

4. Schichteigenschaften, Schichtcharakterisierung

4. 1. Schichtdicke

4.1.1. Einleitung

Um sicherzustellen, dass die mit verschiedenen Technologien hergestellten Beschichtungen den jeweiligen technischen Anforderungen genügen, sind entsprechende Untersuchungs-, Meß- und Prüfverfahren erforderlich. Die physikalischen Eigenschaften einer dünnen Schicht hängen naturgemäß sehr stark von deren Dicke ab. Die Bestimmung der Schichtdicke und der Beschichtungsrate ist daher eine grundlegende Aufgabe der Dünnschichttechnologie.

Bei vielen Anwendungen ist es notwendig, die aktuelle Schichtdicke bereits während der Beschichtung möglichst genau zu kennen (z.B.: optische Schichten). Man unterscheidet daher zwischen Meßverfahren, die bereits während des Beschichtungsprozesses, also in-situ, die Bestimmung der aktuellen Schichtdicke erlauben (Schichtdicken-Monitore) und Verfahren für eine nachträgliche Messung der Schichtdicke.

4.1.2. Gravimetrische Methoden:

4.1.2.1. Allgemeines:

Hierbei handelt es sich um Meßmethoden, die auf der Bestimmung einer Masse beruhen. Die Schichtdicke d läßt sich bei bekannter Dichte ρ aus der Masse m des Schichtmaterials berechnen, die auf der Fläche A niedergeschlagen wird:

$$d = m / (A\rho) \quad (4.1)$$

Es ist jedoch zu berücksichtigen, dass die Dichte einer Beschichtung unter Umständen wesentlich von der des Bulk-Materials abweichen kann (z. B. aufgrund von Porositäten oder implantierten Zwischengitteratomen). Für genauere Messungen ist eine Kalibrierung erforderlich.

4.1.2.2. Wägung

Die wohl einfachste Methode zur Bestimmung der Schichtdicke ist die Bestimmung des Massenzuwachses des beschichteten Werkstückes oder eines Probestättchens mit einer genauen Waage. Obwohl neben der oben bereits angesprochenen Schwierigkeit der nicht genau bekannten Dichte auch noch andere Probleme auftreten (z.B.: Kondensation von Wasserdampf aus der Luft), gelingt es, die Schichtdicke mit ausreichender Genauigkeit für viele praktische Anwendungen zu bestimmen.

4.1.2.3. Schwingquarz-Methode

Diese Methode ist das bei den PVD-Verfahren am häufigsten angewandte Prinzip zur in-situ-Messung und Kontrolle von Schichtdicke und Beschichtungsrate. Sie wird auch oft "Quarz Mikrowaage" (engl. Quartz Micro Balance, QMB) genannt. In den serienmäßig verfügbaren Ausführungen werden die Schichtdicke im Bereich 0,1nm bis 100µm und die Beschichtungsrate im Bereich 0,01 bis 100nm^s⁻¹ laufend digital angezeigt (siehe Abb. 4.1.).

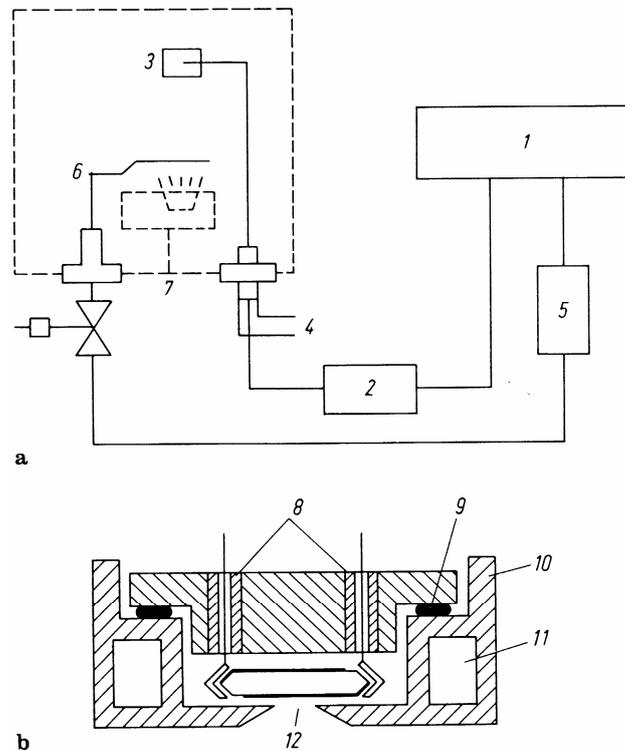


Abb. 4.1.: Schwingquarz-Methode zur Messung und Kontrolle der Schichtdicke und der Beschichtungsrate [1, p. 35]:

a Schema der Meßanordnung, **b** Schwingquarz-Meßkopf (schematisch)

- 1** Schwingquarz-Meßgerät mit D/A-Wandler; **2** Oszillator;
- 3** Meßkopf mit Schwingquarz; **4** Wasserkühlung des Quarzes;
- 5** Blendensteuerung; **6** Blende; **7** Dampfquelle;
- 8** elektrische Zuführungen; **9** Dichtung; **10** Kupferblock;
- 11** Wasserkühlung; **12** Apertur

Die von Sauerbrey bereits 1959 angegebene Methode beruht auf der Änderung der Resonanzfrequenz

$$f = N / d_q \quad (4.2)$$

eines Schwingquarzes, wenn dieser mit einer Schicht der Masse

$$\Delta m = \rho A d \quad (4.3)$$

belegt wird. Dabei ist d_q die Dicke des Quarzes und N seine Frequenzkonstante, die für den sogenannten AT-Schnitt (siehe Abb. 4.2.; geringer Temperaturkoeffizient) 1,67mmMHz beträgt. A ist die Fläche, ρ ist die Dichte und d ist die Schichtdicke des beschichteten Bereiches am Quarz.

Die Belegung Δm wirkt wie eine Dickenzunahme des Quarzes um

$$\Delta d_q = \Delta m / (\rho_q A_q) \quad (4.4)$$

wobei A_q , ρ_q die Fläche bzw. die Dichte des Quarzplättchens bedeuten. Die Resonanzfrequenz nimmt dann für $\Delta f \ll f$ proportional zu d ab, und mit

$$\frac{\Delta f}{\Delta d_q} = -\frac{N}{d_q^2} \quad (4.5)$$

gilt

$$-\Delta f = \frac{N \Delta m}{d_q^2 \rho_q A_q} = \frac{A f^2 \Delta m}{A_q N \rho_q A} = C \frac{\Delta m}{A} = C \rho d \quad (4.6)$$

wobei die Konstante $C = \frac{A}{A_q} \frac{f^2}{N \rho_q}$ ein Maß für die Wäageempfindlichkeit darstellt. Eine

genauere Ableitung liefert statt $\frac{A}{A_q}$ eine Funktion $F(\frac{A}{A_q})$.

Beispiel:

Für einen Quarz mit $f = 6\text{MHz}$ und $d_q = 0,28\text{mm}$ ist $C = 8\text{MHz}/(\text{kgm}^{-2})$. Für eine Schichtdicke $d = 0,1\text{nm}$ und $\rho = 10^4\text{kgm}^{-3}$ ist $\Delta m/A = 10^{-6}\text{kgm}^{-2}$.

Daraus ergibt sich $\Delta f = -8\text{Hz}$, was durchaus gut zu messen ist. Auch für $d = 1\mu\text{m}$ ist Δf mit -80kHz noch klein gegenüber f , so dass der Meßbereich von etwa $0,1\text{nm}$ bis zu einigen μm reicht. In vielen Fällen können mit einem Quarzplättchen 10 bis 100 Beschichtungsprozesse kontrolliert werden.

Eine Temperaturerhöhung ΔT des Quarzes infolge Wärmestrahlung des Verdampfers und/oder der Substrateizung sowie der Kondensationswärme des Aufdampfmaterials ändert die Resonanzfrequenz um $\Delta f = \beta f \Delta T$. Der Temperaturkoeffizient ist für den AT-Schnitt des Quarzes mit $\beta = -1 \cdot 10^{-6}\text{K}^{-1}$ am niedrigsten (siehe Abb. 4.2.), so dass bei $f = 6\text{MHz}$ die Temperaturabhängigkeit $\Delta f/\Delta T = -6\text{Hz/K}$ beträgt. Durch eine Wasserkühlung des

Quarzes wird der Temperatureinfluß so weit reduziert, dass Schichtdicken von 100nm mit einer Auflösung unter 0,1% bestimmt werden können.

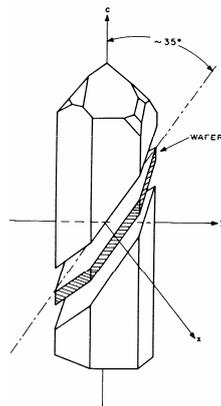


Abb. 4.2.: Quarz AT-Schnitt [5, p. 1-108]

Gute Linearität wird bei herkömmlichen Geräten nur bis zu Massenbelegungen von rund 10% gewährleistet. Schwingquartzgeräte der neueren Generation lassen wesentlich höhere Massenbelegungen zu. Nichtlinearitäten werden durch Mikroprozessoren rechnerisch weitgehend berücksichtigt. Benes (Inst. f. Allg. Physik) hat ein Verfahren zur Messung der Schallkennimpedanzen (z -Werte) angegeben. Durch Anregung unterschiedlicher Oberwellen und deren genaue Messung ist es möglich, die für hohe Massenbelegung notwendigen z -Werte zu bestimmen.

4.1.2.4. Mikrowägung

Diese sehr genaue, aber für die Praxis wenig geeignete Methode wird zur Kalibrierung anderer Verfahren verwendet. Alle Mikrowaagen zur Schichtdickenmessung arbeiten im Vakuum und mit Kompensation des Schichtgewichtes durch eine Gegenkraft. Die Kompensation erfolgt durch mechanische, optische oder elektrische (z.B. Drehspul-) Systeme, mit denen sowohl die Massendicken m/A als auch die Beschichtungsraten \dot{m}/A gemessen werden können (siehe Abb. 4.3.). Die Nachweisgrenze der Schichtdicke liegt bei 0,1nm und die obere Meßgrenze bei 1 μ m.

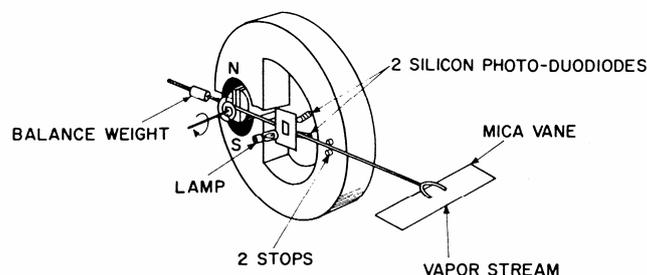


Abb. 4.3.: Mikrowaage [5, p. 1-104]

4.1.2.5. Dosierte Massenzufuhr

Viele Beschichtungsverfahren werden mit dosierter Massenzufuhr, d.h. einem über eine bestimmte Beschichtungsdauer konstant gehaltenen Massenstrom (in kg s^{-1}) bei definierten Abmessungen der Anordnung und konstant gehaltenen Prozessparametern ausgeführt. Das Verfahren erfordert selbstverständlich eine Kalibrierung, z.B. mittels Wägung oder chemischer Mikroanalyse, um den Zusammenhang zwischen Massendicke m/A am Substrat und zugeführter Masse zu ermitteln.

Nach dieser Kalibrierung genügt es dann in vielen Anwendungsfällen, die relevanten Prozeßparameter konstant zu halten, um innerhalb gegebener Toleranzen in gleichen Zeiten gleiche Schichtdicken zu erzielen.

Ist die zugeführte Masse pro Zeiteinheit m bzw. der Schichtdickenzuwachs pro Zeiteinheit d bekannt und ferner die Depositionsrate a konstant:

$$a = \left(\frac{dm}{dt} \right) / A = \rho \left(\frac{dD}{dt} \right) = \text{konst} \quad (4.7)$$

so wächst die Schichtdicke

$$D = \frac{at}{\rho} \quad (4.8)$$

proportional zur Beschichtungszeit t .

Wegen ihrer Einfachheit wird die Methode der dosierten Massenzufuhr in breitem Umfang praktisch eingesetzt: bei den thermischen Spritzverfahren, bei CVD-Prozessen, bei der galvanischen und chemischen Abscheidung, beim Auftragschweißen und vorallem bei vielen Anwendungen der Sputter- und Ionenplattiertechnik.

4.1.3. Optische Methoden

4.1.3.1. Allgemeines

Bei kaum einer anderen Anwendung als bei den optischen Schichten ist es notwendig, die Schichtdicke bereits während des Aufbringens möglichst genau zu messen. Es werden daher vorwiegend Schichtdickenmonitore eingesetzt. Speziell bei Vielschichtsystemen ("Multi-Coated Optics", Interferenzfilter u.ä.) sind Monitore unentbehrlich und werden auch in geschlossene Regelkreise (z.T. über komplexe Computerprogramme) eingebunden.

4.1.3.2. Photometer-Methode

Diese Methode wird vor allem bei PVD-Prozessen zur Herstellung von Ein- und Mehrschichten für optische Anwendungen eingesetzt, weil sie die optische Dicke nd mißt und daher Änderungen der Brechzahl n während der Beschichtung durch entsprechende Änderungen von d kompensiert werden können.

Mit dem Photometer (siehe Abb. 4.4.) wird die Intensität des an den beiden Grenzflächen der Schicht reflektierten und die des durch die Schicht hindurchgehenden Lichtes gemessen. Diese Intensitäten bestimmen, auf die einfallende Intensität bezogen, den Reflexionsgrad R bzw. den Transmissionsgrad T der Schicht. Das einfallende Licht passiert

ein Interferenzfilter, ist also monochromatisch und wird durch einen Chopper moduliert, so dass Störungen durch Fremdlicht vermieden werden. Der Lichtstrahl wird auf das Substrat oder ein Testglas gerichtet, das sich in einem Testglaswechsler befindet. Die Lichtintensitäten werden mit Photomultipliern gemessen. Computer ermöglichen die vollautomatische Steuerung und Regelung der Beschichtungsprozesse (wichtig bei Vielschichtsystemen).

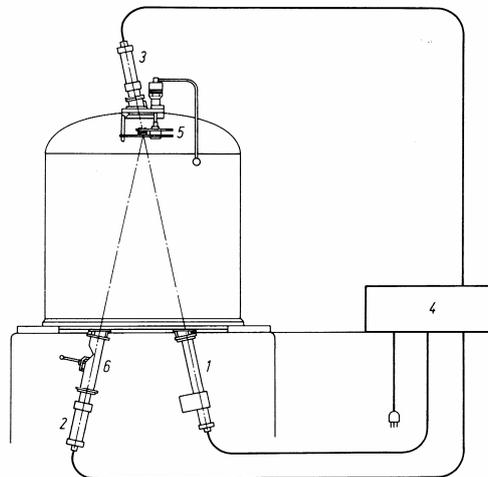


Abb. 4.4.: Photometer-Methode zur Messung der Schichtdicke [1, p. 38]:

- 1** modulierte Lichtquelle; **2** Empfänger für das an der Schicht reflektierte Licht;
- 3** Empfänger für das durch die Schicht transmittierte Licht;
- 4** Anzeige-Instrument; **5** Testglaswechsler; **6** Strahlableitung

Obgleich sich die Schichtdicke d aus R und T bei gegebenen Werten der optischen Konstanten n und k (d. i. Realteil und Imaginärteil des Brechungsindex) berechnen lässt, ist es in der Praxis oft einfacher, d aus Kurven zu entnehmen (siehe Abb. 4.6.). Bei absorptionsfreien oder schwach absorbierenden Schichten ändern sich R und T infolge Interferenzwirkung periodisch mit wachsendem d , falls Schicht und Substrat unterschiedliche Brechzahlen n haben (siehe Abb. 4.5.).

