellierung und Simulation in der Halbleitertechnologie und beschreiben Sie diese kurz.

Antwort 1:

Einleitung:

Früher hat sich die Halbleitertechnologie durch „Versuch und Irrtum“ entwickelt. Heute spart man durch Simulation wertvolle Zeit und Kosten und erreicht dadurch kürzere time-to-market-Zeiten.

Das Moor'sche Gesetz beschreibt das exponentielle Anwachsen der Chip-Komplexität. (verdoppelt sich alle 18 Monate / minimum feature size -30% alle 3-4 Jahre). Die ITRS (erscheint alle 2 Jahre) geht von der weiteren Gültigkeit des MG aus und listet die technologischen Anforderungen für dessen Einhaltung auf. Dieser Zeitplan für die Technologieentwicklung wäre ohne Simulation nicht mehr einzuhalten. Es wird zunehmend schwieriger, das MG einzuhalten (vgl. technolog., funktionelle. und Kostengrenze) => Simulation wird immer wichtiger.

In den 80er Jahren entwickelte sich zunächst die Device- und Schaltkreissimulationen. Die Prozesssimulation (Modellierung der Herstellungsschritte) kam erst später (kompl. Phys. Modelle erf.) Heute sind die führenden Hersteller für TCAD-Produkte: Synopsis, Silvaco und ISE.

Anwendungsgebiete:

- **Equipment-Simulation:** Simuliert die Vorgänge in Prozessanlagen (Öfen, CVD-/Ätz-Reaktoren, Kristallziehen,...)
Wichtig für: Equipmentdesign, Sensordesign und Prozesskontrolle
Wird primär vom Equipmenthersteller angewendet.
- **Prozesssimulation (TCAD)** simuliert Vorgänge auf dem Wafer
 - **Topographie-Simulation:** beschäftigt sich mit all jenen Prozessen, welche die Geometrie der Waferoberfläche verändern (Ätzen, Deposition, Entwickeln von Fotolack,...)
 - **Lithographie-Simulation:** Simuliert die Abbildung der Maske durch die Lichtstrahlen am Wafer sowie die Wechselwirkung mit dem Resist.
 - **Front-End-Process-Modelling:** behandelt alle Prozesse, die direkt am Si ausgeführt werden (Implantation, Oxidation, Diffusion,...) Sie ist meist mit einer Topographiesimulation eng gekoppelt (vgl. Implantationsmasken)
- **Device-Simulation:** übernimmt die Geometrie und Dotierstoffverteilung aus der Prozesssimulation und errechnet elektrische Parameter zur Charakterisierung der Bauteile in der nachfolgenden Schaltkreissimulation. Zunehmend unter Berücksichtigung von Quanteneffekten.

physikalische Prozessmodellierung

- **Schaltkreissimulation (ECAD):** Simuliert das elektrische Verhalten einer Schaltung (vgl. „Spice“).
- **Package Modelling:** Modelliert thermische, mechanische und elektrische Vorgänge die mit der Verkapselung des Chips und den Anschlüssen zusammenhängen.

Der Begriff „Technologiesimulation“ wird oft synonym mit Prozess- und Equipmentsim. verwendet.

Frage 2: --Kap1: Einleitung--

Nennen Sie Methoden zur Simulation der Diffusion und erläutern Sie diese.

Antwort 2:

Einleitung:

Die Simulation der Diffusion ist sehr wichtig, da sie quasi bei jedem Prozess mit höheren Temperaturen oft ungewollt auftritt und analytische Lösungen meist unbrauchbar sind. Die Diffusionsgleichung lässt sich aus den beiden Fick'schen Gesetzen ableiten und kann für einige Fälle leicht analytisch gelöst werden.

Gauß'sche Diffusionsprofile ergeben sich bei Diffusion aus einer unerschöpflichen Quelle mit Delta oder Gauß-Profil als Anfangsbedingung. (kein Ein- u. Ausdiff.)

Fehlerfunktionsprofile erhält man, wenn die Diffusionsglg. für eine unerschöpfliche Quelle mit $\text{rect}(\cdot)$ -Funktion als Anfangsbedingung löst. (z.B.: Ein- und Ausdiffusion, Diffusion zwischen 2 Körpern mit unterschiedl. Konzentration)

Diese Lösungen setzen einen konstanten Diffusionskoeffizienten D voraus sowie bestimmte, mathematisch exakte Randbedingungen. =>

Analytische Lösungen selten brauchbar wegen:

- Anfangsverteilungen meist nicht durch Gauß- oder Deltafunktion gegeben.
- Dotieratome sind geladen => Wechselwirkung zw. Verschiedenen Dotieratomtypen und konzentrationsabhängiger D .
- D ist abh. von lokalen Punktdefektdichten (diese wiederum werden z.B. durch Oxidation, Ionenimplantation oder der Diffusion selbst beeinflusst)
- Es gibt eine Löslichkeitsgrenze im Si => darüber tritt Clusterung auf (unbew.)

Methoden zur Simulation:

Verfahren der finiten Differenzen: (numerische Lösung der Diffusionsgleichung)

In einem vorgegebenen Intervall wird eine örtliche Diskretisierung des Konzentrationsverlaufes vorgenommen. Die Stromdichten sind an den Intervallmitten definiert und die Differentialquotienten werden durch Differenzenquotienten ersetzt. Man erhält für jeden Gitterpunkt eine gewöhnliche Differentialgleichung in t . Danach erfolgt die zeitliche Diskretisierung. Dabei ergeben sich 2 Verfahren:

- **Forward Differences:** Konzentrationsverlauf zum nächsten Zeitpunkt direkt (ohne Lösung eines Gleichungssystems) aus dem vorherigen Zeitpunkt berechenbar. Nachteilig ist die mangelnde Stabilität bei zu großen Δt 's.
- **Backward Differences:** Gleichung hat mehrere Glieder mit $j+1$ => Gleichungssystem muss gelöst werden. Vorteilhaft ist die Stabilität für alle Δt 's.

Verfeinerte Verfahren:

- **Newton'sches Verfahren** zur iterativen Lösung von nichtlinearen Gleichungssystemen
- **Nichtäquidistantes Gitter:** dort wo sich viel tut => feineres Gitter
- **Adaptives Gitter:** Gitter passt sich laufend an.
- **Zeitschrittweitensteuerung:** Zeitschritte werden daran angepasst, wie sich die Lösung zeitlich ändert.

Frage 3: --Kap3: Diffusion-- (Spezialgebiet)

Welche Bedeutung haben Punktdefekte für die Diffusion von Dotieratomen? Beschreiben Sie anomale Diffusionsphänomene, die sich mit Hilfe von Punktdefekten erklären lassen.

Antwort 3:

Einleitung/Einschub:

Es gibt einige Phänomene der Diffusion, die sich nicht mit makroskopischen Modellen erklären lassen.

Phänomene die sich durch makroskopische Modelle erklären lassen:

- **Feldeffekt:** Substitutionelle Dotieratome sind elektrisch geladen und erzeugen deshalb ein Potential (festgelegt über Poisson-Glg). Das erzeugt elektrische Feld ist proportional zum Gradienten der Elektronendichte und dieser wiederum zum Gradienten der Nettodotierung Na-Nd. Die Wirkung des el. Feldes auf die Dotieratome kann durch einen Feldterm in der Glg. für den Diffusionsstrom berücksichtigt werden. Für extrinsische Bedingungen wird die Diffusion beschleunigt. Liegt nur eine Dotieratomart vor, so verdoppelt sich gerade der Diffusionskoeffizient. Dies führt zu „Box-Profilen“ (Stauereffekt der tails)
- **Temperaturabhängigkeit:** Der intrinsische Diffusionskoeffizient hängt von der Temperatur gemäß dem Arrhenius-Gesetz ab. (Dotieratome müssen eine Energiebarriere E_a überwinden). Es gibt keine langsam diffundierenden Akzeptoren => Herstellung seichter p-Zonen schwierig. Punktdefekte diffundieren schneller als Dotieratome.
- **Löslichkeitsgrenze:** Wird C_{ss} überschritten, so beginnen die Dotieratome zu präzipitieren und bilden Cluster. Die Konzentration aktiver Dotieratome ist somit maximal gleich C_{ss} . Mit steigender Temperatur steigt auch C_{ss} . Im Modell werden die geclusterten Atome mit einer bestimmten Wahrscheinlichkeit emittiert und kehren mit einer anderen Wahrscheinlichkeit, die prop. zur Konzentration der aktiven Dotieratome ist wieder in den Cluster zurück. Es ergibt sich eine einfache Differentialgleichung, deren Lösung nach einer bestimmten Zeit einen konstanten Wert annimmt, der maximal gleich C_{ss} ist.
- **Segregation an Grenzflächen:** Da die potentielle Energie der Atome in den beiden Materialien im allg. nicht gleich ist, stimmen auch die Übergangswahrscheinlichkeiten nicht überein. Das Verhältnis der beiden Konzentrationen ist aber eine konstante (Segregationskoeffizient m)
- **Diffusion in Poly-Si:** Zuzolge der schnellen Diffusion an den Korngrenzen ist der Diffusionskoeffizient makroskopisch betrachtet im Poly-Si größer. Ein einfaches Modell berücksichtigt: Diffusion entlang den Grenzen und im Korninneren, Segregation und Clustern.