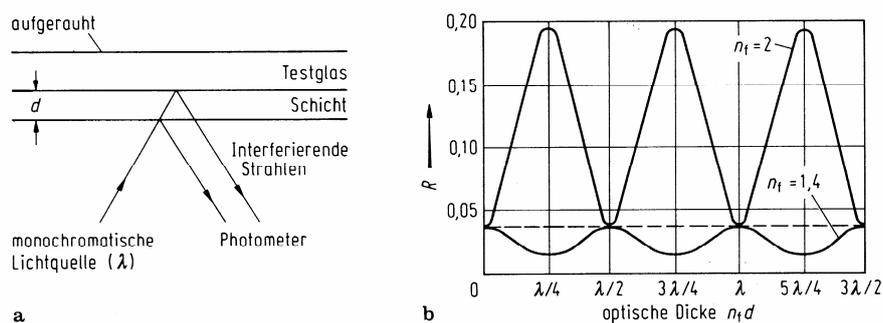


Abb. 4.5.: Zweistrahlinterferenz an einer dielektrischen, absorptionsfreien Schicht [1, p. 39]:

- a** Anordnung; **b** Reflexionsgrad R als Funktion der optischen Dicke $n_f d$ der Schicht bei verschiedenen Brechzahlen n_f . Brechzahl des Testglases $n_s = 1,5$

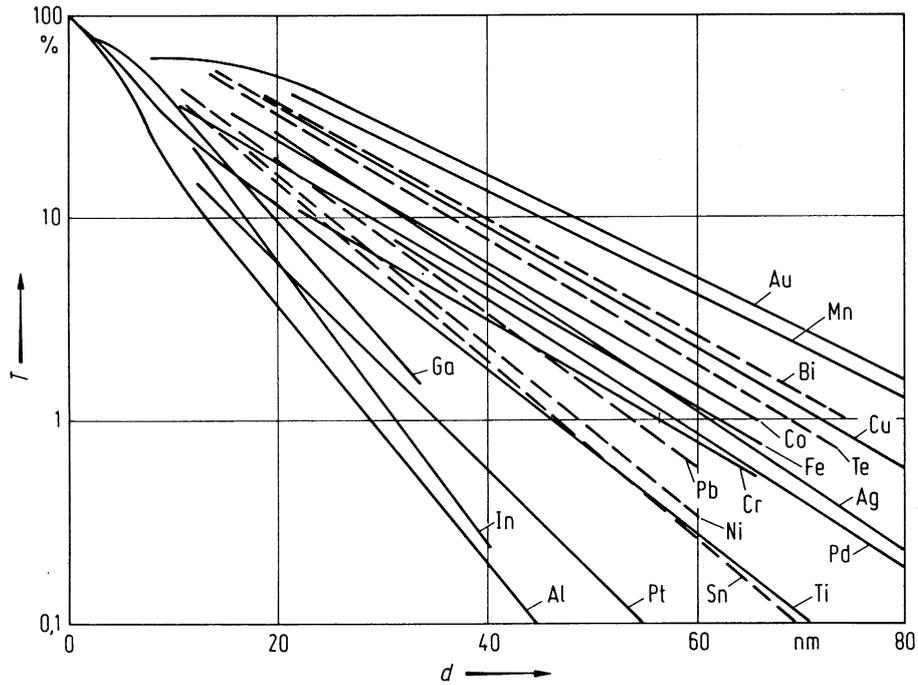


Abb. 4.6.: Transmissionsgrad T in Abhängigkeit von der Schichtdicke d , gemessen bei 550nm; Substrat: Glas [1, p. 39]

Als praktisches Beispiel ist in Abb. 4.7. der Verlauf von R und T bei der Herstellung eines hochreflektierenden Spiegels aus alternierenden hoch- (ZnS) und niedrigbrechenden (MgF_2) $\lambda/4$ -Schichten zu sehen.

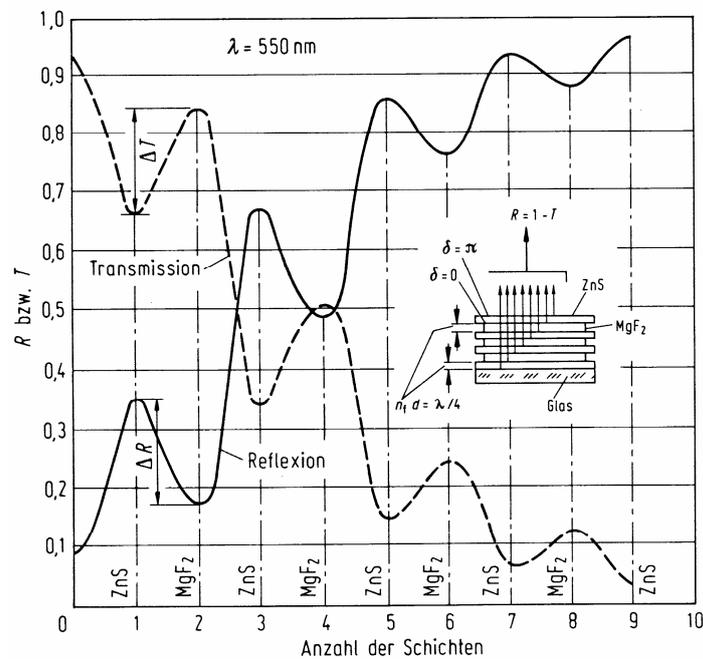


Abb. 4.7.: Reflexionsgrad R und Transmissionsgrad T eines Systems von alternierend aufgedampften hoch- (ZnS, $n=2,3$) und niedrig- (MgF_2 , $n=1,38$) brechenden $\lambda/4$ -Schichten [1, p. 40]

4.1.3.3. Tolansky-Interferometer

Die Funktionsweise dieses Vielstrahl-Interferometers basiert zunächst auf dem Prinzip, dass ein dünnes, transparentes Glasplättchen auf eine hochreflektierende Oberfläche gelegt wird, wobei es mit der Oberfläche einen Winkel α einschließt. Wird diese Anordnung mit monochromatischem Licht bestrahlt, so bildet sich durch Vielfachinterferenz ein Muster aus äquidistanten Interferenzlinien. Je mehr Teilstrahlen zur Überlagerung kommen, desto schärfer sind die Interferenzlinien. Ihr Abstand hängt vom Winkel α ab. Die Dicke einer Schicht wird mittels dieses Prinzips folgendermaßen bestimmt: Zunächst wird ein bis zum Substrat reichender Kratzer in die Schicht geritzt (es kann auch während der Beschichtung ein Teil des Substrates abgedeckt werden, wodurch eine Stufe entsteht). Danach wird die Probe mit einer hochreflektierenden Schicht versehen. Durch den Kratzer oder die Stufe ändert sich der Abstand Reflektor/Interferenzplättchen, was sich in einer Versetzung der Interferenzlinien ausdrückt. Damit ergibt sich die in Abb.4.8. dargestellte Situation. Die Schichtdicke ist gegeben durch

$$d = \Delta N \frac{\lambda}{2} \quad (4.9)$$

wobei ΔN die Anzahl (oder Teil) der Linien ist, um die die Interferenzlinien durch die Stufe oder den Kratzer versetzt werden.

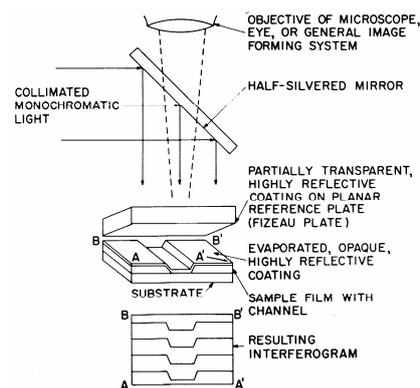


Abb. 4.8.: Schematische Darstellung des Tolansky-Interferometers [5, p. 11-8]

Tolansky-Interferometer sind kommerziell als Zusatzgeräte für Lichtmikroskope erhältlich. Dem Vorteil der einfachen und schnellen Bedienung steht der Nachteil gegenüber, dass eine geeignete Stufe oder ein Kratzer vorhanden sein muß. Zudem muß diese Stufe mit einer hochreflektierenden Schicht (zumeist Ag) überzogen werden. Die Auflösung liegt bei ca. ± 1 nm, wenn mit streng monochromatischem Licht und hochreflektierenden Probe- und Referenzplättchen gearbeitet wird.

4.1.3.4. FECO-Methode

Die FECO (Fringes of Equal Chromatic Order)-Methode bietet eine noch höhere Genauigkeit als das Tolansky-Interferometer. Bei sorgfältiger Messung kann durchaus eine Auflösung von 0,1nm erreicht werden. Die Funktionsweise ist in Abb 4.9. zu sehen. Paralleles weißes Licht fällt auf die Kombination aus Probe und Referenzplättchen. Das reflektierte Licht wird über einen semitransparenten Spiegel auf den Eingangsspalt eines Spektrographen fokussiert. Das Bild der Stufe muß normal zum Spektrographenspalt stehen. Die so erhaltenen Spektren (bzw. Interferogramme) sind in Abb.4.10. zu sehen. Dunkle Interferenzlinien treten bei den Wellenlängen

$$\lambda = 2t / N \quad (4.10)$$

auf, wobei N die Ordnung angibt.

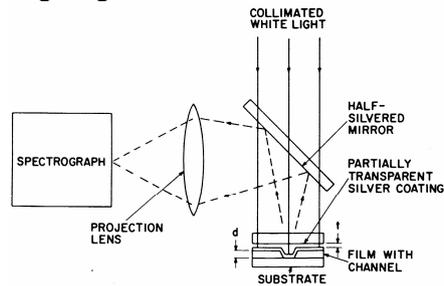


Abb. 4.9.: Schematische Darstellung der FECO-Methode [5, p. 11-9]

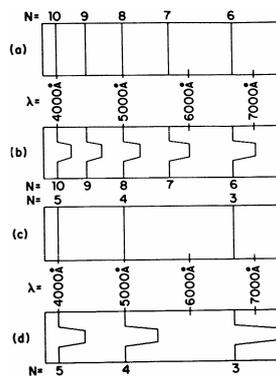


Abb. 4.10.: **a** Interferogramm für einen Plattenabstand $t=2\mu\text{m}$
b Interferogramm für Fall a, aber Kanal mit $d=100\text{nm}$ Tiefe
c Interferogramm für einen Plattenabstand $d=1\mu\text{m}$
d Interferogramm für Fall b, aber Kanal mit $d=100\text{nm}$ Tiefe
 [5, p. 11-9]

Da t in der Praxis jedoch nicht bekannt ist, muß die Auswertung auf folgende Art durchgeführt werden: Wenn zur Ordnung N_1 die Wellenlänge λ_1 gehört, so ist $N_1 + 1$ die Ordnung mit einer kürzeren Wellenlänge λ_0 . Es gilt dann

$$N_1 \lambda_1 = (N_1 + 1) \lambda_0 = 2t \quad (4.11)$$

und

$$N_1 = \frac{\lambda_0}{\lambda_1 - \lambda_0} \quad (4.12)$$

Damit kann nun der Abstand t bestimmt werden

$$t = \frac{N_1 \lambda_1}{2} = \frac{\lambda_1 \lambda_0}{2(\lambda_1 - \lambda_0)} \quad (4.13)$$

Zur Bestimmung der Schichtdicke d wird die Versetzung der Linien der Ordnung N_1 herangezogen. Wenn die Linien bei der Wellenlänge λ_1 zu liegen kommen, muß gelten

$$t + d = \frac{N_1 \lambda_2}{2} \quad (4.14)$$

Damit folgt für die Schichtdicke

$$d = \frac{N_1 \lambda_2}{2} - \frac{N_1 \lambda_1}{2} = \frac{\lambda_0 (\lambda_2 - \lambda_1)}{2(\lambda_1 - \lambda_0)} \quad (4.15)$$

4.1.3.5. Weitere optische Dickenmeßgeräte

Neben den oben dargestellten Verfahren kann die optische Schichtdickenbestimmung noch durch eine Vielzahl anderer Methoden erfolgen. Einige Beispiele: Das Nomarsky-Interferometer, welches als Zusatzgerät für Lichtmikroskope erhältlich ist; die VAMFO (Variable Angle Monochromatic Fringe Observation) -Methode gestattet die in-situ-Bestimmung der Schichtdicke und der Brechungsindizes; mit der Ellipsometrie ist gleiches möglich.

4.1.4. Direkte Methoden

4.1.4.1. Allgemeines

Bei den direkten Methoden zur Schichtdickenbestimmung erfolgt die Messung entweder über eine mechanische Abtastung oder über die Beobachtung in Mikroskopen.

4.1.4.2. Stylus-Methode

Die Schicht muß als Stufe auf einem (ebenen) Substrat vorliegen. Eine Diamantnadel (Krümmungsradius der Spitze ca. 10µm) wird mit konstanter Geschwindigkeit über die zu messende Fläche gezogen und die Höhe der Stufe mit einem elektromechanischen Pick-up-System gemessen. Voraussetzungen für eine genaue Messung sind eine hinreichende Härte der Schicht und ein hinreichend ebenes Substrat. Um Beschädigungen der Probe zu vermeiden, kann die Belastung der Nadel bis auf 10µN verringert werden.

Der Meßbereich liegt zwischen etwa 5nm und einigen 10µm bei einer Auflösung von wenigen Å (im empfindlichsten Bereich). Eine weitere Anwendung der serienmäßig erhältlichen Geräte ist die Reliefdarstellung von Oberflächen und die Erfassung der Rauigkeit.

4.1.4.3. Messung mit dem Licht- und Elektronenmikroskop

Zur Bestimmung der Schichtdicke werden mit dem Lichtmikroskop z.B. metallographische Schiffe, mit dem Transmissions-Elektronenmikroskop (TEM) Oberflächenabdrücke einer durch die Schicht erzeugte Stufe oder Querpräparate, und mit dem Raster-Elektronenmikroskop (REM) eine Stufe selbst oder eine Bruchfläche der Schicht vermessen.

Die Auflösung und der Meßbereich hängt vom Gerät und der Vergrößerung ab. Im Lichtmikroskop können Schichtdicken bis in den mm-Bereich mit einer Auflösung von 0,1µm gemessen werden, im REM erreicht die Auflösung Werte von ca. 5nm und im Hochauflösungs-TEM 0,1nm.

4.1.5. Schichtdickenbestimmung mittels elektrischer oder magnetischer Größen

4.1.5.1. Widerstandsmethode

Diese Methode wird bei PVD-Prozessen vielfach zur Bestimmung der Dicke metallischer Schichten eingesetzt. Als Monitorelement wird eine isolierende Platte mit zwei parallelen Kontaktstreifen benutzt, zwischen denen die Schicht über eine Maske niedergeschlagen wird. Der Widerstand R als Maß für die Schichtdicke wird mit einer Brückenschaltung kontrolliert, und die Kondensationsrate mittels eines elektronischen Differenziergliedes.

Mit Hilfe eines Nullindikators wird die Beschichtung gestoppt, wenn der Sollwert der Schichtdicke erreicht ist. Der Meßbereich liegt zwischen 1nm und 10 μ m. Anwendungen sind Metallschichten für integrierte Schaltungen, Widerstandsschichten aus NiCr, metallisierte Folien u.ä..

4.1.5.2. Kapazitätsmethode

Eine zur vorangehenden analoge Methode zur Bestimmung der Dicke isolierender Schichten besteht darin, ein Monitorelement mit kammförmigen, ineinandergreifenden Flächenelektroden zu verwenden und die beim Beschichten auftretende Kapazitätsänderung zu messen.