Diffusionsmechanismen mit Punktdefekten:

Der direkte Platzwechsel und der Ringtausch sind energetisch sehr ungünstig. Dotieratome können deshalb nur mit Vakanzen oder Interstitials diffundieren:

- **Vakanzen-Mechanismus:** Wenn eine Vakanze zu einem Dotieratom diffundiert, so können diese leicht den Platz tauschen. Die Vakanze ist quasi durch das Dotieratom diffundiert und hat den Wechsel um eine Gitterstelle erlaubt.
- **Indirekter Interstitial-Mechanismus:** Das Dotieratom wird durch ein I von seinem Platz verdrängt und wird dadurch selbst zu einem I. Es verdrängt ein Si-

Atom vom nächsten Gitterplatz und wird wieder substitutionell. Dabei wird ein I emittiert.

- **Dirketer Interstitial-Mechanismus (Paardiffusion):** Es kann auch energetisch günstiger sein, dass sich das Dotieratom zu einem benachbarten Zwischengitterplatz bewegt. Es kann die sog. Migrationslänge zurücklegen ehe es sich wieder ins Gitter einbaut und ein I emittiert. Ein analoger Vakanzmechanismus ist ebenfalls denkbar (Vakanz an Dotieratom räumlich gebunden)

Die kick-in und kick-out Reaktionen lassen sich gut in einem E-x-Diagramm für ein Interstitial veranschaulichen. Dabei muss das I die Bildungsenergie E_f plus die Migrationsenergie E_m überwinden, um von der Oberfläche in den Wafer diffundieren zu können. Der Diffusionskoeffizient lässt sich nach dem Arrhenius-Gesetz mit $E_{act}=E_m$ angeben. Im Fall des Bors ist es so, dass die Energie des interstitiellen Bors B_I kleiner ist als die der separierten B und I. Die Migrationsenergie des B_I ist ebenfalls kleiner als die des I. Deshalb kann sich das B_I um die Migrationslänge bewegen ehe es wieder durch eine kick-out-Reaktion eingebaut wird.

Der direkte Interstitial-Mechanismus führt bei sehr steilen Anfangsprofilen der Konzentration zu exponentiellen und nicht gauß'schen Ausläufern. Nach einigen kick-Reaktionen nähert sich das Profil aber seiner gauß'schen Form an.

Phänomene die sich durch punktdefektbasierte Modelle erklären lassen:

- **OED/ORD:** Bei der thermischen Oxidation werden Interstitials emittiert (Stapelfehler wachsen), da das SiO_2 -Molekül mehr Platz als ein einzelnes Si-Atom braucht. Diese I-Übersättigung führt zu einer hohen kick-in-Rate von Bor-Atomen und erhöht dadurch den Diffusionskoeffizienten. (Herleitung des neuen D mit $C_I=const.$)
B und P-Atome werden bei Oxidation beschleunigt (OED). => diffundieren mit Interstitials.
Über die Frenkel-Paar-Rekombination führt die Oxidation zu einer Vakanz-Untersättigung, was die Diffusion des „Vakanz-Diffundierers“ Sb (Antimon) verlangsamt (ORD).
Da bei As (Arsen) Interstitials und Vakanz eine Rolle spielen, wird die Diffusion manchmal beschleunigt und manchmal verzögert. Unter der Voraussetzung konstanter Punktdefektdichten kann der neue Diffusionskoeffizient mit der Interstitial bzw. Vakanzübersättigung über die Faktoren f_I bzw. f_V verknüpft werden.
- **Chemisches Pumpen von Punktdefekten:** Bei Gold treten U-förmige Diffusionsprofile über den ganzen Wafer auf. Erklärung: Gold diffundiert sehr schnell ins Innere des Wafers und emittiert Interstitials beim kick-in. Dies führt zu einer I-Übersättigung im Wafer, da die I's langsamer zur Oberfläche rückdiffundieren als das Gold eindiffundiert. Nun muss die Annahme $C_I=const$ fallen gelassen werden und es ergibt sich ein zusätzlicher Term in der Au-Diffusionsgleichung der in Richtung abnehmender I-Konzentration gerichtet ist. Das Gold wird deshalb solange in Richtung der Ränder des Wafers gedrückt, bis sich ein Gleichgewicht zwischen den beiden Strömen einstellt.
Das chemische Pumpen tritt aber auch bei Dotieratomen auf (diese diffundieren aber viel langsamer als Au wodurch sie nie die andere Seite des Wafers erreichen können und deshalb ergeben sich auch keine U-förmigen Profile). Das chem. Pumpen führt aber trotzdem zu einer mit der Tiefe zunehmender I-Übersättigung. Deshalb wird D mit zunehmender Tiefe größer (beschleunigte Tail-Diffusion).
- **Emitter-Push-Effekt:** Wird in die Bor dotierte Basiszone eines Transistors der Emitter-Anschluß eindiffundiert (mit P), dann setzt das Phosphor beim kick-in Interstitials frei, die wiederum die Bor-Diffusion beschleunigen. Seitdem man den Effekt kennt nimmt man As für den Emitter.

- **Konzentrationsabhängigkeit des Diffusionskoeffizienten:** Bei extrinsischer Dotierung $n \gg n_i$ erhöht sich durch den Feldeffekt die Vakanzkonzentration. (Erklärung mit Vakanz aus historischen Gründen). Dies wiederum führt zu einer beschleunigten Diffusion. Für diese Erklärung ist die Voraussetzung eine im thermodynamischen Gleichgewicht befindlichen Punktdefektkonzentration. Der neue Diffusionskoeffizient kann als Reihe aus Potenzen von (n/n_i) geschrieben werden.

Anmerkungen:

Für den Formalismus des Paardiffusionsmodells gibt es 5 Spezies zu unterscheiden: A, AI, AV, I und V. I und V reagieren im Frenkel-paar-Mechanismus miteinander, während die Reaktionen $AI+V=A$; $AV+I=A$ den Namen Frank-Turnbull-Mechanismus tragen. Die Reaktionsraten dieser Prozesse sind proportional zu den Dichten der benötigten Spezies. Aus den Reaktionsraten können Generationsraten für jede Spezies abgeleitet werden. Über die Kontinuitätsgleichungen erhält man ein Gleichungssystem für die Konzentration der einzelnen Spezies, dessen Parameter schwer zu bestimmen sind.

Frage 4: -- Kap3: Ionenimplantation --

Erläutern Sie das Phänomen der transienten Diffusion.

Antwort 4:

Erklärung aus Kap3: Neben der thermischen Oxidation und dem chemischen Pumpen gibt es auch noch andere Quellen von Punktdefekten. Eine der Wichtigsten ist die Ionenimplantation, bei der Si-Atome von ihrem regulären Platz gestoßen werden und als Interstitials auftreten. Diese I's diffundieren zur Oberfläche und zum Bulk weg, verursachen aber vorübergehend einen höheren Diffusionskoeffizienten der Dotieratome. Andere Quellen von Punktdefekten: Sauerstoffpräzipitate, Stapelfehler, {311}-Defekte, Punktdefektcluster, Dotieratomcluster, thermische Nitridation und Silizidierung.

Erklärung aus Kap4: Die Diffusion von Bor und Phosphor wird für eine Zeitspanne t_{TED} beschleunigt. Das einfachste Modell ist das „+1-Modell“. Es geht davon aus, dass alle Punktdefekte durch den Frenkel-Mechanismus rekombiniert sind und nur mehr die Interstitials übrig bleiben die durch substitutionellen Einbau der Dotieratome emittiert werden. $\Rightarrow C_I = N_d$

Der Gesamteffekt bei $t = t_{TED}$ kann mit Hilfe eines einfachen Modells berücksichtigt werden, das folgende Annahmen trifft:

- Alle Defekte bei der mittleren Eindringtiefe der Implantation R_p
- +1-Modell: $Q = N_d$
- Oberfläche des Wafers ist ideale Senke für I's

Durch $R_p \ll d$ (Waferdicke) \Rightarrow I's diffundieren nur in eine Richtung zur Oberfläche.

Nach Implantation sind nicht alle I's mobil (ein Großteil clustert sofort)

Für konstantes \hat{C}_I ergibt sich: $t_{TED} = Q \cdot R_p / (D_I \cdot \hat{C}_I)$

Als charakteristische Größe der Diffusion wird eine Diffusionslänge λ definiert.

Diskussion der Ergebnisse dieses einfachen Modells:

- $\lambda \sim \sqrt{N_d}$ wenn Implantation nicht amorphisiert.
- $\lambda \sim \sqrt{E}$ für $x \geq R_p$ mit E...Implantationsenergie
- $\lambda \sim \sqrt{x}$ für $x \leq R_p$
- λ und t_{TED} nehmen mit abnehmender Temperatur zu!!! (vgl. normale Diffusion)

Die Realität weicht vom +1-Modell aus folgenden Gründen ab:

- es gibt mehr netI als Dopanden
- netI liegt im Mittel tiefer als Dopanden
- netI können mit netV rekombinieren \Rightarrow nicht alle I's können bis zur Oberfläche diffundieren und haben deshalb eine kleine Lebensdauer $\Rightarrow t_{TED}$ kleiner

verfeinertes +1-Modell: $Q = p \cdot N_d$ mit p...Plusfaktor

Wenn der Plusfaktor bekannt ist \Rightarrow es kann genauer simuliert werden durch die Anfangsbedingung $C_I(x) = p \cdot C_B(x)$.

Verteilung der I's liegt tiefer als jene der V's. Bei niedrigen Implantationsdosen sind die I und V im Mittel weiter voneinander entfernt, sodass einige V davondiffundieren und die entsprechende Anzahl von I zurücklassen ($D_V \gg D_I$). Der Plusfaktor ist Dosisabhängig (steigt mit höherer Dosis).

Transiente Diffusion für $t < t_{TED}$: Hier ist es erforderlich, die Defekte zu simulieren weil:

- jeder Defekttyp erzeugt ein anderes $\hat{C}_I \Rightarrow t_{TED}$ wird beeinflusst.
- Defekte wachsen mit der Zeit $\Rightarrow \hat{C}_I$ nimmt ab.
- I können im Laufe der Zeit von einem Defekttyp in einen anderen übergehen.

Reverse Short Channel Effekt: Unerwartete Erhöhung der Schwellspannung bei Verringerung der Gatelänge. Erklärung: Die während der S/D-Implantation erzeugten I diffundieren und werden vom Gateoxid absorbiert \Rightarrow Gradient zur Oberfläche für die Dotieratome im ganzen Kanal \Rightarrow pile-up $\Rightarrow V_t$ steigt.

Frage 5: --Kap4: Ionenimplantation--

Nennen Sie Methoden zur Simulation der Ionenimplantation und erläutern Sie diese.

Antwort 5:

Einleitung:

In diesem wichtigsten Dotierverfahren wird ein Strahl hochenergetischer Ionen auf den Wafer gerichtet. Die Ionen werden an den Kristallatomen zufällig gestreut (I's werden emittiert und es kann auch Recoils geben) => Implantationsprofile
Ein großer Vorteil dieses Dotierungsverfahrens ist die Reproduzierbarkeit der Implantationsprofile durch leicht kontrollierbare Parameter (Dosis, Energie). Nachteilig ist die Notwendigkeit der thermischen Ausheilung von Gitterschäden. Die Simulation der Ionenimplantation soll mit Hilfe der Monte-Carlo-Simulationstechnik die Profile und Schäden vorhersagen können.

Der Trend in der Ionenimplantation geht zu niedrigeren Energien hin.

Die Monte-Carlo-Simulation muss Stoßpartner für die eindringenden Ionen unter Berücksichtigung der Gitterschwingung und den Schäden finden. Der Stoß selbst wird nach der klassischen Mechanik berechnet (interatomares Potential). Ein Recoil findet für Restenergien der Ionen von 15eV statt. Ebenso müssen die Damage-Rekombination und die elektronische Abbremsung berücksichtigt werden.

Die freie Weglänge wird so festgelegt, dass 1 Targetatom pro aufgespannten Zylinder vorhanden ist. Es gilt: $R_n = F(p_n) \Rightarrow p_n = p_{max} \cdot \sqrt{R_n}$ (vgl. Erzeugung von Zufallsvariablen x aus gegebener Wahrscheinlichkeitsdichtefunktion $f(x)$)

Bei großer freier Weglänge sind wenige Stöße zu berücksichtigen => schnelle Simulation. Die Gitterschwingung wird über drei gaußverteilte Zufallsvariablen angegeben.

Bei hohen Energien ist die elektronische Abbremsung dominant. Unter „stopping power“ versteht man den Energieverlust pro Å. Schwere Ionen haben eine kleinere Eindringtiefe wegen ihrer größeren Ladung. Analytische Modelle versagen beim Übergang zwischen 2 Materialien.

Analytische Modelle:

Für amorphes Si bei dem kein Channeling (bevorzugt in monokristallinen Si in [110]-Richtung) auftritt lassen sich die Implantationsprofile gut durch die Pearson-Funktion (verallgemeinerte Gauß-Funktion) annähern. Aus der Verteilungsfunktion können Parameter wie die Schiefe, mittlere Eindringtiefe, Standardabweichung und Kurtosis berechnet werden (diese Parameter dienen zum Anpassen der Pearson-Funktion an den tatsächlichen Verlauf).

Der Konzentrationsverlauf hängt mit der Pearson-Funktion $f(z)$ wie folgt zusammen:
 $C(z) = N_d \cdot f(z)$ N_d ...Dosis

Die Gaußfunktion ist eine spezielle Pearson-Funktion (Schiefe=0, Kurtosis=3). Allgemein erhält man die Pearson-Funktion als Lösung der Gleichung:

$$d/dy [\ln(f(y))] = (y-b_1)/(b_0+b_1y+b_2y^2)$$

Die Dual-Pearson-Funktion hat 9 Fitparameter und ermöglicht die Beschreibung von Channeling: $C(z) = N_{d1} \cdot f_1(z) + N_{d2} \cdot f_2(z)$ mit $N_{d1} + N_{d2} = N_d$

Dies entspr. der Aufteilung in eine „channeling“- und „random scattering“-Komponente. Durch das Amorphisieren des Si-Kristalls nimmt der Effekt des Channelings mit steigenden Dosen ab.

Zweidimensionale Implantationsprofile werden mit Hilfe des Superpositionsprinzips berechnet. Dabei beschreibt eine Funktion den Verlauf der Oberfläche. Und es wird über eine sogenannte Punktantwort $f(x,z)$, deren laterale Verteilung eine Gauß-Form hat, integriert => Konzentration $C(x,z)$

Physikalische Modelle:

Es wird der Stoß mit den Kristallatomen (Kernabbremung) ab einem Abstand p durch Wechselwirkung der interatomaren Potentiale nach der klassischen Mechanik modelliert. Dabei besteht das Potential $V(r)$ aus einem Coulomb-Term und einem Korrekturfaktor, der die Elektronenhüllen beschreibt. Der Streuwinkel, Recoil-Winkel und der Energieübertrag T lassen sich über den Umweg eines Massenmittelpunktsystems berechnen.

Bei der Berechnung von $V(r)$ wird angenommen, dass sich die Elektronenhüllen beim Vorbeiflug eines Ions durchdringen. Es kommen deshalb noch Potentialterme für das Pauli-Verbot und den Wechselwirkungen der Hüllen untereinander bzw. der Hüllen mit den jeweiligen Atomkernen hinzu.

Elektronische Abbremsung: Ist bis heute nicht ganz verstanden, da viele elektronische Prozesse (z.B.: Ionisation von Targetatomen, Übertragung von kinetischer Energie usw.) eine Rolle spielen. Die el. Abbremsung ist wichtig bei hohen Energien, leichten Ionen (fliegen weit!) und unter Channeling-Bedingungen.

Modell von Lindhard: Setzt eine konstante „stopping power“ voraus. Mathematische Beschreibung: $\Delta E_e = N_{Si} \cdot S_e \cdot \Delta R$ mit $S_e = k \cdot \sqrt{E}$ mit Parameter: k

Modell von Oen, Robinson: kann durch den Parameter a angefügt werden:

Beschreibung: $\Delta E_e = S_e / (2\pi a^2) \cdot \exp(-p/a)$ p ...Stoßparameter

Beide Modelle haben Probleme bei der Beschreibung von Implantationen in $[110]$ -Richtung (Channeling).

Verbessertes Modell: Kombiniert die beiden vorhergehenden Modelle mit einem entsprechenden Gewichtsfaktor, der durch Parameter festgelegt wird. Der Anteil des Lindhard Modells wird als Nichtlokaler Anteil bezeichnet (Oen, Robinson als Lokaler).