4.1.5.3. Wirbelstrommeßmethode

Die Dicken von Isolationsschichten auf Nichteisenmetall (NE)-Substraten oder die von NE-Schichten auf Isolierstoff-Substraten können mit dieser Methode gemessen werden. Die Meßgröße ist z.B. die an einer HF-Spule liegende Spannung, die durch die Wirbelströme im NE-Metall modifiziert wird. Da die Meßgröße außer von der Schichtdicke noch von der elektrischen Leitfähigkeit des NE-Metalls abhängt, bedarf es einer Kalibrierung. Die Anwendungen liegen in der Aluminisierung von Fernsehbildröhren oder der Metallisierung von Kunststoffen.

4.1.5.4. Magnetische Meßmethode

Dieses Verfahren wird auf Schichten angewendet, die sich auf einer planen Unterlage aus ferritischem Stahl befinden. Die Methode beruht darauf, dass die Haftkraft eines auf die Schicht (aus NE-Metall, Lack, Kunststoff etc.) gesetzten Magneten von der Schichtdicke abhängt. Da das Messergebnis auch von der Permeabilität des Stahls abhängt, ist eine Kalibrierung erforderlich. Auch Nickel als Schichtmaterial ist nach entsprechender Kalibrierung einer Dickenmessung zugänglich.

4.1.6. Dicken-Messung nach Teilchen-Wechselwirkung

4.1.6.1. Verdampfungsraten-Monitor

Schichtdicken- bzw. Beschichtungsraten-Monitore wurden speziell für Anwendungen der Aufdampftechnik entwickelt. Dabei wird zur Kontrolle der Dampfstromdichte der Dampfstrom in der Nähe des Substrates durch Elektronenstoß ionisiert und der Ionenstrom gemessen. Einige ältere - kaum mehr in Verwendung befindliche - Ausführungen sind in Abb. 4.11. zu sehen.

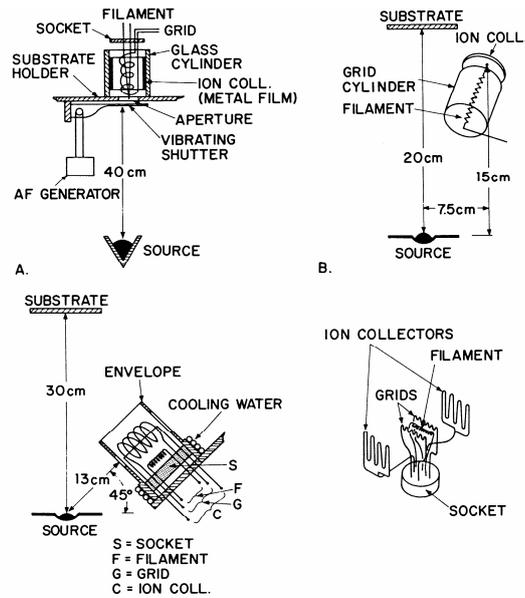


Abb. 4.11.: Dicken- und Raten-Monitore nach dem Ionisationsprinzip [5, p. 1-98]

Neuere Geräte analysieren den Ionenstrom in einem Quadrupol-Massenspektrometer. Auf diese Weise können die Verdampfungsraten simultan verdampfender Materialien auf ihrem Sollwert gehalten werden.

4.1.6.2. Weitere Methoden

- ... Beta (Elektronen)-Rückstreuung: besonders zur Dickenmessung an dünnen Edelmetallschichten und an Metallschichten auf gedruckten Schaltungen geeignet.
- ... Röntgen-Fluoreszenzmethode: auch zur Bestimmung der Dicke von Mehrschichten geeignet. Da diese Methode zerstörungsfrei ist und einen hohen Durchsatz ermöglicht, wird sie vor allem in der Qualitätskontrolle (z. B. von Verschleißschutzschichten) verwendet.
- ... Tracer-Methoden: entweder das Schichtmaterial oder das Substrat muß radioaktive Atome enthalten.

4.2. Schichtrauhigkeit

4.2.1. Einleitung

Bei immer geringer werdenden Schichtdicken beginnt die Oberflächenstruktur eine immer größere Rolle zu spielen. Im Extremfall sehr dünner Schichten kann die Oberflächenrauigkeit in der Größenordnung der Schichtdicke liegen und damit Einfluss auf alle Schichteigenschaften, seien sie mechanisch, elektrisch, magnetisch oder optisch, haben. Auch Schichtmorphologie, innere Struktur, Textur und Kristallinität sind stark mit der Rauigkeitsentwicklung verbunden. Dieses Unterkapitel soll daher einen kurzen Überblick über Rauigkeitstypen, deren Entstehungsmechanismus, Rauigkeitsmessung und die Quantifizierung von Rauigkeiten geben.

4.2.2. Rauigkeitstypen

Ganz allgemein ist der Grund für die Entstehung von Rauigkeiten bei der Abscheidung von Beschichtungen die endliche Ausdehnung der Teilchen, aus denen sich die Schicht zusammensetzt sowie deren zufälliges, zeitlich und örtlich unkorreliertes Auftreffen an der Wachstumsfront. Dabei müssen die "Grundbausteine" der Schicht nicht notwendigerweise Einzelatome sein, wie es bei PVD-Schichten meist der Fall ist. Es kann sich auch um komplexe Moleküle (z. B. bei organischen Schichten), um Mehrteilchenaggregate wie z. B. bei der Cluster-Deposition oder um makroskopische Aggregate wie die Metall- oder Keramiktröpfchen beim thermischen Spritzen handeln.

4.2.2.1. Stochastische Rauigkeit

Das einfachst mögliche Modell zum Aufbau von Rauigkeiten besteht darin, nicht näher spezifizierte Teilchen endlicher Ausdehnung a an zufälligen Positionen eines quadratischen Gitters zu zufälligen Zeitpunkten auf eine zunächst völlig glatte Oberfläche auftreffen zu lassen. Die Auftreffrichtung der Teilchen ist dabei immer senkrecht zum Substrat. Ein Teilchen wird zum Materialverbund der gesamten Schicht (im folgenden auch "Aggregat" genannt) dazugezählt, sobald es einen nächsten Nachbarn unter sich hat. Teilchen, welche zum Aggregat zählen, aber an ihrer Oberseite keinen nächsten Nachbarn haben zählen zur sogenannten "aktiven Oberfläche" und bilden die Wachstumsfront der Schicht. Nur an Teilchen der aktiven Oberfläche können sich neue Teilchen am Aggregat anlagern. Die Situation ist für ein eindimensionales Profil $h(x)$ in Abb. 4.12. dargestellt.

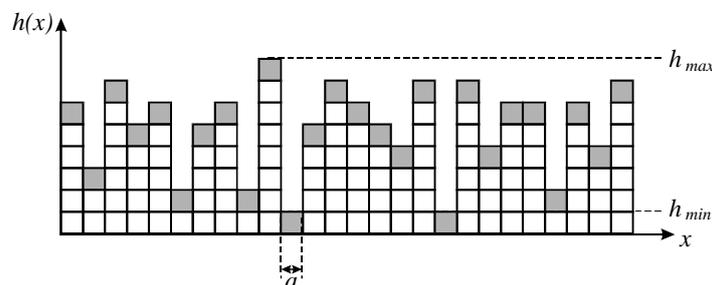


Abb. 4.12.: Stochastisches Wachstum: grau markierte Teilchen stellen die aktive Oberfläche dar. h_{max} ist der maximale, h_{min} der minimale Höhenwert

Der oben skizzierte einfache Wachstumsmechanismus führt zu einem Aggregat, welches aus nebeneinanderliegenden Säulen völlig unkorrelierter Höhe besteht. Unkorreliert bedeutet, dass aus der Höhe einer Säule nicht auf die Höhe der benachbarten Säulen geschlossen werden kann. Das Aggregat ist unterhalb der aktiven Oberfläche dicht, enthält also keine Volumsfehlstellen, wie aus Abb. 4.12. ersichtlich ist. Modelle der Rauigkeitsentwicklung, bei der die aktive Oberfläche eine einwertige Funktion der Position über dem Substrat ist, werden als "Solid on Solid" (SOS)-Modelle bezeichnet. Überhänge

oder abgeschlossene Poren im Bulkmaterial kommen in diesen Modellen per definitionem nicht vor.

Die aktive Oberfläche der Schicht wird mit der Zeit immer rauher, wobei der Begriff Rauigkeit hier noch nicht quantifiziert ist, aber im wesentlichen ein Maß für die Differenz zwischen dem minimalen und dem maximalen Höhenwert im Aggregat darstellt. Die zeitliche Entwicklung der Rauigkeit ist abhängig vom spezifischen Modell der Rauigkeitsentwicklung und erlaubt somit die Zuordnung von experimentell ermittelten Rauigkeitswerten zu verschiedenen Wachstumsmodellen.

4.2.2.2. Selbstähnliche Oberflächen

Bereits geringe Modifikationen des im Vorigen beschriebenen Modells des stochastischen Wachstums haben weitreichende Konsequenzen für die detaillierte Gestalt der aktiven Oberfläche. Dabei soll zunächst der SOS-Charakter der Wachstumsmodelle erhalten bleiben. Eine erste Annäherung an realistische Bedingungen kann dadurch realisiert werden, dass Teilchen innerhalb einer vordefinierten Umgebung um ihren Auftreffpunkt Bindungspositionen mit einer möglichst hohen Zahl an nächsten Nachbarn einnehmen. Dieser Mechanismus ist ein Modell für Teilchenmigration entlang der Oberfläche durch Oberflächendiffusion sowie die Anlagerung an Positionen mit möglichst hoher Nächster-Nachbar-Koordination wie z. B. Stufen, Kink-Sites oder Punktdefekten.

Die beschriebenen Relaxationsmechanismen führen zum Auftreten lateraler Korrelationen in der aktiven Oberfläche. Das bedeutet, dass sich Höhenwerte in der Umgebung einer bestimmten Position an der Oberfläche nicht abrupt ändern können. Innerhalb der, als Modellparameter gegebenen, Zone, in welcher die Teilchen günstige Bindungspositionen aufsuchen können, kann man also davon ausgehen, dass ähnliche Höhenwerte vorliegen.

Ein wesentlicher Unterschied zum Modell des stochastischen Wachstums ist, dass die Rauigkeit der Oberfläche nicht nur von der Beschichtungszeit abhängt, sondern auch von der Länge des Messintervalls, in dem sie bestimmt wird. Dieses Verhalten zeigt Abb. 4.13. für eine Oberfläche nach längerer Beschichtungszeit.

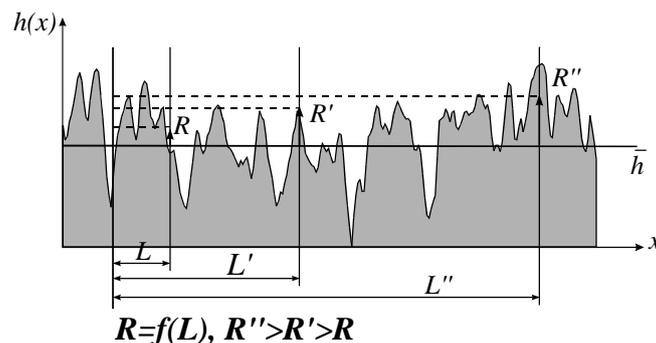


Abb. 4.13.: Abhängigkeit der Rauigkeit R vom Messintervall L bei einer selbstähnlichen Oberfläche. \bar{h} ist die mittlere Höhe der Oberfläche

Solche Oberflächen werden als "selbstähnlich" bezeichnet, da ihre Gestalt zu einem Zeitpunkt t in einem Längenintervall L sich zur Gestalt zu t' in L' nur durch einen zeitlichen Skalierungsfaktor T und einen räumlichen Skalierungsfaktor X unterscheidet.

4.2.2.3. Porenbildung und nicht "Solid-on-Solid"-Oberflächen

Selbst das Modell des stochastischen Wachstums kann durch eine simple Modifikation seinen Solid-on-Solid Charakter verlieren. Fügt man ein Teilchen nämlich zum Aggregat hinzu sobald es einen nächsten Nachbarn hat, der nicht notwendigerweise unter dem auftreffenden Teilchen liegen muss, so kommt es zur Ausbildung von Überhängen und Poren, wie Abb. 4.14. zeigt.

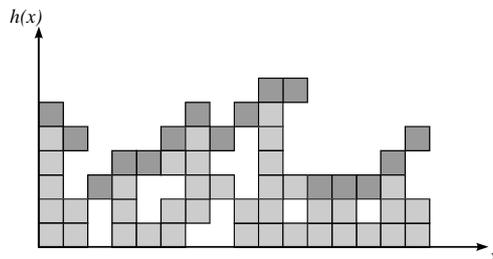


Abb. 4.14.: Ballistische Aggregation: hellgrau markierte Teilchen sind Teil des Aggregates, die aktive Oberfläche ist dunkelgrau markiert. Man erkennt deutlich die Existenz abgeschlossener Poren.

Das so entstehende Wachstumsmodell wird als "ballistische Aggregation" bezeichnet und weist, im Gegensatz zum stochastischen Wachstum, laterale Korrelationen in der selbstähnlichen aktiven Oberfläche auf. Die Funktion, welche die Aktive Oberfläche beschreibt ist allerdings, wie aus Abb. 4.14. ersichtlich, immer noch einwertig. Erst wenn all jene Teilchen zur Oberfläche gezählt werden, die über einen durchgehenden Bereich mit dem Vakuum über dem höchsten Punkt des Aggregates verbunden sind, wird der nicht Solid-on-Solid Charakter des Modells offensichtlich.

Auch ein Abgehen von der Forderung nach senkrechter Teilcheninzidenz führt über den Mechanismus der Abschattung zur Bildung langgezogener, tiefer Poren in der wachsenden Schicht, wenn die Beweglichkeit der Teilchen entlang der Wachstumsfront hinreichend gering ist. Treffen die Beschichtungsteilchen aus einem sehr breiten Winkelbereich auf eine strukturierte Wachstumsfront auf, so ist die normale Wachstumsgeschwindigkeit v_n an einer Spitze höher als in einem Tal, wie Abb. 4.15.b zeigt. Ist die Einfallswinkelverteilung hingegen schmal, so tritt dieser Effekt nicht auf (Abb. 4.15.a).

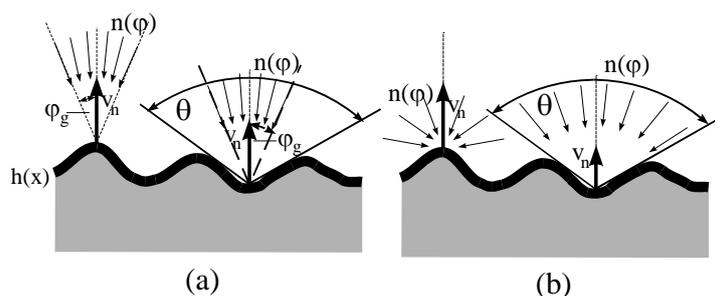


Abb. 4.15.: Abschattungseffekte bei verschiedenen Einfallswinkelverteilungen $n(\varphi)$;

θ ist der Sichtwinkel aus einem Tal

a Teilchen fallen nur aus $\varphi < \varphi_g$ ein

b Teilchen fallen unter allen Winkeln φ ein;

Die Form des Startprofils $h(x)$ kann durch Inselbildung in den Anfangsphasen des Schichtwachstums, stochastische oder ballistische Aufrauung oder durch willkürliche Profilierung der Substratoberfläche gegeben sein.

Durch das raschere Wachstum der Spitzen wird der Abschattungseffekt immer mehr verstärkt und es kommt zur Bildung säulenförmiger Schichtbereiche, welche durch tiefe,

schmale Spalten voneinander getrennt sind. Abb. 4.16. veranschaulicht diesen Prozess für ein sinusförmiges Anfangsprofil.

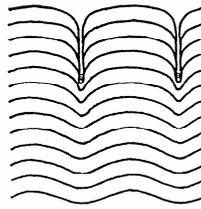


Abb. 4.16.: Auswirkung der Abschattung für ein sinusförmiges Anfangsprofil [7, p. 692]

Abb. 4.17. zeigt, dass Abschattung im Falle selbstähnlicher Oberflächen auch zur Ausbildung abgeschlossener Poren in einer Schicht führen kann. Daher wird auch abschattungsdominiertes Schichtwachstum zu den nicht-SOS Mechanismen gezählt.

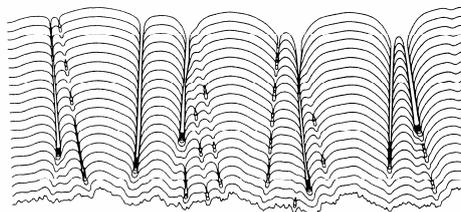


Abb. 4.17.: Ausbildung geschlossener Porositäten durch Abschattung bei einem selbstähnlichem Anfangsprofil [8, p. 148]

Breite Teilcheneinfallswinkelverteilungen treten insbesondere bei Sputterprozessen oder Plasmaunterstützten PVD-Verfahren auf, wo die Beschichtungsteilchen relativ ausgedehnte Bereiche erhöhten Gasdruckes durchqueren müssen und daher oft gestreut werden. Generell gilt, je höher der Gasdruck desto breiter die Einfallswinkelverteilung bei unverändertem Abstand zwischen Quelle und Substrat. Daher tritt das abschattungsdominierte Wachstum im Falle von Sputterprozessen bei hohen Arbeitsgasdrücken verstärkt auf, wie es auch aus den Strukturzonenmodellen des Schichtwachstums bekannt ist (siehe Abschnitt 4.3.2.).

4.2.2.4. Welligkeiten und definierte Strukturen

Material- und Schichtoberflächen können naturgemäss auch durch eine Unzahl von anderen Mechanismen Rauigkeiten bzw. ganz allgemein Strukturierungen aufweisen. Selbst eine ideale Einkristalloberfläche zeigt Korugationen durch die regelmäßige Anordnung der Atome. Inselbildung in den ersten Phasen der Schichtbildung kann zur Ausbildung mehr oder weniger regelmässiger dreidimensionaler Strukturen führen.

Abtragende Verfahren wie Drehen, Fräsen oder Honen erzeugen verschiedenste, oftmals sehr komplexe wellige Strukturen, die nur schwer quantitativ zu beschreiben sind. Schlussendlich können, wie z. B. in der Mikroelektronik, Oberflächen gezielt mit wohldefinierten, regelmäßigen Strukturen versehen werden. Dies geschieht meist mit photolithographischen Methoden, die meist auf eine selektive Materialabtragung hinauslaufen ("Top-Down-Methoden"). In letzter Zeit werden definierte Mikro- und Nanostrukturen allerdings auch durch geschicktes Ausnutzen von Schichtwachstumsmechanismen dargestellt ("Bottom-Up-Methoden").

4.2.3. Rauigkeitsmessung

Die Grundlage jeder Art von Rauigkeitsmessung ist, im Falle einer zweidimensionalen Oberfläche, die Erfassung aller Höhenwerte $h(x,y)$. Theoretisch muss die Voraussetzung einer Solid-on-Solid Oberfläche nicht gemacht werden, in der Praxis liefern aber so gut wie alle Methoden nur die obersten Höhenwerte. Bereiche der Oberfläche, welche sich unter Überhängen befinden, werden nicht abgebildet. Weiters ist die intrinsische Rauigkeit einer Oberfläche, das ist jene Rauigkeit, welche nur auf die Art und die Parameter des Beschichtungsprozesses zurückzuführen ist, immer mit der Rauigkeit, Neigung oder Welligkeit des Substrates überfaltet. Die Trennung der beiden Morphologien ist nicht trivial und muss fast immer vom Beobachter vorgenommen werden. Moderne Messsysteme bieten dazu eine Vielzahl von Datenmanipulationsmöglichkeiten (z. B. korrigieren von beliebig geneigten Substraten oder Filtern von langwelligen Komponenten durch Fourieralgorithmen) an, welche aber nur von einer geübten Person angewendet werden sollten, da es sonst zu massiven Verfälschungen der Messergebnisse kommen kann. Grundsätzlich können Daten im Realraum oder im reziproken Raum erfasst werden. Beide Methoden werden im folgenden kurz beschrieben.

4.2.3.1. Stylus- oder Sondenmethoden

Dieser Satz von Messmethoden erlaubt die Erfassung von Oberflächendaten im Realraum durch die Annäherung einer Sonde ("Stylus") an eine Oberfläche. Die laterale Auflösung ist dabei von der Spitzenverrundung und auch von der Form der Sonde begrenzt. Generell wird so eine Oberfläche, welche durch eine kontinuierliche Funktion $h(x,y)$ in einer Fläche der Seitenlänge L gegeben ist, in einen diskreten Satz von Höhenwerten $h_i(x_i \pm \Delta x/2, y_i \pm \Delta y/2)$, welche auf Flächen der Ausdehnung Δx und Δy definiert sind, umgewandelt. Dieser Sachverhalt ist schematisch in Abb. 4.18. dargestellt.

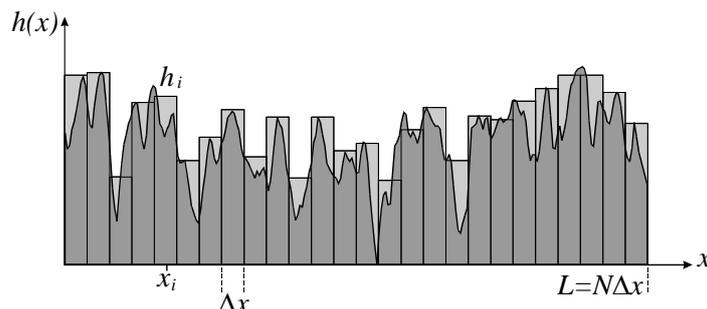


Abb. 4.18.: Umwandlung eines kontinuierlichen Profils $h(x)$ in diskrete Höhenwerte h_i , welche in einem Intervall Δx definiert sind. Die h_i kommen durch eine Überfaltung von $h(x)$ mit der Geometrie der Sonde zustande. Man erkennt deutlich den Verlust von Details der Oberflächenmorphologie durch den Messvorgang.

Die Intervalle Δx und Δy werden als "Sampling Interval" bezeichnet. Die abgerasterte Fläche der Seitenlänge L enthält $L/\Delta x$ Messpunkte in x -Richtung und $L/\Delta y$ Messpunkte in y -Richtung. Sind Δx und Δy gleich, so enthält die gesamte Messfläche $N = (L/\Delta x)^2$ Messpunkte. L begrenzt die maximale laterale Ausdehnung von Oberflächenmerkmalen im Messbereich, Δx und Δy bestimmen die minimale laterale Ausdehnung von Oberflächenmerkmalen. Die Anzahl N der Messpunkte wird bei der Quantifizierung der Rauigkeit (siehe Abschnitt 4.2.4) eine wesentliche Rolle spielen.

Die Erfassung von Oberflächenprofilen war bis ca. 1985 weitgehend auf eindimensionale Oberflächenschriebe begrenzt. Die rapide Steigerung der Geschwindigkeit und, vor allem, der Speicherkapazität moderner Rechner seit Mitte der 80-er Jahre des vorigen Jahrhunderts erlaubt nun die problemlose Speicherung und Manipulation von zweidimensionalen Oberflächenbereichen mit hoher Messpunktanzahl N . Im folgenden

sollen die wichtigsten Sondenmikroskopischen Techniken mit ihren Vor- und Nachteilen kurz besprochen werden:

Stylusprofilometer: Dieses bereits in Abschnitt 4.1.4.2 kurz beschriebene Instrument ist auch die älteste Form aller Messgeräte zur Erfassung von Oberflächenprofilen. Eine Diamant- oder Saphirspitze gleitet über eine Oberfläche. Ihre Auslenkung wird über eine Induktive Messbrücke in ein elektrisches Signal umgewandelt. Die vertikale Auflösung einer solchen Anordnung kann bis in den Sub-Nanometerbereich gehen. Die laterale Auflösung (sampling interval) liegt in der Größenordnung von sub-Millimeter bis μm . Als Vorteile des Gerätes seien die großen Scanlängen, L , von bis zu einigen 10 cm genannt. Das macht Stylusprofilometer insbesondere in der Halbleiterindustrie zu gebräuchlichen Messinstrumenten, wo sie z. B. zur Bestimmung der Krümmung von großen Halbleiterwafern eingesetzt werden. Als Nachteile können die relativ hohe mechanische Belastung der Oberfläche durch die Spitze sowie die Beschränkung auf eindimensionale Linienprofile genannt werden.

Rastertunnelmikroskop: Das von Binnig und Rohrer 1986 entwickelte Rastertunnelmikroskop ("Scanning Tunneling Microscope", STM) ermöglichte zum ersten mal die atomar aufgelöste Abbildung von zweidimensional ausgedehnten Oberflächenbereichen. Das Messprinzip basiert auf dem Tunneleffekt: wird eine leitende Spitze (meist W) sehr nahe an eine leitende Oberfläche gebracht, so beginnt ein quantenmechanischer Tunnelstrom zu fließen. Dessen Stärke hängt exponentiell vom Abstand der Nadel von der Oberfläche ab. Wird die Nadel über die Oberfläche bewegt, so kann aus der Variation des Tunnelstromes die Gestalt der Oberfläche abgeleitet werden. Die vertikale Auflösung liegt dabei im pm-Bereich (10^{-12} m), die laterale Auflösung liegt im Å-Bereich. Die genaue laterale und auch vertikale Positionierbarkeit der Spitze wird heute, wie auch beim Prototypen des STM, mit piezokeramischen Positionierelementen erreicht.

Das STM bietet auch ein gutes Beispiel, um das heute in der Messtechnik weitverbreitete Rückkopplungs- ("Feedback") Prinzip, wie es in Abb. 4.19. schematisch dargestellt ist, zu erläutern.

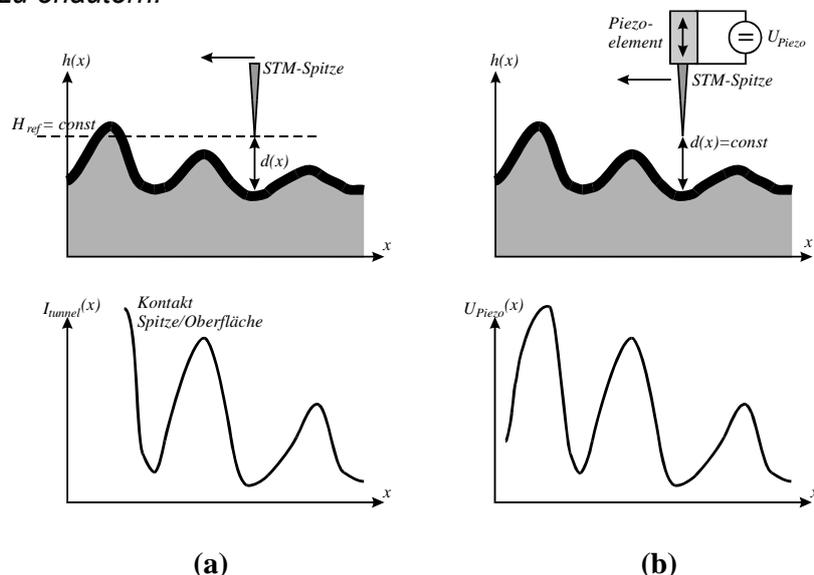


Abb. 4.19.: Schematische Darstellung des Rückkopplungsprinzips für das STM

- a** Starre Spitze, Messung des Tunnelstromes I_{tunnel} ; Es kann zum Kontakt von Spitze und Oberfläche kommen
- b** Konstanter Abstand Spitze/Oberfläche d ; Messung der Piezospannung U_{Piezo} , die erforderlich ist, um d konstant zu halten

Wie bereits erwähnt, führt eine Bewegung der STM-Spitze über eine Oberfläche in einer konstanten Höhe relativ zu einem beliebigen Referenzniveau, H_{ref} , zu einer Variation

des Tunnelstromes, I_{tunnel} (Abb. 4.19.a). Dieses Messprinzip ist allerdings aus mehreren Gründen ungünstig: steigt z. B. die Oberfläche über das Referenzniveau an, so kommt es zum (unerwünschten) Kontakt zwischen Spitze und Probe. Auch ist der zu messende Tunnelstrom sehr klein und daher das Signal-Rauschverhältnis ungünstig. Es ist daher günstiger, den Tunnelstrom, und damit den Abstand der Spitze von der Oberfläche, konstant zu halten und z. B. jene Spannung U_{Piezo} zu messen, die dem vertikalen Piezoelement zugeführt werden muss, um den Tunnelstrom konstant zu halten (Abb. 4.19.b). Ein Kontakt Spitze/Oberfläche wird dabei vermieden und das Signal wird deutlicher, da typische Piezospannungen in der Größenordnung von 1-1000 V liegen, was leicht messbar ist.

Die Vorteile des STM sind sein robuster Aufbau und seine hohe Abbildungsgenauigkeit. Nachteile sind die Beschränkung auf leitende Oberflächen sowie das Faktum, dass der Tunnelstrom nicht nur von der Entfernung Spitze/Oberfläche abhängt, sondern auch von der Spitzengeometrie, der chemischen Zusammensetzung der Spitze (welche nicht genau kontrollierbar ist) und von der chemischen Zusammensetzung der zu untersuchenden Oberfläche. Besonders der letzte Punkt kann aber auch ein Vorteil sein, da er eine Unterscheidung von einzelnen Atomsorten an der Oberfläche erlaubt.

Rasterkraftmikroskop: Das Rasterkraftmikroskop ("Atomic Force Microscope", AFM) ist die Erweiterung des STM auf nichtleitende Oberflächen. Ein mikromechanisch gefertigter Hebelarm ("Cantilever", meist aus Si oder Si_3N_4) trägt an seinem Ende eine Spitze mit einem Öffnungswinkel in der Größenordnung von $10-50^\circ$. Die Spitzenverrundung liegt bei ca. 20-50 nm. Wird die Spitze in Kontakt mit der Oberfläche gebracht, so verbiegt sich der Hebelarm. Die Verbiegung wird gemessen und, gemäss dem vorher beschriebenen Rückkopplungsprinzip, konstant gehalten. Aus dem Signal an das vertikale Stellelement wird auf die Oberflächengestalt zurückgerechnet.

Das heute häufigste Prinzip zur Messung der Verbiegung des Cantilevers ist die Detektion der Ablenkung eines an der Cantileverrückseite reflektierten Laserstrahles mittels einer vier-quadranten Photodiode, welche die Detektion von vertikalen und lateralen Cantilever-Deformationen erlaubt. Es kann sich allerdings auch eine leitfähige Spitze an der Rückseite des Hebelarmes befinden, welche die Durchbiegung in ein Tunnelstromsignal umsetzt. Diese Bauform ist vor allem bei UHV-AFM's gebräuchlich.

Ein Vorteil des AFM ist, wie bereits erwähnt, seine Unabhängigkeit von der elektrischen Leitfähigkeit der Probe. Weiters liefert der direkte Kontakt zwischen Spitze und Probe auch orts aufgelöste Informationen über die mechanischen Eigenschaften der Oberfläche wie z. B. den lokalen Reibungskoeffizienten oder die Härte. Leider bringt der direkte Kontakt zwischen Spitze und Probe auch Nachteile: atomare Auflösung kann aufgrund der hohen Spitzenverrundung nicht erreicht werden. Die Oberfläche wird mechanisch belastet und es besteht eine hohe Wahrscheinlichkeit der Spitzenverschmutzung. All diese Punkte können allerdings durch den Betrieb des AFM im sogenannten "non-contact-mode" zumindest abgemildert werden. Dabei wird der freie Cantilever in seiner Resonanzfrequenz zum schwingen gebracht. Wird nun die Spitze an die Oberfläche angenähert, kommt es zu einer Verringerung der Amplitude und zu einer Phasenverschiebung gegenüber der freien Schwingung durch kurzzeitigen Kontakt der Spitze mit der Oberfläche. Die Amplituden- oder Phaseänderung wird als Feedback-Signal konstant gehalten und kann wiederum zur Rekonstruktion der Oberflächengestalt herangezogen werden.