Implantationsschäden: Wenn ein Targetatom die Energie $E > E_d (15\text{eV})$ erhält, wird es zu einem Recoil \Rightarrow Vakanz und Interstitial entstehen. Die richtige Beschreibung der Dosisabhängigkeit von Channeling-Tails erfordert die Korrektur der Anzahl der generierten Defekte. Schwere Ionen haben eine effiziente Schadenserzeugung wogegen bei leichten Ionen die Punktdefektrekombination eine Rolle spielt. Ab einer kritischen Punktdefektdichte von $1.15 \cdot 10^{22} \text{cm}^{-3}$ bezeichnet man den Kristall als amorphisiert. Die Grenze zwischen amorpher und kristalliner Schicht liegt bei P-Implantation am tiefsten im Wafer (P ist leichter als As u. Sb \Rightarrow dringt tiefer ein).

Modifiziertes Kinchin-Pease-Modell: Die Energie T wird auf ein Targetatom übertragen. Trotzdem werden nicht T/E_d Defekte erzeugt. Es geht Energie durch elektronischen Energieverlust verloren. Für die Kernstöße steht deshalb abhängig vom Targetmaterial nur die Energie $E_v < T$ zur Verfügung.

Für $E_v < E_d$ entstehen keine Defekte. Bei $E_v > 2.5E_d$ berechnet sich die Anzahl der Defekte zu $v = E_v / (2.5 \cdot E_d)$

Frage 6: --Kap5: thermische Oxidation--

Nennen Sie Einflussgrößen der thermischen Oxidation und erläutern Sie die Rolle der mechanischen Spannungen.

Antwort 6:

Einleitung:

Si ist deshalb so beliebt, weil eine gut isolierende Oxidschicht einfach aufgewachsen werden kann. Bei thermischer Oxidation ist diese an der Grenzschicht auch sehr hochwertig. Die Schicht lässt sich gut strukturieren und die meisten Atome diffundieren schlecht in $\text{SiO}_2 \Rightarrow$ Maske für Diffusionsprozesse möglich. ITRS sagt voraus, dass SiO_2 als Gate-Dielektrikum abgelöst werden muss. Für Isolationen zwischen den Leiterbahnen sucht man dagegen Materialien mit niedrigem ϵ .

Bei der Oxidation wird das Substrat auf 1000°C geheizt und es laufen an der Grenzschicht die chem. Reaktionen ab (mit $\text{H}_2\text{O}/\text{O}_2$). Das Volumen eines SiO_2 -Moleküls ist 2,27 mal so groß wie jenes eines Si-Atoms \Rightarrow durch den Einbau hebt sich die Oberfläche und Si wird verbraucht (44%/56%). Die Silizidierung (z.B.: mit W zur Senkung von Serienwiderständen) ist ein verwandter Prozess, bei dem vorher die Metallschicht aufgebracht wird, die dann bei höheren Temperaturen mit dem Si reagiert. Nur bei 1D-Strukturen (keine Kanten, usw.) sind empirische Modell (z.B.: Deal u. Grove) anwendbar, da in 2D oder 3D-Strukturen sich mechanische Spannungen ausbilden, die wiederum Einfluss auf die Oxidationsrate haben. Da aber auch Diffusion der Dotieratome bei diesen hohen Temperaturen die Oxidation beeinflusst und umgekehrt, sind oft numerische Methoden erforderlich.

Analytische Modelle:

Modell von Deal und Grove:

Annahmen:

- oxidierende Spezies muss durch die bestehende Oxidschicht hindurchdiffundieren und reagiert an der Si/SiO₂-Grenzschicht
- Konzentration an der Oberfläche im Gleichgewicht mit jener der Gasphase
- Diffusion im Oxid ist stationär (linearer Konzentrationsverlauf)

Bei der Herleitung bezieht man sich auf 2 Stromdichten. J_1 ist die Lösung der Diffusionsgleichung im Oxid (Oberflächenkonzentration C^*) und J_2 ist der Strom durch die chemische Reaktion an der Grenzschicht und ist proportional der Konzentration an dieser Stelle ($k_s \cdot C^i$). Diese beiden Stromdichten müssen gleich sein. Daraus lässt sich C^i eliminieren. Über eine Gleichung für die Anzahl der pro Zeit dt auf der Fläche A entstehenden SiO_2 -Moleküle erhält man eine Differentialgleichung in d_{ox} , die man durch Integration auf eine quadratische Gleichung bringen kann. Für kleine Oxiddicken verläuft das Wachstum näherungsweise linear und geht dann in ein parabolisches Wachstum über. Die Lösung wird mit den Faktoren $B/A = k_s \cdot V_{\text{SiO}_2} \cdot C^*$ und $B = 2D \cdot V_{\text{SiO}_2} \cdot C^*$ dargestellt.

Einflussgrößen der therm. Oxidation:

Temperatur: Die Oxidationsrate nimmt gemäß dem Arrhenius-Gesetz mit steigenden Temperaturen zu. Man nimmt an, dass dies auf die Temperaturabhängigkeit der Parameter k_s und D zurückzuführen ist.

Dry/Wet: Wet-Oxidation ist schneller, weil das H_2O -Molekül kleiner als das O_2 ist und deshalb leichter in das Oxid eindringen kann. ($C^*(\text{wet}) \gg C^*(\text{dry})$)

Partialdruck p des Gases: Die Oxidationsrate steigt mit zunehmendem Druck (C^* steigt). Bei der trockenen Oxidation ist dies allerdings nur eine sublineare Abhängigkeit.

Orientierung der Oberfläche: Da der Faktor k_s von der Kristallrichtung abhängt, ist damit nur der Faktor B/A beeinflusst. Es gilt: $k_s(100) < k_s(110) < k_s(111)$.

HCl-Partialdruck: Zu Verbesserung der Oxideigenschaften wird manchmal HCl beigefügt. Die Beeinflussung dadurch ist schlecht verstanden und erfolgt durch tabellarische Korrektur.

Dotierung am SiO₂/Si-Interface: Bei Bor und Phosphor-Dotierung (bei nicht zu hohen Temperaturen) nimmt die Oxidationsrate zu. Ursachen: Diffusionskoeffizient vom H₂O wird durch das ins Oxid eingebautes Bor (wird vom Oxid angesaugt) vergrößert. Beim P glaubt man, dass durch den „pile-up“ (P wird vor dem Oxid hergeschoben) die extrinsische Dotierung eine erhöhte Vakanzenkonzentration zur Folge hat, die wiederum die Reaktionsrate über k_s linear erhöht.

Modifikation für dünne Oxide: Für sehr dünne Oxide stellt man ein beschleunigtes Oxidwachstum fest, das sich nicht im Rahmen des Deal/Grove'schen Gesetzes erklären lässt. Man fügt deshalb einen empirisch ermittelten Korrekturterm für dd_{ox}/dt ein.

Mechanische Spannungen im Oxid!

Grenzen des Deal u. Grove'schen Gesetzes:

Dieses analytische Gesetz ermöglicht nur die Beschreibung von planarem Oxidwachstum. Beim lateralen Wachstum am Zylindermantel ergeben sich periodisch variierende Oxiddickenschwankungen, die sich mit der Abhängigkeit der Oxidationsrate von der Kristallorientierung, erklären lassen. An konkaven Strukturen ist das Oxidwachstum geringer als an Konvexen. (mech. Spannungen hemmen die Reaktion)
Diese Effekte nehmen bei steigenden Temperaturen ab.

Numerische Modelle:

Sind unbedingt erforderlich bei 2- oder 3D-Strukturen aber auch bei lokaler Oxidation. Folgende gekoppelte Phänomene müssen simuliert werden:

- Diffusion des Sauerstoffs im Oxid
- Berechnung des neuen Oxids und der Verschiebung des darüber liegenden.
- Behandlung von Dotierstoff-Diffusion und Segregation an der Grenzschicht.

Wichtig ist die Simulation der entstehenden Spannungen, da diese die lokale Oxidationsrate stark beeinflussen und die Überschreitung eines kritischen Wertes Kristalldefekte zur Folge hat.

Grundbegriffe der Kontinuumsmechanik:

Befindet sich ein Körper in einem bestimmten Spannungszustand, so wirken auf jede beliebige Fläche im Körper entsprechende Kräfte. Normalspannung und Schubspannung sind die Normal- bzw. Tangentialkomponente der mech. Spannung an einem Flächenstück. Man definiert einen Spannungstensor als Matrix von den speziellen Zeilen-Spannungsvektoren σ_x , σ_y und σ_z , die für die speziellen Flächen A_x , A_y und A_z festgelegt wurden. Diesen Tensor kann man in einen sphärischen Anteil, der für die Volumenänderungen verantwortlich ist und in einen deviatorischen Anteil, der die Formänderungen beschreibt, zerlegen.