In letzter Zeit haben sich aus STM und AFM eine Unzahl von Sondentechniken für die Untersuchung verschiedenster Oberflächeneigenschaften in den verschiedensten Medien entwickelt. Hier seien stellvertretend noch die "Magnetic Force Microscopy" (MFM) zur lokal aufgelösten Untersuchung magnetischer Oberflächen und die optische Nahfeldmikroskopie, welche z. B. zur Einzelmolekülspektroskopie verwendet werden kann, erwähnt.

4.2.3.2. Optische Methoden und Streuung

Die Rauigkeitsbestimmung mittels optischer Methoden ist das komplementäre Verfahren zu den Sondentechniken. Hier liegt die Information über die Oberflächenmorphologie nicht im Realraum, sondern im reziproken Raum vor. Generell sind bei der Reflexion einer elektromagnetischen Welle (sei es optische oder Röntgenstrahlung) an einer Oberfläche zwei Komponenten vorhanden: der spekulare und der diffuse Anteil. Diese sind in Abb. 4.20. dargestellt. Bei der Reflexion an einer glatten Oberfläche ist der spekulare Anteil dominierend (Abb. 4.20.a), für rauhe Oberflächen dominiert der diffuse Anteil (Abb.4.20.b).

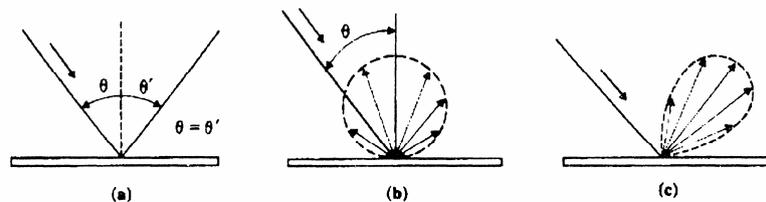


Abb. 4.20.: Reflexionstypen an Oberflächen [9, p. 35]:

a spekulare Reflexion

b diffuse Reflexion

c Kombination von spekulärer und diffuser Reflexion; reales Messsignal

Die Veränderung der Intensität des spekularen Anteils während des Schichtwachstums kann daher bereits qualitative Informationen über die Rauigkeitsentwicklung liefern. Quantitative Informationen über die Oberflächenmorphologie können aus dem diffusen Anteil extrahiert werden, und zwar in Form der Fourierkoeffizienten des Fourierpektrums der Oberfläche. Diese können aus der orts aufgelösten Intensität des diffusen Anteils ermittelt werden. Die exakte Gestalt der Oberfläche, welche (der Einfachheit halber für ein eindimensionales Profil) in ihrer Fourieranalyse gegeben ist durch

$$h(x) = \sum_{k=0}^N A_k \sin(kx + \varphi_k) \quad (4.16)$$

kann allerdings nicht rekonstruiert werden, da bei der Reflexion die Phaseninformation verloren geht. Es können zwar die A_k , nicht aber die φ_k ermittelt werden. Was dieser Sachverhalt bedeutet ist für ein einfaches Oberflächenprofil mit nur 3 Fourierkomponenten in Abb. 4.21. illustriert.

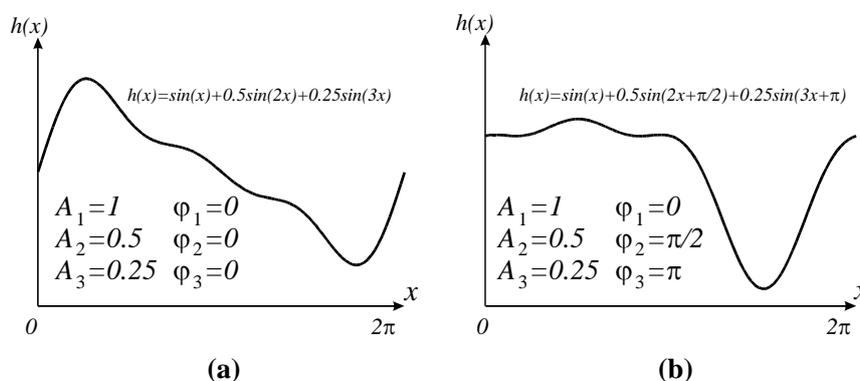


Abb. 4.21.: Oberflächenapproximation mit drei Fourierkomponenten im Intervall $[0, 2\pi]$

a alle Phasenfaktoren $\varphi_k=0$

b $\varphi_1=0, \varphi_2=\pi/2, \varphi_3=\pi$

Die Kenntnis der Faktoren A_k reicht nicht aus, um die Oberflächengestalt eindeutig zu rekonstruieren.

Als optische Strahlung kommen sowohl sichtbares Licht als auch Röntgenstrahlung zum Einsatz. Insbesondere die Röntgenstrahlung erlaubt durch ihre hohe Eindringtiefe auch die Bestimmung von Rauigkeiten an inneren Grenzflächen von beliebigen Materialien. Auch Teilchenstrahlung wie z. B. Elektronen oder Ionen kommen für Streuexperimente in Frage und liefern aufgrund des Welle/Teilchendualismus ähnliche Ergebnisse wie die Streuung elektromagnetischer Strahlung, allerdings aufgrund ihrer wesentlich geringeren Wellenlänge auf einer anderen Längenskala.

Die auf Streuung basierenden Techniken sind zerstörungsfrei und erlauben die Bestimmung von Oberflächenparametern in Echtzeit und auch in situ während des Beschichtungsprozesses. Nachteile sind der bereits erwähnte Verlust der Phaseninformation sowie, bei Röntgenstrahlung, die geringen Reflexionskoeffizienten bei schiefen Einfallswinkeln. Die geringe Intensität besonders der diffus reflektierten Röntgenstrahlung macht daher oft den Einsatz einer Synchrotronquelle notwendig, was mit hohem apparativen, administrativen und zeitlichem Aufwand verbunden ist.

4.2.4. Quantifizierung von Rauigkeiten

Wie bereits in Abschnitt 4.2.3., Rauigkeitsmessung, erwähnt, spielt für die Messung von Rauigkeiten im Realraum das Messintervall L und die laterale Auflösung (sampling interval), Δx , des Messgerätes eine wesentliche Rolle. Im Falle einer eindimensionalen Messstrecke ergibt sich die Anzahl der Messpunkte, N , aus $L/\Delta x$, bei Abrastern eines quadratischen Bereiches der Seitenlänge L ergibt sich N , bei gleicher lateraler Auflösung in x - und y , zu $N=(L/\Delta x)^2$. Daher kann L bei bekanntem Δx durch N ersetzt werden, was dazu führt, dass die im folgenden gegebenen Rauigkeitsdefinitionen für ein- und zweidimensionale Datenfelder gleichlautend sind.

4.2.4.1. Globale Messgrößen

Bevor ein Rauigkeitswert ermittelt werden kann, muss ein Referenzniveau in der aktiven Oberfläche definiert werden. Dieses ist gegeben durch den Mittelwert \bar{h} aller gemessenen Höhenwerte h_i ,

$$\bar{h} = \frac{1}{N} \sum_{i=1}^N h_i \quad (4.17)$$

Die Rauigkeit R ist dann durch die Abweichungen von diesem Referenzniveau definiert, wie Abb. 4.22. schematisch zeigt.

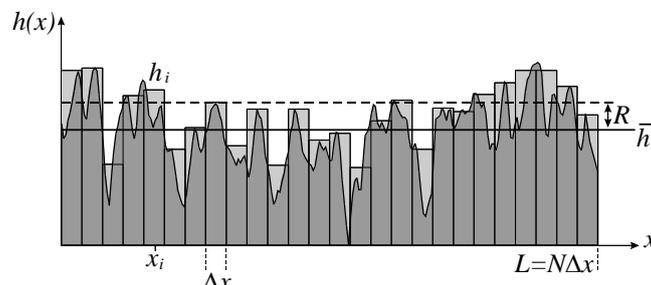


Abb. 4.22.: Schematische Darstellung der Rauigkeit R als Abweichung von der mittleren Höhe eines Oberflächenprofils \bar{h}

Die Rechenvorschrift zur Bestimmung von R kann allerdings unterschiedlich sein. Aus historischen Gründen haben sich mehrere Definitionen von R herausgebildet. In den Ingenieurwissenschaften wird hauptsächlich der sogenannte R_a -Wert verwendet, welcher durch den Mittelwert der absoluten Abweichungen von \bar{h} definiert ist:

$$R_a = \frac{1}{N} \sum_{i=1}^N |\bar{h} - h_i| \quad (4.18)$$

In der wissenschaftlichen Literatur wird hingegen hauptsächlich der mathematisch aussagekräftigere R_q , R_{RMS} oder RMS-Wert verwendet. Dieser ist die Wurzel aus der mittleren quadratische Abweichung von \bar{h} :

$$R_q = R_{RMS} = RMS = \sqrt{\frac{1}{N} \sum_{i=1}^N (\bar{h} - h_i)^2} \quad (4.19)$$

R_q hat eine enge Verbindung zu den sogenannten Korrelationsfunktionen und deren Fouriertransformierten, welche im wesentlichen dem diffusen Anteil bei den im Abschnitt 4.4.3.2. beschriebenen Streuexperimenten entsprechen. Daher ist er für die theoretische Beschreibung von Rauigkeiten im Rahmen der Streutheorie vorzuziehen. Generell unterscheiden sich R_a und R_q numerisch nicht stark voneinander. Für den einfache Fall eines sinusoidalen Profils gilt $R_a/R_q = 2^{3/2} \pi \cong 0.9$, für allgemeinere Profilformen ist $R_a/R_q \cong 0.8$.

Verschiedene Profilformen können dieselben R_q bzw. R_a -werte aufweisen, wie Abb. 4.23. anhand einiger einfacher Beispiele schematisch zeigt.

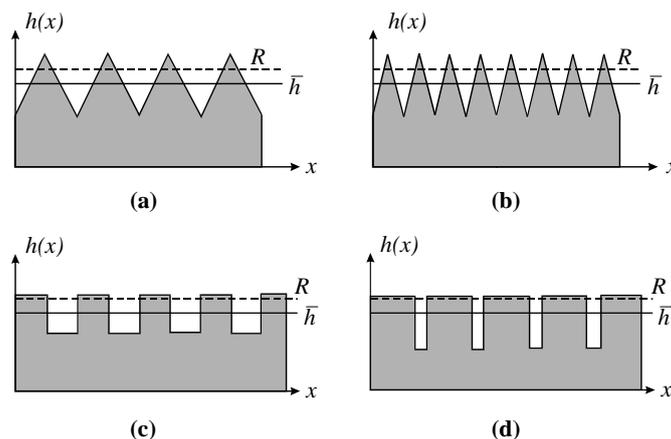


Abb. 4.23.: Oberflächen mit verschiedenen Gestalten, aber gleichen Rauigkeiten
a, b verschiedene Periodizität
c, d verschiedene Symmetrien

Parameter, welche zumindest im begrenzten Masse formspezifisch sind, sind die sogenannte "Skewness" Sk ,

$$Sk = \frac{1}{NR_q^3} \sum_{i=1}^N (h_i - \bar{h})^3, \quad (4.20)$$

deren Vorzeichen (aufgrund der ungeraden dritten Potenz in Gl. 4.20) ein Maß für die Symmetrie eines Profils um \bar{h} darstellt, d. h. eine Aussage darüber trifft, ob mehr Höhenwerte des Profils über oder unter \bar{h} liegen. Die sogenannte "Kurtosis" K

$$K = \frac{1}{NR_q^4} \sum_{i=1}^N (h_i - \bar{h})^4, \quad (4.21)$$

quantifiziert die mittlere Flankensteilheit eines Profils. Trotz der Definition dieser formsensitiven Parameter können weiterreichende Informationen, insbesondere über die laterale Gestalt eines Profils, nur mit mathematischen Werkzeugen gewonnen werden, die

auch den Abstand verschiedener Bereiche einer Oberfläche einbeziehen. Diese werden im folgenden beschrieben.

4.2.4.2. Korrelationsfunktionen

In vielen Gebieten der Physik spielen räumliche und zeitliche Korrelationen eine große Rolle. Unter Korrelation versteht man das vermehrte Auftreten bestimmter Merkmale (seien es Höhenwerte, die Anzahl von radioaktiven Zerfallsereignissen oder von bestimmten Signalpegeln bei einer Messung) in einem bestimmten Intervall um einen Messpunkt herum. Dieses Intervall wird als Korrelationslänge bezeichnet. Bei einem Oberflächenprofil kann dies z. B. bedeuten, dass um einen großen Höhenwert das Profil nicht beliebig steil abfallen kann wie es bereits bei der Diskussion der selbstähnlichen Oberflächen (Abschnitt 4.2.2.2.) erwähnt wurde. Im Gegensatz dazu stehen stochastische, unkorrelierte Profile (Abschnitt 4.2.2.1.), wo beim betrachten eines Höhenwertes h_i keinerlei Rückschluss auf den Höhenwert $h_{i\pm 1}$ gezogen werden kann.

Zur Quantifizierung von Korrelationen können mehrere Funktionstypen herangezogen werden. Die sogenannte Autokovarianzfunktion $R(X)$ ergibt sich für ein Intervall mit N Messpunkten mit dem Abstand Δx zu

$$R(X) = R(n \cdot \Delta x) = \frac{1}{N-n} \sum_{i=1}^{N-n} (h_i - \bar{h}) \cdot (h_{i+n} - \bar{h}). \quad (4.22)$$

Dabei ist X das n -fache des sampling intervals Δx wobei n Werte von 0 bis $N-1$ annehmen kann. Die Autokovarianzfunktion ist der Mittelwert der Produkte zweier Höhenwerte h_i und h_{i+n} , welche die Strecke X voneinander entfernt sind. Naturgemäss können mit steigendem X bei endlicher Profillänge nur immer weniger Wertepaare zur Mittelwertbildung herangezogen werden, wie Abb. 4.24. schematisch zeigt.

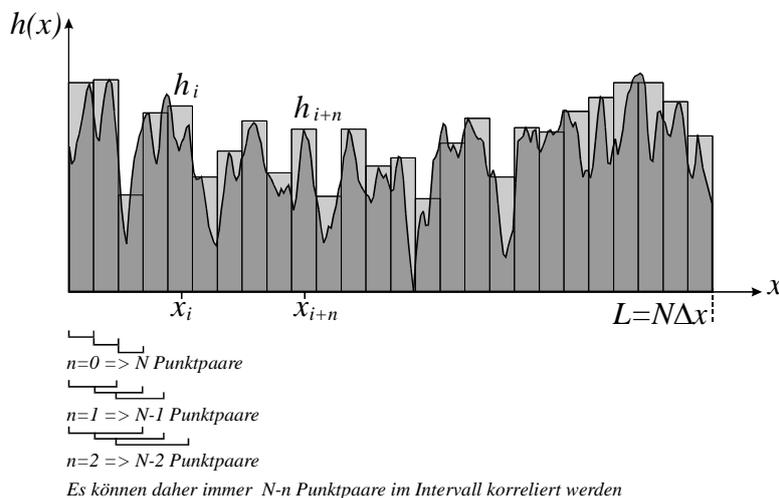


Abb. 4.24.: Bildung der Autokovarianzfunktion für ein diskretes Profil. Je weiter die zu korrelierenden Punkte entfernt sind, desto weniger Punktpaare können zur Mittelung herangezogen werden

Aus Abb. 4.24. ergibt sich die in Gleichung (4.22) gegebene Struktur von $R(X)$. Für ein kontinuierliches Profil geht $R(X)$ in die in Gleichung (4.23) gegebene Integralform

$$R(\tau) = \frac{1}{L-\tau} \int_0^{L-\tau} (h(x) - \bar{h}) \cdot (h(x+\tau) - \bar{h}) dx \quad (4.23)$$

über, wobei L die Länge des Profils ist und τ die kontinuierliche Entfernung der Punktepaare. Aus den Gleichungen (4.22) bzw. (4.23) ist leicht zu sehen, dass $R(0)$ dem RMS-Wert entspricht. Dieser Zusammenhang von R_q mit der Autokovarianzfunktion erklärt die höhere mathematische Signifikanz der RMS-Rauhigkeit. Oft wird auch die normierte Autokovarianzfunktion oder Autokorrelationsfunktion

$$\rho(X) = R(X) / R(0) \text{ oder } \rho(\tau) = R(\tau) / R(0) \quad (4.24)$$

verwendet. Für Profile, welche Korrelationen aufweisen, jedoch keine Periodizitäten, sind die vorher definierten Funktionen monoton fallend. Die Korrelationslänge ξ ist dabei durch die Geschwindigkeit des Abfallens vom Wert bei X (oder τ) = 0 gegeben. Dieses Verhalten ist in Abb. 4.25. schematisch dargestellt.

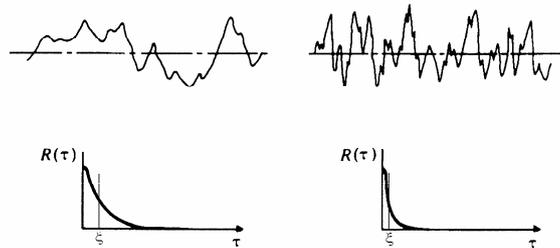


Abb. 4.25.: Autokovarianzfunktion für Oberflächenprofile gleicher Rauigkeit, aber mit verschiedener Korrelationslänge ξ [9, p.151]

Oft wird ξ als jener Wert X oder τ definiert bei dem $R(X)$ (oder $R(\tau)$, $\rho(X)$, $\rho(\tau)$) auf das 1/e-fache ihres Wertes bei X , $\tau = 0$ abgefallen sind. Das Auftreten von Periodizitäten äußert sich hingegen im Auftreten von Maxima bei Werten von X , $\tau \neq 0$.

Eine zu $R(X)$ analoge Funktion ist die sogenannte Strukturfunktion $S(X)$,

$$S(X) = S(n \cdot \Delta x) = \frac{1}{N-n} \sum_{i=1}^{N-n} [(h_i - \bar{h}) - (h_{i+n} - \bar{h})]^2. \quad (4.25)$$

in der diskreten Darstellung und

$$S(\tau) = \frac{1}{L-\tau} \int_0^{L-\tau} [(h(x) - \bar{h}) - (h(x+\tau) - \bar{h})]^2 dx \quad (4.26)$$

in der Integraldarstellung für kontinuierliche Profile. $R(\tau)$ und $S(\tau)$ sind über die Beziehung

$$S(\tau) = 2R_q^2 [1 - \rho(\tau)] \quad (4.27)$$

miteinander verknüpft. Obwohl beide Funktionen dadurch die gleiche Information enthalten, ist $S(X)$ (oder $S(\tau)$) leichter zu berechnen und mathematisch robuster. Während z. B. hochfrequente Anteile $R(\tau)$ bei kleinen τ stark beeinflussen, ist $S(\tau)$ davon nicht betroffen, da es für kleine τ gegen 0 geht. Weiters ist $S(\tau)$ wesentlich stärker sensitiv für die Oberflächenrauigkeit, da es für große τ gegen $2R_q^2$ strebt. Dies macht Abb. 4.26. deutlich, welches einen Vergleich von $R(\tau)$ und $S(\tau)$ für unverschlossene ("unworn") und verschlossene ("worn") Oberflächen zeigt.

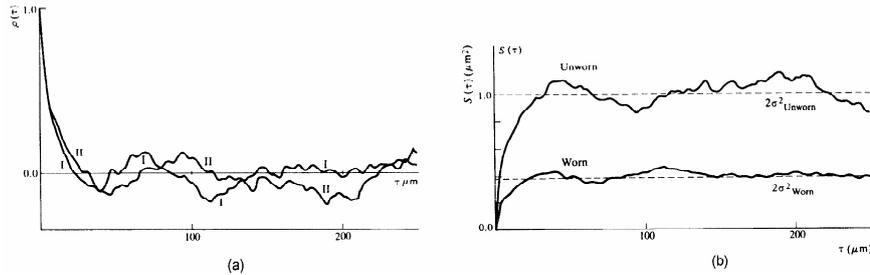


Abb. 4.26.: Korellationsfunktionen für Oberflächen verschiedener Rauigkeiten
a Autokorrelationsfunktion: keine signifikanten Unterschiede feststellbar
b Strukturfunktion: Unterschiede in der RMS-Rauigkeit (hier als σ bezeichnet) treten deutlich zu tage. [9, p., 156]

Die Verminderung der Rauigkeit durch den Oberflächenverschleiss ist bei $S(\tau)$ deutlich erkennbar, während $R(\tau)$ praktisch keine Unterscheidung der Oberflächentypen ermöglicht.

Für zweidimensionale Oberflächenscans besteht die Möglichkeit, diese in M eindimensionale Profile zu zerlegen und deren Korrelationsfunktionen nochmals zu mitteln. Die eindimensionalen Profile können dabei parallel oder normal zur Scanrichtung liegen oder z. B. auch aus radialen Linien eines kreisförmigen Bereiches auf der Oberfläche gebildet werden. Im letzteren Fall spricht man von "radial gemittelten Korrelationsfunktionen".

4.2.4.3. Methoden auf der Basis von Fourieranalysen

Wie in Abschn. 4.3.2.2., Optische Methoden und Streuung, erwähnt, liefert die Streuung elektromagnetischer Strahlung im wesentlichen die Fouriertransformierte des Oberflächenprofils. Vom mathematischen Standpunkt aus ist das Ergebnis des Streuexperimentes die sogenannte "Power Spectral Density", $P(k)$. $k=2\pi/\lambda$ ist dabei ein Wellenvektor, welcher die Wellenlänge λ einer isolierten Fourierkomponente des Oberflächenprofils repräsentiert. $P(k)$ ist über die Beziehung

$$P(k) = \lim_{L \rightarrow \infty} \frac{1}{L \cdot (2\pi)^2} \left| \int_{-\infty}^{\infty} h(r) e^{ikr} dr \right|^2 \quad (4.28)$$

mit dem Oberflächenprofil verknüpft. Außerdem ist $P(k)$ auch die Fouriertransformierte der Autokovarianzfunktion,

$$P(k) = \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} R(r) e^{ikr} dr. \quad (4.29)$$

Somit können aus dem Streuexperiment wesentliche globale Parameter der Oberfläche wie z. B. Korrelationslängen und Rauigkeiten durch die Rücktransformation von $P(k)$ in den Ortsraum gewonnen werden. Die detaillierte Gestalt der Oberfläche kann jedoch, wie bereits in Abschn. 4.3.2.2. besprochen, aufgrund des Verlustes der Phaseninformation, aus dem Streuexperiment nicht rekonstruiert werden.

Die in den vorhergehenden Abschnitten beschriebenen Mechanismen der Rauigkeitsentwicklung, Messmethoden und mathematischen Werkzeuge zur Oberflächencharakterisierung sind naturgemäß nicht vollständig. Die Strukturentwicklung von Schichten während des Wachstums ist noch von zahlreichen anderen Prozessen abhängig, welche in den sogenannten Strukturzonenmodellen enthalten sind. Diese werden im Folgenden behandelt und mit verschiedenen Schichteigenschaften korreliert.

4.3. Mechanische Eigenschaften

4.3.1. Einleitung

Für die mechanischen Eigenschaften von dünnen Schichten ist weitgehend deren Struktur verantwortlich. Es soll daher vorerst der Zusammenhang zwischen Schichtwachstum, Struktur (Morphologie) und den daraus folgenden mechanischen Eigenschaften aufgezeigt werden

4.3.2. Strukturzonen-Modelle

4.3.2.1. Allgemeines

Für das Wachstum der Schicht und die Ausbildung ihrer Struktur sind drei Faktoren maßgebend: die Rauigkeit technischer Oberflächen, die Aktivierungsenergien für die Oberflächen- und die Volumsdiffusion der Schichtatome und schließlich die Adatom/Substrat-Bindungsenergie. Die Rauigkeit des Substrates bewirkt Abschattungen der im allgemeinen aus einer Vorzugsrichtung einfallenden Atome und damit eine ungleichmäßige Belegung, so dass eine poröse Struktur entsteht. Dieser Abschattungseffekt kann jedoch durch Oberflächendiffusion ausgeglichen werden.

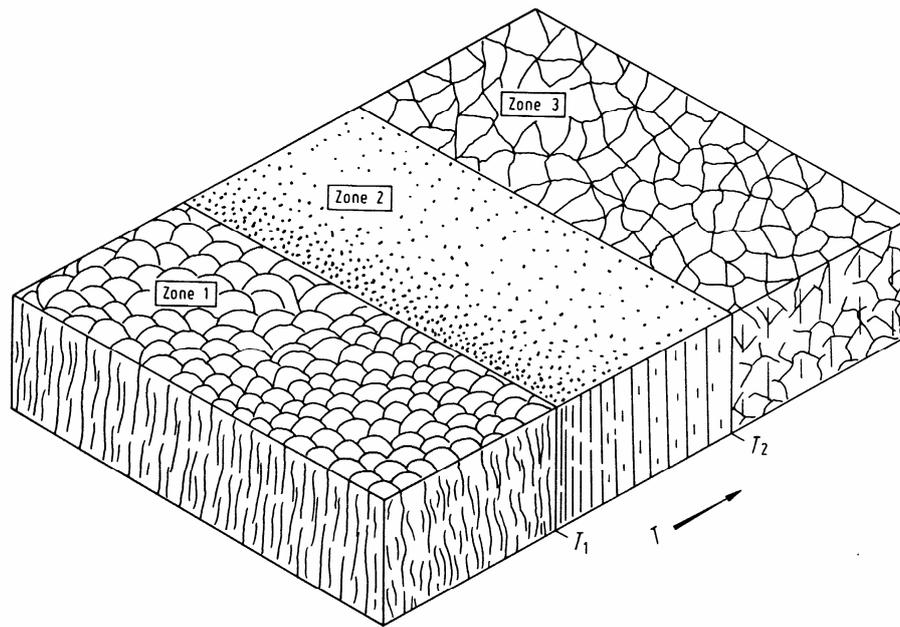
Die oben genannten Energien sind für viele reine Metalle ihrer absoluten Schmelztemperatur T_m [K] proportional. Daher ist zu vermuten, dass von den drei Effekten: Abschattung, Oberflächendiffusion und Volumsdiffusion - jeweils einer in einem bestimmten Bereich von T/T_m , d.h. der auf T_m [K] bezogenen Substrattemperatur T [K] dominiert und den Typ der entstehenden Mikrostruktur prägt. Dies ist die Grundlage für die beiden folgenden Strukturzonen-Modelle.

4.3.2.2. Modell nach Movchan und Demchishin

Movchan und Demchishin untersuchten 1969 die Struktur und Eigenschaften von im Hochvakuum (bei 10^{-4} bis 10^{-3} Pa) aufgedampften, vergleichsweise dicken (bis zu 2mm!) Schichten aus Ti, Ni, W, ZrO_2 und Al_2O_3 als Funktion von T/T_m . Sie beschrieben ihre Ergebnisse durch das Dreizonen-Modell (M.-D.-Modell), welches in Abb. 4.27.a zu sehen ist.

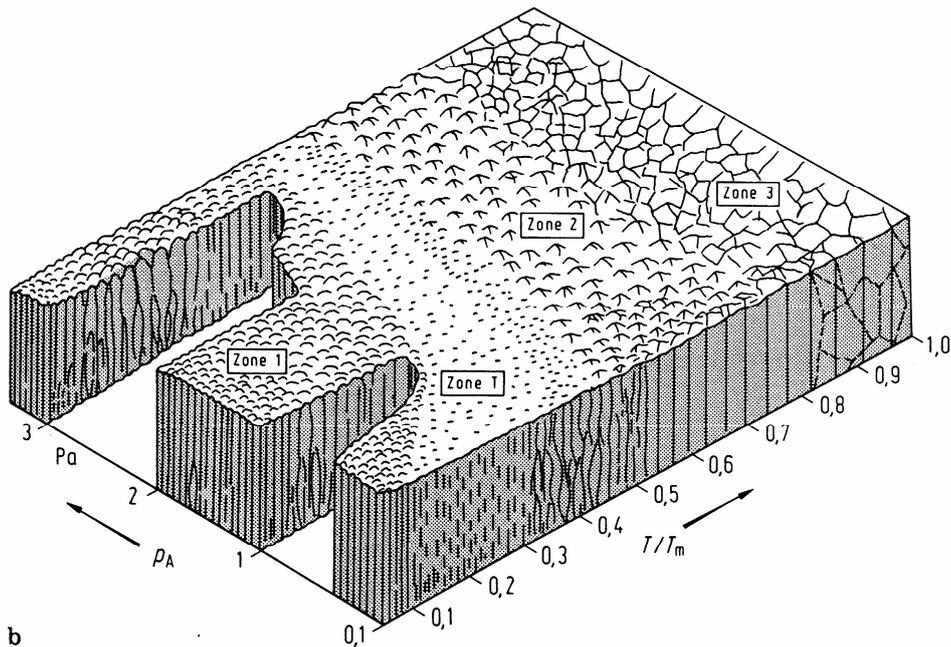
4.3.2.3. Thornton-Modell

Das M.-D.-Modell wurde von Thornton 1974 aufgrund von Experimenten mit einer Hohlkathoden-Sputteranordnung bei Drücken im Bereich 0,1 bis 4Pa Argon erweitert. Thornton fügte erstens als weitere Variable den Argon-Druck hinzu, um den Einfluß einer Gasatmosphäre (ohne Ionenbombardement!) auf die Struktur zu beschreiben. Zum anderen wurde eine Übergangs (Transition)-Zone T zwischen den Zonen 1 und 2 eingefügt (siehe Abb. 4.27.b). Diese Übergangszone ist bei Schichten aus Metallen und einphasigen Legierungen nicht deutlich ausgeprägt, wohl aber bei Schichten aus refraktären Verbindungen und mehrphasigen Legierungen, die durch Aufdampfen im Hochvakuum oder in Gegenwart von inerten oder reaktiven Gasen durch Sputtern oder Ionenplattieren hergestellt werden. Im übrigen stimmen die Zonen beider Modelle in ihren Merkmalen überein.



a

	Zone 1	Zone 2	Zone 3
Metalle T/T_m	$< 0,3$	$0,3 \dots 0,45$	$> 0,45$
Oxide T/T_m	$< 0,26$	$0,26 \dots 0,45$	$> 0,45$



b

Abb. 4.27.: Strukturzonen-Modelle: **a** nach Movchan und Demchishin; **b** nach Thornton;
Zone 1: poröse Struktur aus nadelförmigen Kristalliten
Zone T: dichtes faserförmiges Gefüge
Zone 2: kolumnares Gefüge
Zone 3: rekristallisiertes Gefüge
 T Substrattemperatur [K]; T_m Schmelztemperatur [K]; p_A Argon-Druck
 [1, p. 22]

Zone 1

Die Zone 1 umfasst die sich bei niedrigem T/T_m bildende Struktur. Die Adatom-Diffusion reicht nicht aus, um die Wirkung der Abschattung auszugleichen. Daher entstehen aus einer relativ geringen Zahl von Keimen nadelförmige Kristallite, die mit zunehmender Höhe durch Einfangen von Schichtatomen breiter werden und sich zu auf der Spitze stehenden Kegeln mit gewölbten Basisflächen entwickeln. Die Schicht ist porös, und die einzelnen Kristallite haben bei einer gegenseitigen Distanz von einigen 10nm eine hohe Versetzungsdichte und hohe innere Spannungen (Druck, wegen hoher Defektdichte). Generell stehen Schichten der Zone 1 allerdings unter Zugspannungen, da es mit steigender Schichtdicke zur Koaleszenz benachbarter Kristallite kommen kann.