Bei der Oxidation verschwindet die Summe der Kräfte und Momente auf ein Volumenelement woraus sich ein symmetrischer Spannungstensor ergibt. Die Summe der partiellen Ableitungen der Spannungsvektoren muss Null sein (siehe Skriptum). Daraus ergeben sich drei Gleichungen für die Komponenten des Spannungstensors.

Verzerrungen berücksichtigen keine reine Translation und Rotation sondern setzen sich aus Längenverzerrungen (Dehnungen) und Winkelverzerrungen zusammen. Die Verschiebungen der Punkte eines Körpers können durch ein Vektorfeld angegeben werden. Ein Objekt, das die Verzerrungen charakterisiert und Rotations- und Translationsinvariant ist bezeichnet man als Verzerrungstensor. Die Hauptdiagonalelemente dieses symmetrischen Tensors geben die relative Dehnung einer infinitesimal kleinen, parallel zur entsprechenden Achse orientierten Strecke an. Die Winkelverzerrung ist durch die restlichen Elemente charakterisiert.

Wenn man den Zusammenhang zwischen dem Spannungs- und Verzerrungstensor kennt, hat man 3 partielle Differentialgleichungen in den Unbekannten u_x , u_y und u_z .

Materialverhalten:

- **Elastisches Verhalten:** Bei linearem Zusammenhang zwischen Spannung und Dehnung (Hook'sches Gesetz) lässt sich eine Steifheitsmatrix c angeben (spezielle Anordnung der Elemente siehe Skriptum). Diese symmetrische Matrix führt im Allgemeinen auf ein elliptisches Differentialgleichungssystem 2. Ordnung. In kubischen Systemen (z.B.: Si) sind nur 3 Steifheitskonstanten unabhängig und in isotropen Medien gibt es nur 2 Konstanten. Im letzten Fall definiert man auch den Kompressionsmodul und den Schubmodul.
- **Viskoses Verhalten:** In einer Flüssigkeit gleiten die Atome aneinander vorbei. Die innere Reibung kann als proportional der Relativgeschwindigkeit aneinander vorbeigleitender Atome modelliert werden. Der deviatorische Anteil des Spannungstensors ist proportional zur zeitlichen Ableitung des Verzerrungstensors über eine Konstante, die als Viskosität bezeichnet wird. Wird ein spannungsfreier Zustand wiederhergestellt, so werden nur die Volumenänderungen nicht aber die Verzerrungen vollständig rückgängig gemacht
- **Viskoelastisches Verhalten:** SiO₂ zeigt teilweise elastisches und viskoses Verhalten bei typischen Prozesstemperaturen. Für rasche Änderungen ist das Verhalten elastisch. Man nennt solche Objekte auch als Maxwell-Körper (müssen einer bestimmten Beschreibung genügen: siehe Skriptum)

Spannungsabhängige Parameter der Oxidation:

- **Reaktionskonstante des Sauerstoffs:** Hängt von der Normalspannung am Si/SiO₂ Interface ab. (prop. $\exp(-\dots)$) => für steigenden Druck nimmt k_s ab.
- **Diffusionskoeffizient des Sauerstoffs im Oxid:** Nimmt mit steigendem hydrostatischem Druck p ab. (wie k_s)
- **Gleichgewichtskonzentration des Sauerstoffs C^* :** Auch diese Konzentration ist abhängig vom Druck p im Oxid und folgt der gleichen Proportionalität. Allgemein werden empirische Parameter festgelegt.
- **Viskosität des Oxids:** Sie ist ebenfalls spannungsabhängig (von-Mises-Schubspannung: siehe Skriptum) und führt daher zu einem komplexen mechanischen Sachverhalt bei der Simulation.

Simulationsbeispiele:

Ein wichtiges Beispiel ist die lokale Oxidation im CMOS-Prozess. Im Bereich des „bird's head“ treten sehr hohe Spannungen auf (es kann die gesamte Spannungsverteilung in Abhängigkeit der Oxidationszeit simuliert werden). Unterschiedliche Ausführungsformen der Nitridschicht können untersucht werden.

Wechselwirkung zwischen Oxidation und Diffusion:

OED/ORD: Durch die Größe der SiO₂-Moleküle bauen sich Spannungen auf, die durch Emission von Interstitials abgebaut werden können. Über den Frenkel-paar-Mechanismus führt die I-Übersättigung zu einer V-Untersättigung. Bor und Phosphor werden dadurch beschleunigt. Die Diffusion von Antimon wird hingegen gebremst. Die Generationsrate der Interstitials hängt von der Geschwindigkeit der Si/SiO₂-Grenzschicht nach einem Potenzgesetz ab dessen Parameter dem Arrhenius-Gesetz folgen. Die Interstitial-Übersättigung hängt aber auch von der Rekombinationsrate der I's an der Si/SiO₂-Grenzschicht ab. Diese Rate wird auch über ein Potenzgesetz mit temperaturabhängigen Parameter modelliert.

Dotierungsabhängigkeit der Oxidationsrate: B/A erhöht sich bei extrinsischer Konzentration der Dotieratome an der Si/SiO₂-Grenzschicht. Erklärung: über Vakanzensättigung siehe weiter oben.

Segregation von Dotieratomen an der Si/SiO₂-Grenzschicht: Je nach Dotieratomart ist es energetisch günstiger, dass das Atom im Si oder im Oxid ist. Das Verhältnis der Konzentrationen an den Seiten der Grenzschicht ist konstant und heißt Segregationskoeffizient m . Unter oxidierenden Bedingungen bewegt sich die Grenzschicht über die Dotieratome hinweg und es bauen sich gerade so viele Atome ins Oxid ein, wie dem Segregationskoeffizienten entspricht. Da die meisten Atome schlecht in SiO₂ diffundieren, sieht man diesen Effekt nur unter oxidierenden Bedingungen. Bor hat $m=0.3$ und hält sich bevorzugt im Oxid auf. Es wird also vom SiO₂ regelrecht angesaugt wodurch eine Verarmung im Si entsteht. Die meisten anderen Dotieratome haben $m \gg 1 \Rightarrow$ sie werden von der Grenzschicht vor sich her geschoben („pile-up“). Für die Dotierungsabhängigkeit der Oxidation ist die Konzentration der Dotieratome an der Grenzschicht maßgeblich. Diese verändert sich aber im Laufe des Prozesses \Rightarrow gekoppelte Simulation erforderlich.

Frage 7: --Kap6: Topographie-Simulation--

Was versteht man unter Topographiesimulation? Welche Prozesse der Halbleitertechnologie werden im Rahmen der Topographiesimulation behandelt? Welche Methoden gibt es zur Bewegung der Oberfläche?

Antwort 7:

Einleitung:

Unter Topographiesimulation versteht man die Simulation der Geometrieänderung der Waferoberfläche zufolge von Schichtabscheidungen oder Ätzen.

Die Oxidation wird häufig getrennt behandelt, da sie durch Sauerstoffdiffusion und mechanischen Spannungen ins Innere wirkt. (Unterschied Oxidation zu anderer Deposition: bei anderer Deposition werden die Materialien ausschließlich von außen geliefert)

Typische Aufgabenstellungen sind die Frage der Kantenbedeckung (step coverage; Verhältnis von vertikaler zu horizontaler Schichtdicke) sowie die Frage ob ein Loch vollständig ausgefüllt wird. Auch Schichtzusammensetzung oder Schädigung des Substrats sind gefragt.

Oft genügen einfache Modelle mit isotroper und oder unidirektionaler Depositions/Ätzrate. Es genügt die Angabe eines Algorithmus zur Bewegung der Oberfläche.

Depositions- und Ätzverfahren:

Allgemeines Prinzip: Moleküle werden zum Wafer transportiert, interagieren dort mit der Oberfläche oder mit anderen adsorbierten Molekülen und schließlich werden wieder Moleküle emittiert.

CVD-Verfahren: Moleküle diffundieren aus der Gasphase zur Oberfläche und werden adsorbiert. Sie können noch entlang der Oberfläche diffundieren und mit anderen adsorbierten Molekülen reagieren.

Der Haftkoeffizient steht für die Wahrscheinlichkeit mit der sich ein adsorbiertes Molekül in die Schicht einbaut bevor es emittiert wird. Ein sehr kleiner Haftkoeffizient ist Vorteilhaft beim Füllen von Löchern.

- **APCVD:** Die Diffusionsgeschwindigkeit der Moleküle zur Oberfläche ist durch die große Stoßwahrscheinlichkeit mit Luftmolekülen sehr klein. Die entstehenden Schichten sind abhängig von den Strömungsverhältnissen des Gases am Wafer.
- **LPCVD:** Die Abscheideraten sind hier ausschließlich durch die chemischen Reaktionen und der Temperatur bestimmt (300-900°C). (Temperatur kann man leichter kontrollieren als die Strömungsverhältnisse)
- **PECVD:** Ein Teil der für die Reaktionen benötigter Energie wird hier durch ein Bombardement von Ionen die durch ein Wechselfeld aus einem Plasma beschleunigt werden, bereitgestellt. Damit kann der Prozess schon bei 200-350°C ablaufen.

Sputter-Deposition(PVD): Das abzuschneidende Material wird entweder verdampft oder durch Ionenbeschuss eines Targets abgesputtert. Die abgesputterten Atome durchfliegen geradlinig(durch $p \ll 1$) das Plasma und bleiben am Wafer haften. Durch Beheizung des Wafers wird ein Haftkoeffizient < 1 erzeugt und somit die Stepp coverage verbessert. Mit dem Verfahren der Sputter-Deposition kann fast jedes Material abgeschieden werden.

Nasschemisches Ätzen: Ist ein rein chemisches Verfahren => hohe Selektivität aber isotropes Ätzen => Unterätzen von Masken. Es wird deshalb nur zum großflächigem Abtragen von Schichten verwendet.

Eine Variante dieses Verfahrens ist das chemisch-mechanische Polieren (CMP), das zur Planarisierung eingesetzt wird. (Schleifkörner 20-500nm tragen bevorzugt die höchsten Erhebungen ab; durch chemischen Anteil ergibt sich eine gewisse Selektivität)

Plasmaätzen: (Trockenätzen): Funktioniert ähnlich wie Sputter-Deposition nur dass die Ionen nicht auf das Target sondern auf den Wafer fliegen und dort Material abspalten. Beim RIE-Prozess wirkt zusätzlich noch ein chemischer Anteil. Dabei finden die chemischen Ätz-Reaktionen nur an denjenigen Stellen statt, die mit Ionen beschossen werden. Damit ist die Selektivität des chemischen Ätzens mit der Anisotropie des Ionenätzens kombiniert worden.

Physikalische Modelle für Deposition und Ätzen:

Teilaufgaben der Modelle:

- Berechnen der Winkel- u. Energieverteilung der einfallenden Teilchen.
- Berechnung des Flusses an jedem Punkt der Oberfläche (berücksichtigt Sichtbarkeit, Re-Emission und Oberflächendiffusion)
- Umrechnung des Flusses auf eine Depositions- oder Ätzrate.
- Bewegung der Oberfläche.

Winkel- u. Energieverteilung:

- **Transport aus der Gasphase** (CVD-Verfahren): isotrope Winkelverteilung und vernachlässigbare Energie.
- **Gesputterte Teilchen:** Durch Monte-Carlo-Simulation werden die das Target verlassende Recoils untersucht => Winkel- u. Energieverteilung
Häufig wird angenommen: $F \sim \cos^n(\Theta)$, wobei Θ der Winkel zwischen der Flugrichtung des Teilchens und der Vertikalen ist.
- **Ionentransport aus dem Plasma:** Elektrische Spannung fällt an der Grenzschicht zwischen Plasma und Wafer ab. In diesem Feld werden die Ionen beschleunigt. (Kollisionen sind in diesem Bereich nicht vernachlässigbar!) Durch Kollisionen können die Ionen ihren Ladungszustand verändern => Winkel- u. Energieverteilung werden durch Monte-Carlo-Simulation bestimmt.

Fluss einfallender Teilchen:

- **Sichtbarkeit:** Für eine kleine freie Weglänge L ist die Diffusionsgleichung für den Fluss der Teilchen zu lösen wogegen bei großen freien Weglängen sich die Teilchen nahezu geradlinig bewegen => „ray tracing“-Methode. Ist L vergleichbar mit den Strukturgrößen, so kommt man mit Monte-Carlo-Simulation ans Ziel.
- **Re-Emission:** Durch dieses Phänomen ergibt sich ein gekoppeltes System, dessen mathematische Beschreibung auf ein Gleichungssystem führt. Es können nämlich Teilchen durch Haftkoeffizient < 1 bzw. je nach Sputterausbeute (=gesputterte Atome/einfallendem Atom) emittiert werden und als „Quellen“ von Flüssen für andere Teile des Wafers, die einen überlappenden „Sichtbarkeitsraum“ beanspruchen, dienen.
- **Oberflächendiffusion:** Wenn die adsorbierten Moleküle über größere Distanzen am Wafer diffundieren, so ist ein entsprechender Term in den Flüssen vorzusehen. Der Diffusionskoeffizient wird häufig als abhängig von der Krümmung der Oberfläche modelliert (in den Mulden geringer).

Depositions- und Ätzrate:

- **Nicht-energetische Teilchen:** Typ. Neutralteilchen aus der Gasphase. Die Depositionsrate ist durch das Volumen der Moleküle und dem Haftkoeffizient gegeben. Die Ätzrate ist abhängig vom Volumen der geätzten Atome und dem Ätzertrag (geätzte Atome pro einfallendem Teilchen)

- **Energetische Teilchen:** Sind Teilchen, die aus einem Plasma beschleunigt oder von einem Target abgesputtert werden. Die Sputter-Ausbeute Y hängt vom Winkel ab und wird häufig folgendermaßen modelliert:
 $Y = c_1 \cdot \cos(\Theta) + c_2 \cdot \cos^2(\Theta) + c_4 \cdot \cos^4(\Theta)$. Y hängt aber auch von der Energie ab (es gibt einen Schwellwert unter dem keine Sputterung stattfindet)

Anwendungen: Multiscale Modelling bedeutet die Simulation mit verschiedenen Längenmaßstäben (atomic-scale, feature-scale, reactor-scale). Eine Anwendung der Topographiesimulation ist das Berechnen der an den Resistwänden angelagerten Substratatome (durch Sputtern entstanden). Die Herstellung von Spacer für die LDD-Technik durch CVD kombiniert mit anisotropem Ätzen. Eine ausgedehnte Quelle führt zu einer größeren Winkelverteilung der einfallenden Teilchen. Beim Sputterätzen können Probleme auftreten: an Flanken können grabenartige Vertiefungen entstehen, der Photoresist kann abgesputtert werden, die Flugbahn kann durch Aufladung der Maske verändert werden,...

Algorithmen zur Bewegung der Oberfläche:

String-Algorithmus: Die Oberfläche wird durch einen Polygonzug beschrieben. Die Eckpunkte werden je nach Ätzrate und Ätzrichtung um eine bestimmte Länge in die gegebene Richtung verschoben (beim anisotropem Ätzen in Richtung der Winkelsymmetralen).

Auftretende Probleme:

- Wenn Streecken zu lang oder zu kurz sind => Punkte einfügen oder löschen
- Winkel zunspitz => es können keine Rundungen entstehen => Punkte einfügen
- Zu spitze Winkel an konvexen Ecken => Zurückbleiben der Ätzfront => Korrekturfaktor $1/\sin(\Theta/2)$ für Ätzrate.
- Entstehung von Oberflächenloops => Problem im 3D schwer lösbar.

Cell-Removal-Algorithmus: Das Simulationsgebiet wird in rechteckige Zellen unterteilt und für jede Zelle wird der Füllgrad $0 < f < 1$ gespeichert. Mit Hilfe der Ätzrate $v(x,y)$ wird der Füllungsgrad abhängig von der Anzahl der offenen Seiten (1 bis 3) der Zellen in eine Zeit umgerechnet, die erforderlich ist um die Zelle wegzuätzen. Nach der kleinsten Zeit wird der Zeitschritt festgelegt. Danach: Updaten der Füllungsgrade.

Problem: Facettenbildung

Modifizierter Cell-Removal-Algorithmus: Der Füllungsgrad in den Zellen ist entweder 1 oder 0 (0 für Luft, 1 für Material). Dabei lässt man ein strukturierendes Element über die Oberfläche gleiten. Alle Zellen, die dabei überstrichen werden erhalten +1 oder -1 zum Füllungsgrad je nachdem ob geätzt oder abgeschieden wird (boolesche Operation). Das strukturierende Element muss in allen Abmessungen größer als eine Zelle sein. Bei der Ätzung über eine Materialgrenze hinweg kann ein zusammengesetztes verallgemeinertes SE verwendet werden (unterschiedliche Größe bestimmt individuelle Ätz- oder Abscheiderate)

Level-Set-Methode: Die Oberfläche wird als Höhengichtslinie einer $n+1$ dimensionalen Funktion definiert. Danach wird die Level-Set-Funktion nach unten bewegt (Schnitt $\Phi=0$ wächst). Geschwindigkeit der Bewegung bestimmt die Rate; Richtung legt fest ob Ätzen oder Deposition. Anfangsbedingung $\Phi(x,y,t=0)$: Vorzeichenbehafteter Abstand d des Punktes (x,y) von der Oberfläche.