Zone T

Die Zone T ist dadurch gekennzeichnet, dass die Adatome durch Oberflächendiffusion die Wirkung der Abschattungen zum Teil ausgleichen. Weiters ist, vor allem bei geringen Arbeitsgasdrücken, ein permanenter Strom energiereicher Teilchen vorhanden, welche durch die Erzeugung von Oberflächendefekten die Keimdichte stark erhöhen. Es entsteht dadurch eine faserförmige und gegenüber der Zone 1 dichtere Struktur.

Zone 2

Die Zone 2 ist durch den Bereich T/T_m definiert, in dem die Oberflächendiffusion für das Wachstum bestimmend ist. Es bildet sich eine kolumnare Struktur, wobei der Säulendurchmesser mit der Substrattemperatur T wächst und die Porosität entsprechend abnimmt.

Zone 3

Die Zone 3 umfaßt schließlich den T/T_m -Bereich, in dem das Wachstum durch die Volumendiffusion bestimmt wird. Es entsteht ein rekristallisiertes dichtes Gefüge aus dreidimensionalen Kristalliten. Dieser Temperaturbereich ist auch für das epitaktische Wachstum von Halbleitern durch Aufdampfen, Sputtern und CVD von Bedeutung.

Einfluß des Inertgasdruckes auf die Struktur

Nach dem Thornton-Modell sinken die Übergangstemperaturen T_1 und T_2 mit fallendem Inertgasdruck p_A . Das liegt im wesentlichen daran, dass, wie bereits erwähnt, ein permanenter Strom energiereicher Teilchen vorhanden ist. Grund dafür ist die geringe Anzahl von Kollisionen der Beschichtungsteilchen mit den Arbeitsgasatomen. Dadurch werden einerseits Oberflächendefekte erzeugt, welche die Keimdichte stark erhöhen. Andererseits übertragen sie ihren Impuls an locker adsorbierte Teilchen (z. B. Adatome) und erhöhen so deren transiente Mobilität. Weiters kommt es naturgemäß durch den Energieeintrag zu einer zusätzlichen Heizung des Substrates. All diese Effekte bewirken eine T-Reduktion besonders beim Übergang von Zone 1 zu Zone T.

Einfluß des Ionenbombardements auf die Struktur

Ein Ionenbombardement erzeugt auf dem Substrat Punktdefekte und erhöht damit die Keimdichte; zum anderen wird durch Energieübertragung auf die Adatome deren Beweglichkeit erhöht. Daher entsteht bei gegebenen T/T_m gegenüber dem Fall ohne Ionenbeschuß eine Struktur aus Kristalliten, die dichter gepackt sind.

Mit anderen Worten: das Ionenbombardement beeinflusst die Schichtstruktur in dem Sinne, dass sich die Zonengrenzen, insbesondere zwischen Zone 1 und Zone T zu tieferen T/T_m -Werten hin verschieben. Dieser Effekt, der durch das zusätzliche Aufheizen des Substrates allein nicht erklärt werden kann, wird in der Tat beobachtet (siehe Abb. 4.28.).

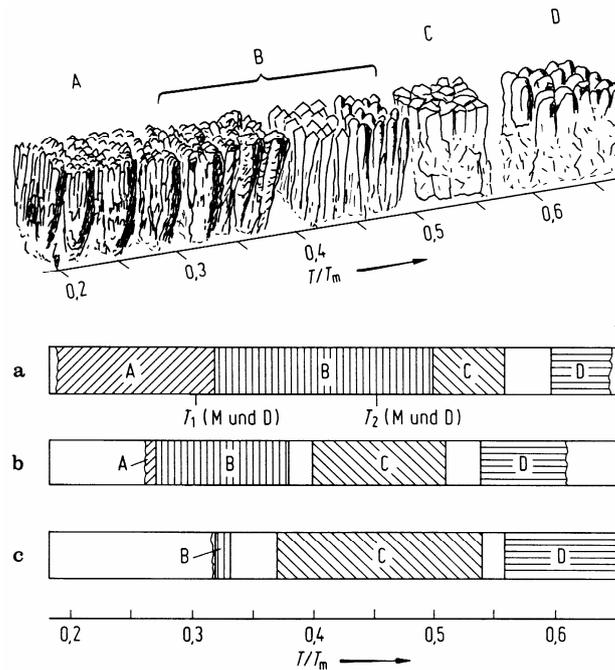


Abb. 4.28.: Einfluß der Substrattemperatur T und der Biasspannung U_B auf die Struktur ionenplattierter Titan-Schichten bei $p_A = 2,7\text{Pa}$
a $U_B = 0\text{V}$; **b** $U_B = -5\text{kV}$; **c** $U_B = -10\text{kV}$ [1, p. 24]

Die Ti-Schichten in Abb. 4.28 wurden durch Verdampfen aus einer Elektronenkanone hergestellt. Das Substrat lag auf Biasspannungen bis zu -10kV und 2mAcm^{-2} Stromdichte (das entspricht ca. 1 Ion pro Beschichtungsteilchen). Man sieht, dass bei einer Substrattemperatur, bei der durch reines Aufdampfen ($U_B = 0$) poröse Schichten mit Nadelstruktur entstehen, durch Ionenplattieren Schichten mit einem dichten Gefüge erhalten werden können.

Allerdings besteht ein wichtiger Unterschied zwischen dem dichten Gefüge der Aufdampfschichten und dem der Ionenplattierschichten: Die durch Aufdampfen bei entsprechend hoher Temperatur entstandene Struktur ist das Ergebnis von Rekristallisation und Kornwachstum durch Volumendiffusion. Beim Ionenplattieren spielt die Volumendiffusion wegen der vergleichsweise tiefen Substrattemperaturen eine wesentlich geringere Rolle. Hingegen werden bei intensivem Ionenbombardement kontinuierlich neue Keime gebildet, so dass ein feinkörniges, dichtes Gefüge entsteht.

4.3.3. Inkorporation von Fremdatomen

In die wachsende Schicht inkorporierte Fremdatome können mit den Schichtatomen Verbindungen eingehen, interstitielle oder substitutionelle Positionen einnehmen und dadurch die Schichteigenschaften, z.B. die inneren Spannungen, verändern. Verunreinigungen können auf Korngrenzen zur Versprödung und im Korninneren zu polymorphen Phasen führen. Fremdstoffe können auch das kolumnare Wachsen der Schicht durch kontinuierliche Renukleation unterbinden, so dass ein feinkörniges Gefüge entsteht.

Ein Beispiel für die ständige Renukleation ist die Wirkung der "Glanzstoffe", die galvanischen Bädern zur Erzielung glänzender Schichten zugesetzt werden. Üblicherweise haben galvanisch abgeschiedene Schichten eine dendritische Struktur, die aufgrund der großen Beweglichkeit der hydratisierten Ionen und der bevorzugten Kondensation auf gewissen Kristallflächen - in weitgehender Analogie zu den PVD-Schichten - entsteht. Durch Zugabe der Glanzstoffe, d.h. bestimmter organischer Additive zum Elektrolyten, bildet sich ein sehr feinkörniges Gefüge, das den gewünschten Glanz bewirkt.

Bei allen Beschichtungsverfahren werden Fremdatome in die Schicht eingebaut - sei es als nicht erwünschte, aber unvermeidliche Verunreinigung oder als gezieltes Mittel, um besondere Schichteigenschaften zu erreichen. So liegen nicht erwünschte Verunreinigungen z.B. vor bei galvanisch erzeugten Schichten, die Wasser und mitabgeschiedene organische oder anorganische Additive enthalten können. Bei der stromlosen (chemischen) Abscheidung werden Katalysatorsubstanzen (Phosphor, Bor) in die Schicht eingebaut.

Bei den Vakuumprozessen sind im Restgas H_2O , O_2 , N_2 , H_2 und Kohlenwasserstoffe vorhanden, die durch Gasabgabe der Materialien, Einströmungen durch Lecks und Rückströmungen von Pumpentreibmitteln in die Apparatur gelangen und z. B. beim Aufdampfen in die Schicht aufgenommen werden können. Bei den Sputterverfahren, insbesondere beim Bias-Sputtern, kommt die Implantation von Sputtergas hinzu.

Bei den CVD-Verfahren kann Trägergas und nicht vollständig umgesetztes Reaktionsgas inkorporiert werden. Da selbst geringe Mengen an Verunreinigungen sich stark auf die Schichteigenschaften auswirken können, ist der Aufwand zur Erzielung einer sauberen Apparatur oft enorm, aber gerechtfertigt.

Die Inkorporation von Fremdstoffen in die Schicht wird aber als gezieltes Mittel bei den folgenden Verfahren eingesetzt:

- ... Die reaktiven PVD-Verfahren erlauben, in Gegenwart reaktiver Gase ein elementares Beschichtungsmaterial in Form einer Verbindung (Nitrid, Karbid, Oxid, Borid) niederzuschlagen.
- ... Bei der CVD-Abscheidung von Chrom oder Wolfram auf Stahl und Hartmetallen diffundiert der Kohlenstoff des Substrates in die Schicht und bildet Karbide, was bei Herstellung von Hartstoffschichten von großer Bedeutung ist.
- ... Die Ionenimplantation von Fremdatomen findet in der Halbleitertechnologie und im zunehmendem Maße auch zur Herstellung abriebfester tribologischer Schichten Verwendung.

4.3.4. Innere Spannungen

4.3.4.1. Allgemeines

Alle Schichten befinden sich im Zustand mehr oder weniger großer innerer mechanischer Spannungen (auch Eigenspannungen genannt), die sich aus zwei Anteilen, dem thermischen σ_T und dem intrinsischen σ_i zusammensetzen:

$$\sigma = \sigma_T + \sigma_i \quad (4.30)$$

Die thermische Spannung σ_T hat ihre Ursache in den unterschiedlichen thermischen Ausdehnungskoeffizienten von Schicht und Substrat. Es gilt:

$$\sigma_T = E_S (\alpha_S - \alpha_U) (T_B - T_M) \quad (4.31)$$

Dabei ist

E_S = Elastizitätsmodul der Schicht

α_S = mittlerer thermischer Ausdehnungskoeffizient der Schicht

α_U = mittlerer thermischer Ausdehnungskoeffizient der Schichtunterlage

T_B = Substrattemperatur während der Beschichtung

T_M = Substrattemperatur während der Messung

Eine Berechnung der thermischen Spannung σ_T ist also - im Gegensatz zu den intrinsischen - prinzipiell möglich.

Die intrinsische Spannung σ_i hängt von den Beschichtungsparametern ab und rührt von struktureller Unordnung her, d.h. von inkorporierten Fremdatomen und von Schichtatomen, die außerhalb des Potentialminimums liegen. Diese Spannungen äußern sich, je nach Depositionsbedingungen, entweder als Druck- oder als Zugspannung, so dass sich der Film parallel zur Oberfläche auszudehnen bzw. zu kontrahieren sucht.

Werden niedrig schmelzende Materialien bei hinreichend hoher Substrattemperatur ($T/T_m > 0,5$) aufgebracht, so bleiben die intrinsischen Spannungen wegen der Begünstigung der Platzwechselfvorgänge relativ gering, während die thermischen Spannungen dominieren. Ein Ausgleich der letzteren kann erfolgen, wenn die Schichten getempert werden. Das dabei auftretende Fließen des Materials kann zur Bildung von Hügeln ("Hillocks") oder Löchern (μm - bzw. sub- μm -Bereich) in der Schicht führen, je nachdem, ob diese durch die Temperaturänderung unter Druck oder Zug gesetzt wird. Die Ausbildung solcher Hillocks wird z.B. bei Al-, Pb- und Au-Schichten beobachtet, die durch PVD-Methoden bei erhöhter Substrattemperatur aufgebracht werden.

Hochschmelzende Materialien werden im allgemeinen bei relativ geringen T/T_m -Werten ($< 0,25$) abgeschieden, so dass die intrinsischen Spannungen gegenüber den thermischen dominieren. Bei hinreichend dünnen Schichten ($< 500\text{nm}$) sind die intrinsischen Spannungen über die gesamte Schichtdicke annähernd konstant. Sie treten bei Aufdampfschichten üblicherweise als Zug- bei Sputterschichten als Druckspannungen in Erscheinung. Die Spannungen können der Zerreifestigkeit des kompakten Schichtmaterials nahe kommen und bei refraktären Metallen Werte von 10^3Nmm^{-2} erreichen. Oft übersteigt die intrinsische Spannung auch die Zerreifestigkeit des kompakten Schichtmaterials, was auf gewisse Verfestigungsmechanismen in der Schicht hinweist.

Die Bindungskräfte in der Interfacezone zwischen Substrat und Schicht müssen den Scherkräften widerstehen, die durch die intrinsischen und thermischen Spannungen verursacht werden. Da der Beitrag der intrinsischen Spannung zur Scherkraft mit der Schichtdicke wächst, kann die Schicht beim Überschreiten einer kritischen Dicke von der Unterlage abplatzen. Unter ungünstigen Bedingungen kann dies bereits bei 100nm Dicke geschehen.

Die intrinsischen Spannungen können hinsichtlich Art (Druck oder Zug) und Betrag durch die Prozessparameter wie Substrattemperatur, Depositionsrate, Einfallswinkel und Energieverteilung der Schichtatome, Gasinkorporation und Restgasatmosphäre beeinflusst werden (siehe Abb. 4.29.). Dies wurde für Aufdampfschichten und Sputterschichten in umfangreichen Arbeiten nachgewiesen. Damit ergibt sich die Möglichkeit, weitgehend spannungsfreie und damit auch relativ dicke Schichten durch PVD-Methoden herzustellen.

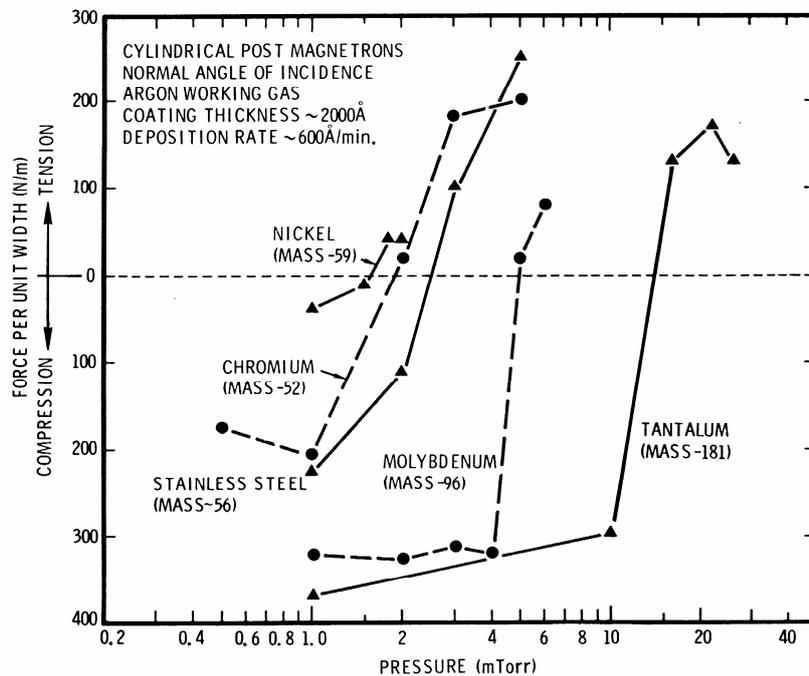


Abb. 4.29.: Untersuchung der inneren Spannungen von gesputterten Schichten in Abhängigkeit vom Ar-Druck [2, p. 222]

4.3.4.2. Messung der Inneren Spannungen

Allgemeines

Alle Meßmethoden beruhen im Prinzip darauf, dass ein relativ dünnes Substrat (Quarz oder Glas) beschichtet und die durch die mechanische Spannung der Schicht bedingte Durchbiegung in situ oder auch zu einem anderen Zeitpunkt oder bei einer anderen Temperatur gemessen wird. Die Temperaturmessung während der Beschichtung erfolgt meist auf einem Zusatzsubstrat unmittelbar neben dem eigentlichen Substrat. Die Meßsubstrate sind i.a. rechteckig; sie werden entweder an einem Ende oder an beiden Enden eingespannt. Im ersten Fall wird die Durchbiegung des freien Endes, im anderen die Durchbiegung in der Mitte der Probe gemessen. Aber auch kreisförmige Substrate finden durchaus Anwendung.