Vorteile: einfache Erweiterung auf 3D, geringe Diskretisierungsfehler, keine Probleme beim Verschmelze oder Auftrennen von Gebieten.

Nachteile: erhöhter Rechenaufwand (Dimensionalität um 1 erhöht)

Den Nachteil kann man durch die „Narrow Band Level-Set-Methode“ weitgehend beseitigen (Level-Set-Funktion nur in einem Band um die Oberfläche definiert)

Frage 8: --Kap7: Equipmentsimulation --

Bei welchen Prozessen der Halbleitertechnologie spielen strömende Gase und Wärmetransport eine wesentliche Rolle? Welche Gleichungen löst man bei der Equipment-Simulation dieser Prozesse?

Antwort 8:

Die Equipmentsimulation beschäftigt sich mit der Simulation von Prozessreaktoren (Diffusionsöfen, CVD-Reaktor,...) und wird überwiegend vom Equipmenthersteller angewendet.

Sie hat folgende Modellierungsaufgaben:

- **Strömungsmechanik:** Geschwindigkeits- und Druckverteilung
- **Wärmetransport:** Wärmeleitung, -konvektion, -strahlung
- **Chemische Spezies:** Transport durch Diffusion und Konvektion, chemische Reaktionen

Diffusionsöfen: (Diffusion) Das Gas strömt kalt ein und wird durch die Wände geheizt. Gesucht ist die genaue Temperaturverteilung im Inneren und auch als Funktion der Wafermenge, Waferdurchmesser usw.

RTP-Öfen: (thermisches Annealing) Hier wird der Wafer in einigen Sekunden durch Wärmestrahlung auf einige hundert Grad gebracht. Die Temperaturverteilung in Abhängigkeit der Zeit gilt es zu simulieren.

CVD-Reaktoren: (CVD-Prozesse) Beim APCVD spielt die genaue Druck- und Geschwindigkeitsverteilung der Gase eine wesentliche Rolle. Beim LPCVD steht die Temperaturverteilung im Vordergrund. Der Transport der chemischen Spezies muss in beiden Fällen simuliert werden.

Trocken-Ätz-Reaktoren: (Ätzen) Die Ätzrate ist abhängig von den Strömungsverhältnissen des Ätzgases, der Temperaturverteilung, usw.

Kristallziehen: (Waferherstellung) Die Anzahl der eingebauten Defekte und der Durchmesser des Waferzylinders ist bestimmt durch den Temperaturgradienten zwischen Schmelze und Zylinder sowie der Ziehgeschwindigkeit.

Für Strömungsmechanische Probleme wird die **Kontinuitätsgleichung** gelöst. Sie besagt: „Änderung der Masse im Volumen $V =$ Aus V ausströmender Masse“. Mit Hilfe des Gaußschen Integralsatzes lässt sich leicht eine lokale Darstellung finden. Die „**Navier Stokes-Gleichung**“ geht vom Impulserhaltungssatz aus und bringt Volumenkräfte (z.B.: Schwerkraft) und Druckkräfte ins Spiel.

Zur Berechnung der Temperaturverteilung löst man die **Wärmetransportgleichung** und Gleichungen, die die **Wärmestrahlung** berücksichtigen. Die Wärmetransportgleichung wird über den Energieerhaltungssatz formuliert und berücksichtigt Wärmeleitung und -konvektion. Um die abgestrahlte Energie eines schwarzen Körpers berechnen zu können, muss das Produkt aus Besetzungswahrscheinlichkeit der Photonenzustände mit der Energie entsprechender Photonen und der Anzahl der Zustände über alle Frequenzen integriert werden. Die Besetzungswahrscheinlichkeit wird durch die Bose-Einstein-Statistik beschrieben. Es ergibt sich die Energiedichte zu $E/V = (4/c) \cdot k \cdot T^4$. Pro Flächeneinheit wird die Leistung $J = k \cdot T^4$ angestrahlt. Die Simulation der Wärmestrahlung in einem Ofen ist sehr komplex, da Absorptionen, Reflexionen (auch diffus) und Emissionen der grauen Körper berücksichtigt werden müssen. Es ergibt sich ein Gleichungssystem mit sogenannten Gebhart-Faktoren G_{ji} . Sie entsprechen jenem Anteil der von A_i ausgesandten Strahlung, die auf A_j landet, unter Berücksichtigung von Reflexionen.

Frage 9: --Kap8: Lithographie-Simulation--

Erläutern Sie die drei Bereiche der Lithographiesimulation.

Antwort 9:

Einleitung:

Die optische Lithographie wird bis zu Strukturbreiten von 0.1µm eingesetzt werden können. Je kürzer die Wellenlänge ist, umso schwieriger ist es, transparente Materialien für die Maske zu finden.

Röntgenstrahl-, Elektronenstrahl- und Ionenstrahlithographie sind Alternativen zur konventionellen optischen Litho. Bei Röntgenstrahlen beträgt die Wellenlänge nur 1nm und statt Linsen werden Spiegel verwendet. Projection-X-Ray = EUV.

Direkt schreibende Elektronenstrahlitho eignet sich wegen des Durchsatzes nicht für die Massenfertigung (an einem Projektionsverfahren SCALPEL wird derzeit gearbeitet).

Ionenstrahlithographie hat den Vorteil einer wesentlich kleineren lateralen Streuung der Ionen im beschossenen Substrat.

3 Bereiche sind zu simulieren:

- Simulation des Belichtungssystems
- Ausbreitung des Lichts im Resist
- Entwicklung des Resists

Durch die Simulation des Belichtungssystems erhält man die Intensitätsverteilung des Lichts am „areal-image“ (oberflächennahes Bild).

Die Simulation des 2. Bereiches liefert die Intensitätsverteilung im gesamten Resist sowie die Veränderung der optischen Parameter des Resists bedingt durch chemische Reaktionen.

Die Simulation der Resistentwicklung liefert schließlich die Kontur der Resistoberfläche nach dem Entwickeln.

Simulation des Belichtungssystems:

Die von einer Quelle ausgehenden Lichtstrahlen werden in einer 1. Linse parallelgerichtet. Danach durchsetzen sie die Maske, die mathematisch durch eine Transmissionsfunktion $t(x,y)$ charakterisiert ist. Das Fernfeld des durch die Maske gelassenen Lichts ergibt sich als Fouriertransformation der Transmissionsfunktion $t(x,y)$ (Fernfeld hängt nur mehr vom Winkel der Lichtstrahlen und nicht mehr von der z-Koordinate ab!). Dieses Fernfeld entspricht dem Fraunhofer'schen Beugungsmuster. Vor der 2. Linse ist eine Blende angeordnet, die durch die sogenannte Pupillenfunktion beschrieben wird (hat Werte 0 oder 1). Die 2. Linse projiziert das Bild auf den Wafer. Da nicht alle Lichtstrahlen an der zweiten Linse ankommen, ergibt sich ein Fehler bei der inversen Fourier-Transformation. ($F^{-1}\{p(x,y) \cdot E(x,y)\}$)

Diese Fehler äußern sich als Verschmieren von Detail der Transmittanzfunktion, die die Größenordnung der Wellenlänge und darunter haben => Strukturen sind nach dem Entwickeln im Allgemeinen abgerundet (Optical Proximity Correction der Maske möglich). In der Praxis gibt es weitere Effekte: Schärfentiefe, Off-axis-Belichtung, nur teilweise kohärentes Licht, Abbildungsfehler der Linsen, Endliche Dicke des Resists

Lichtausbreitung im Resist:

Wird durch Lösung der Maxwell-Gleichungen berechnet. Dabei müssen Reflexion und Absorption an der Resist/Substrat-Grenzfläche berücksichtigt werden (stehende Wellen möglich). Man muss gemeinsam mit der Lichtausbreitung die Änderung im Resist berechnen (Absorptionskoeffizient ändert sich im Laufe der Belichtung durch chemische Reaktionen im Resist).

Modelle: $M(x,y,z)$ steht für die Konzentration der lichtempfindlichen Komponente des Resists (Sensitizer) und $I(x,y,z)$ für die lokale Lichtintensität.

Für den 1D-Fall gilt mit den empirischen Konstanten A, B und C:

1. $dM/dt = -C \cdot I \cdot M$... Differentialgleichung für M (exp-Verlauf)
2. $dI/dz = -(A \cdot M + B) \cdot I$... DG für I

Wichtig: Intensität nimmt umso schneller ab je größer M ist (zu Beginn ist Absorptionskoeffizient am größten!)

Ein nachfolgender Post Exposure Bake hat eine Diffusion der lichtaktiven Komponente zur Folge => Auswirkungen von stehenden Wellen gemildert.

Resist-Entwicklung:

Wird als isotropes Ätzen mit einer von der lichtaktiven Komponente $M(x,y,z)$ abhängigen Ätzrate modelliert.

Modell: $v(x,y,z) = \exp(E_1 + E_2 \cdot M + E_3 \cdot M^2)$

(Hohe Ätzgeschw. für großes M => die Modelle entsprechen einem Negativ-Lack)

Ein Photo-Resist Post-Bake härtet das Resist. Dabei tritt ein Verfließen auf, welches durch eine krümmungsabhängige Oberflächendiffusion oder durch viskoses Verhalten modelliert wird.

Anwendungen:

Berechnung der Streuung des Lichts bei der Belichtung von CMOS-Strukturen => Resist wird durch gestreutes Licht auch an anderen Stellen entwickelt.

Simulation der Ausbildung von stehenden Wellen und der en Wirkung auf die Flankengeometrie des Resists.

...

Frage 10: --Kap9: Atomic Scale Modeling --

Beschreiben Sie atomistische Simulationsmethoden und geben Sie Beispiele für ihre Anwendung.

Antwort 10:

Einleitung:

Eine Kontinuumstheorie macht keinen Sinn, wenn sich die Wahrscheinlichkeit, ein Atom an seinem Platz anzutreffen zwischen benachbarten Plätzen stark ändert, oder wenn sich die Kraft zwischen zwei benachbarten Atom-Paaren stark ändert. (makroskopisch betrachtet wären dann die Ableitungen der Größen nicht stetig)

Ein anderer Grund für die Verwendung des Atomic-Scale-Modeling kann sein, dass sich in einigen Fällen das Verhalten eines einzelnen Teilchens leichter beschreiben lässt als das zugehörige Differentialgleichungssystem zu lösen ist. (vgl. Monte-Carlo-Simulation der Ionenimplantation).

Atomistische Simulationmethoden:

Monte-Carlo-Methode: Kennt man einen Katalog von möglichen Ereignissen und deren mittlere Zeitspanne bis zu ihrem Eintritt, so kann man mit Hilfe eines Pseudozufallszahlengenerators diese Ereignisse in der Simulation auswählen. Der große Vorteil gegenüber der Kontinuumsmechanik ist jener, dass der rechnerische Aufwand unabhängig von der Anzahl der simulierten Spezies ist (man muss nicht für jede Spezies eine Differentialgleichung lösen). Die Simulation entspricht insofern nicht der Wirklichkeit, als nur verhältnismäßig wenige Teilchen verwendet werden.

Anwendungen:

- kinetische Monte-Carlo-Simulation bei thermisch aktivierten Prozessen (z.B.: Diffusion: Übergangswahrscheinlichkeit p prop. zur Migrationsenergie und der Temperatur) Ein Ereignis wäre ein sogenannter „hop“.
- bei Ionenimplantation: im amorphen Substrat -> Stoßpartnersuche; im kristallinen Substrat -> Simulation der thermischen Gitterschwingung.
- Simulation von Bor-Interstitial-Cluster. Problem: man muss alle Bindungsenergien kennen (erklärt, dass das B-Clustern verstärkt dort auftritt, wo durch Implantation Interstitials erzeugt wurden)
- Simulation der Wärmestrahlung. Abstrahlungsenergie und -Richtung sowie Emissions-, Absorptions- und Reflexionsgrad werden stochastisch behandelt.
-

Molekulardynamik-Methode: Hier werden die Bewegungsgleichungen numerisch aufintegriert, wobei Kräfte als Ableitung eines Potentials geschrieben werden. Das Potential hängt aber von der Lage aller simulierten Teilchen ab und die Bewegungsgleichung muss für jedes Teilchen integriert werden. Der Vorteil dieser Methode ist die Möglichkeit der Simulation der Wechselwirkung von Teilchen untereinander ohne einen Katalog von Ereignissen aufstellen zu müssen. Ein großer Nachteil ist der enorme Rechenaufwand, da ein Zeitschritt deutlich unter der Periode der Gitterschwingung liegen muss. => Simulationszeiten maximal im ns-Bereich.

Bei der Zeitintegration gibt es 2 Methoden:

- **Central Difference Scheme:** Differentialquotienten werden durch Differenzenquotienten ersetzt. Die Geschwindigkeit wird jeweils an den Punkten $t+\Delta t/2$ bestimmt.
- **Verlet Scheme:** $x(t+\Delta t)$ und $x(t-\Delta t)$ werden als Taylorreihen entwickelt. => Die Beschleunigung lässt sich direkt aus $x(t)$, $x(t+\Delta t)$ und $x(t-\Delta t)$ berechnen (ohne Umweg über die Geschwindigkeit).

Methoden zur Berechnung der Systemenergie:

Aus der Energieverteilung eines Festkörpers lassen sich wichtige Parameter für die Simulation ableiten (z.B.: Energiebarrieren für die Monte-Carlo-Simulation oder das Potential für die Molekuldynamik-Simulation).

Um die Energie berechnen zu können, muss die Vielteilchen-Schrödingergleichung für die Elektronen bei vorgegebenen Positionen der Atomkerne gelöst werden. Diese direkte Lösung ist aussichtslos (Wellenfunktion hängt von allen Elektronenkoordinaten und vom Elektronenspin ab -> $4n$ Koordinaten) weshalb man sich mit Näherungen begnügen muss.

Dichtefunktionaltheorie: Soll die Berechnung der Systemenergie dadurch vereinfachen, dass das Problem mit Hilfe von Einelektronenwellenfunktionen formuliert wird, sodass für jedes Elektron eine separate Schrödingergleichung angeschrieben werden kann. Das Potential ist dann ein Funktional der lokalen Elektronendichte und hängt nicht mehr vom genauen Ort der Elektronen ab. Da die Elektronendichte aus den Wellenfunktionen der Elektronen berechnet wird und umgekehrt muss iteriert werden um Konsistenz zu erreichen.

Tight-Binding-Methode: Ist eine Vereinfachung der Dichtefunktionaltheorie. Die Einelektronen-Wellenfunktionen werden als gewichtete Summe von Orbitalfunktionen geschrieben, die jeweils einem Atom zugeordnet sind. Es wird keine Iteration durchgeführt => 100-1000mal schneller als DFT.

Empirisch interatomare Potentiale: Sind Funktionen der Abstände und Winkel der Atome, welche materialabhängige Parameter enthalten. Diese Parameter werden entweder empirisch ermittelt oder quantenmechanisch berechnet.

Beispiele: Morse-Potential (Zweikörper-Potential; nicht für Si geeignet)

Stillinger-Weber-Potential (berücksichtigt Zwei- und Dreikörperwechselwirkungen; ermöglicht die Beschreibung kovalenter Bindungen)

Anwendungsbeispiele:

Berechnung der Verteilung der kinetischen Energie im Si-Substrat (Energie jedes Atoms je nach Höhe mit unterschiedlich großen Kugeln dargestellt) bei Ionenbeschuss und die daraus resultierenden Punktdefekte. Eine Simulation von As mit 2keV in Si zeigt z.B. eine große Dichte der kinetischen Energie, die zum Aufschmelzen des Kristalls führt.

Simulation der Kollision eines Si-Clusters mit einem Substrat (lässt sich durch eine binary-collision-Simulation nicht beschreiben)

Aufbau einer Schicht mit MBE.

Implantationsschäden (siehe Frage 5):

Sind nicht vollständig durch das einfache Modell mit Ed beschreibbar, da einige zusätzliche Phänomene auftreten: => Atomistische Simulation!!

- **Massenabhängigkeit:** Die Schäden nehmen bei schwereren Ionen stärker zu als es das einfache Modell voraussagt => schwere Ionen erzeugen dichtere Stoßkaskaden, die Teilgebiete leicht amorphisieren (weniger Energie/Atom nötig)
- **Polyatomarer Effekt:** Schäden durch Moleküle größer als Summe der Schäden durch einzelne Atome => siehe oben.
- **Temperaturabhängigkeit:** Abnahme der Schäden mit wachsender Temperatur oberhalb von 150K => Punktdefekte sind oberhalb von 150K beweglich und können rekombinieren.
- **Dosisabhängigkeit:** Nichtlineare Zunahme der Schäden bei steigender Dosis => Stärker als linear: Kollaps des Kristallgitters
Schwächer als linear: Rekombination von Punktdefekten mit schon existierenden Defekten.
- **Dosisratenabhängigkeit:** Zunahme der Schäden mit wachsender Dosisrate oberhalb von 150K => weniger Zeit für Rekombination der Punktdefekte.