Die Trennung der beiden Anteile σ_T und σ_i kann folgendermaßen geschehen: Die Spannung wird bei Raumtemperatur und Beschichtungstemperatur gemessen. Bei Beschichtungstemperatur ist $T_B = T_M$, $\sigma_T = 0$ und nach Gleichung (4.30) ist $\sigma = \sigma_i$. Für Raumtemperatur ist dann $\sigma_T = \sigma - \sigma_i$, wenn σ die gemessene Spannung bei Raumtemperatur ist. Die Verformung wird entweder durch optische oder elektrische Methoden gemessen.

Interferenzoptische Messung

Das beschichtete und durchgebogene Substrat wird mit der Schichtseite auf eine ebene Glasplatte gelegt. Bei Zugspannung liegt es an den Enden, bei Druckspannung in der Substratmitte auf (siehe Abb. 4.30.).

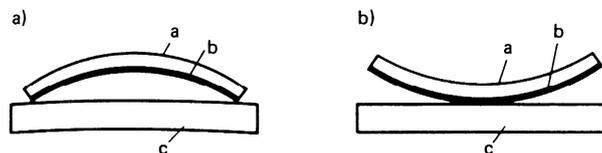


Abb. 4.30.: Substratverformung: **a)** Zugspannung in der Schicht
b) Druckspannung in der Schicht
a Substrat; **b** Schicht; **c** Referenzplatte
[3, p. 295]

Mit der Anordnung, die in Abbildung 4.31. zu sehen ist, kann mit ausreichender Genauigkeit der Krümmungsradius R_S des Substrates ermittelt werden.

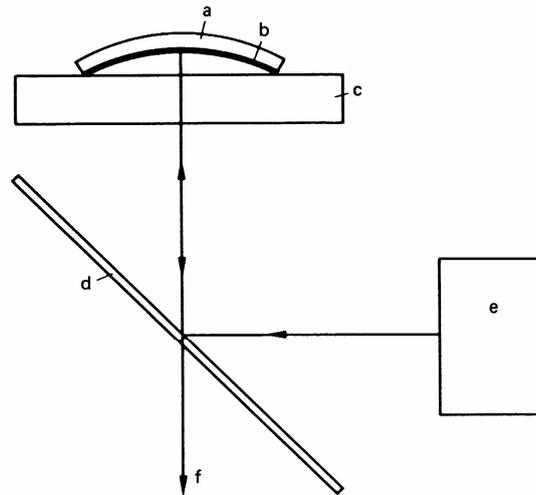


Abb. 4.31.: Schema zur Messung der Schichtspannungen durch Interferenz:
a Substrat, **b** Schicht; **c** plane Glasplatte; **d** Strahlenteiler; **e** Lichtquelle;
f zur Aufnahmeoptik [3, p. 295]

Für obige Anordnung gilt für den Krümmungsradius der Probe

$$R_s = \frac{D_m^2 - D_n^2}{4\lambda(m-n)} \quad (4.32)$$

mit

D_m = Durchmesser des m-ten Newton Ringes

D_n = Durchmesser des n-ten Ringes

λ = Wellenlänge des eingestrahlt Lichtes

Die Gesamtspannung einer dünnen Schicht berechnet sich nach:

$$\sigma = \frac{E_s d_s^2}{6(1-\nu_s)d_F} \left(\frac{1}{R_{s1}} - \frac{1}{R_{s2}} \right) \quad (4.33)$$

In der Gleichung bedeuten

E_s = Elastizitätsmodul des Substrates

ν_s = Poissonsche Konstante (Querkontraktionszahl) des Substratmaterials

d_s = Dicke des Substrates

d_F = Dicke der Schicht

R_{s1} bzw. R_{s2} = Krümmungsradius des Substrates vor bzw. nach den Beschichten

Messung durch geometrische Optik

Während die Anwendung der interferometrischen Methode eher zeitaufwendig ist, nimmt die Messung von Schichtspannungen mit Hilfe der geometrischen Optik weit weniger Zeit in Anspruch.

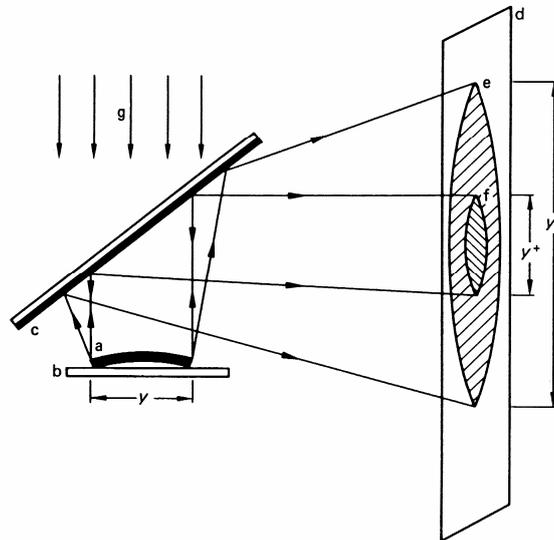


Abb. 4.32.: Schema einer Anordnung zur optischen Messung von Schichtspannungen:
a beschichtetes Substrat; **b** Glasplatte mit dünner Silberschicht;
c teildurchlässiger Silberspiegel; **d** Bildschirm; **e** Bild des beschichteten Substrates; **f** Bild des unbeschichteten Substrates; **g** einfallendes Licht;
y Durchmesser des Substrates; **y⁺** Durchmesser des Bildes des nicht beschichteten Substrates (für paralleles Licht ist $y=y^+$); **y[']** Durchmesser des Bildes des beschichteten Substrates
[3, p. 296]

Nach der in Abbildung 4.32. dargestellten Anordnung ist bei Proben mit kleinen Abmessungen die Auflösung allerdings gering. Näherungsweise gilt:

$$R = \frac{2yD}{y' - y^+} \quad (4.34)$$

Darin bedeutet

y = Durchmesser der Probe

y^+ = Durchmesser des Bildes der nicht beschichteten Probe

y' = Durchmesser des Bildes der beschichteten Probe

D = Strecke zwischen Probe und Bildschirm

Die Schichtspannung kann nach Gleichung (4.34) berechnet werden, wobei auch Asymmetrien in der Spannungsverteilung leicht erkenn- und messbar sind.

Elektrische Messung

Bei dieser Meßmethode wird das Substrat auf der der Beschichtungsquelle abgewandten Seite vor dem Aufbringen der zu prüfenden Schicht mit einer elektrisch leitenden Schicht (z.B. Silber) belegt. Das Substrat und eine in einem Abstand von ca. 1mm gegenüberliegende metallische Platte bilden einen Kondensator, der wiederum Teil eines hochfrequenten Schwingkreises ist. Wenn sich nun das Substrat aufgrund von inneren Spannungen durchbiegt, ändert sich der Abstand dieser beiden Kondensatorplatten, was eine Verstimmung des Schwingkreises zur Folge hat. Diese Verstimmungen sind meßtechnisch sehr genau erfaßbar.

Röntgenographische Spannungsmessung

Bei dieser Methode wird die Änderung der Gitterkonstante mit einem Röntgendiffraktometer gemessen.

4.3.5. Haftfestigkeit

4.3.5.1. Allgemeines

In allen Einsatzbereichen von beschichteten Werkstoffen wird gefordert, dass die Schichten mit einer der technischen Anwendung entsprechenden, genügend großen Haftfestigkeit auf die artverschiedene Unterlage aufgebracht werden und das Werkstück möglichst lange funktionstüchtig bleibt. Die durch ihre Haftfestigkeit gekennzeichnete Adhäsion der Schicht ist eine makroskopische Eigenschaft, die von den atomaren Bindungskräften in der Übergangszone (Interface) zwischen Substrat und Schicht, den inneren Spannungen im Verbund Schicht/Substrat und den jeweiligen Beanspruchungen abhängt. Letztere können durch mechanische (Zug, Scherung), thermische (hohe und niedrige Temperaturen, auch in zyklischer Folge), chemische (Korrosion, sowohl chemisch als auch elektrochemisch) und andere Einwirkungen erzeugt werden.

Nach Mattox wird eine "gute Adhäsion" im allgemeinen dann erreicht, wenn

1. eine starke Atom-Atom-Bindung in der Interfacezone,
2. geringe innere Spannungen in der Schicht,
3. kein leichter Deformations- oder Bruchmodus, und
4. keine Langzeit-Degradation im Schicht-Substrat-Verbund vorliegen.

Die Haftfestigkeit hängt daher insbesondere von den Werkstoffpartnern dieses Verbundes, dem Typ der Interfacezone, der Mikrostruktur und damit den Herstellungsbedingungen der Schicht sowie der Art der Vorbehandlung des Substrats ab. Diese Zusammenhänge sollen im folgenden in ihrer Anwendung mit besonderem Augenmerk auf die PVD-Beschichtungsverfahren dargestellt werden.

4.3.5.2. Interface zwischen Substrat und Schicht

Keimbildung und Schichtaufbau

Die Vorgänge bei der Keimbildung und dem Schichtaufbau wurden bereits ausführlich in Kapitel 3 behandelt. Es wurde gezeigt, dass durch Anlagerung von Adatomen die Keime zu Inseln wachsen, welche wiederum - je nach den Bedingungen - zu einem mehr oder weniger zusammenhängenden Film koaleszieren.

Die Keimdichte und das Keimwachstum bestimmen die Kontaktfläche in der Interfacezone bzw. die auf die Hohlräume entfallende Fläche. Bei geringer Keimdichte ist die Haftfestigkeit wegen der geringen Kontaktfläche und der leichten Bruchausbreitung durch die Hohlräume gering. Die Keimdichte kann durch Ionenbombardement, Oberflächendefekte, Verunreinigungen, die Gasatmosphäre und damit durch die Beschichtungstechnik wesentlich beeinflusst werden.

Nach Mattox unterscheidet man die im folgenden beschriebenen fünf Typen von Übergangszonen (Interface-Typen):

Mechanischer Übergang

Die Substratoberfläche ist aufgerauht und weist Poren auf, in denen das Schichtmaterial verankert wird (siehe Abb. 4.33.a). Dadurch kommt eine für viele Fälle ausreichende, rein mechanische Haftung zustande. Zu beachten ist, dass Rauigkeiten des Substrates zu Abschattungen beim Beschichten und damit zu einer porösen Struktur führen. Andererseits kann sich ein Bruch in einer Schicht, die auf einem durch Sandstrahlen aufgerauhten Werkstück aufgebracht wurde, nur sehr schwer ausbreiten, weil er oft seine Richtung ändern oder festeres Material passieren muss. Beispiele für mechanische Übergangszonen liegen auf dem Gebiet dicker Schichten, die durch thermisches Spritzen oder galvanisch auf Metalle sowie chemisch auf Sinterkeramik oder angeätzten Kunststoff aufgebracht wurden.

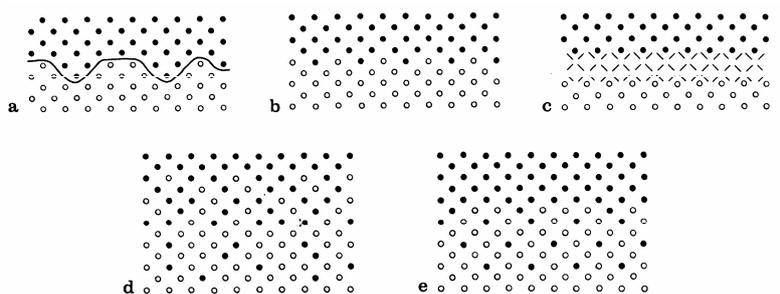


Abb. 4.33.: Schematische Darstellung der fünf Übergangszonen zwischen Substrat und Schicht: **a** Mechanisch; **b** Monoschicht/Monoschicht; **c** chemische Bindung; **d** Diffusion; **e** Pseudodiffusion [1, p. 20]

Monoschicht/Monoschicht-Übergang

Hier besteht ein abrupter Übergang vom Substrat zum Filmmaterial innerhalb weniger Atomlagen (siehe Abb. 4.33.b). Ein solcher Übergang bildet sich, wenn keine Diffusion oder chemische Reaktion zwischen den Partnern auftritt, d.h. keine gegenseitige Löslichkeit besteht, eine zu geringe Energie verfügbar ist oder Verunreinigungen auf dem Substrat vorhanden sind. Unter solchen Bedingungen durch Aufdampfen hergestellte Schichten können Beispiele für Monoschicht/Monoschicht-Übergänge darstellen.

Verbindungsübergang

Dieser Übergang, der durch chemische Reaktion zwischen Substrat- und Schichtmaterial entsteht, setzt chemische Bindung der Partner voraus und wird durch erhöhte Temperatur begünstigt (siehe Abb. 4.33.c). Dabei kann sich in einer viele Atomlagen umfassende Zone eine intermetallische Verbindung, ein Oxid oder eine andere Verbindung bilden. Intermetallische Verbindungen sind meist spröde, brechen leicht und können im Werkstoffverbund Brüche auslösen. Oxide können Wärmebarrieren und günstige Übergänge von Metall zu Keramik darstellen. Beispiele für letzteres sind durch CVD-Verfahren oder thermisches Spritzen erzeugte oxydische Zwischenschichten.

Diffusionsübergang

Bei gegenseitiger Löslichkeit von Substrat- und Schichtmaterial und entsprechender Erwärmung bildet sich durch Interdiffusion ein kontinuierlicher Übergang der chemischen Zusammensetzung, der Gitterparameter und der inneren Spannungen vom Substrat zum Schichtmaterial (siehe Abb. 4.33.d). Bei unterschiedlichen Diffusionsraten beider Materialien können aufgrund des Kirkendall-Effektes Porositäten in der Übergangszone entstehen, es sei denn, letztere ist hinreichend dünn.

Der Diffusionsübergang tritt überall dort auf, wo zwei ineinander lösliche Materialien mit atomar sauberen Flächen unter Erwärmung längere Zeit miteinander in Kontakt sind. Das ist bei vielen Beschichtungsverfahren (Ionenplattieren, CVD, Vakuum-Plasmaspritzen und Schmelztauchen) der Fall. Es kann aber auch beim Aufdampfen und Sputtern erreicht werden, wenn das Substrat entsprechend vorbehandelt ("Ionen-Ätzen") und aufgeheizt wird.

Pseudodiffusionsübergang

Dieser Übergang bildet sich bei Materialien ohne gegenseitige Löslichkeit unter starker Energieeinwirkung auf das Substrat bzw. die aufwachsende Schicht, etwa durch Ionenbombardement oder Ionenimplantation (siehe Abb. 4.33.e). Hochenergetische Ionen oder Neutralteilchen dringen je nach Energie bis zu einer bestimmten Tiefe in das Gitter des Substrates ein und bleiben dort stecken, ohne dass Diffusion eintritt, wie dies z.B. beim Ionenplattieren der Fall ist.

Auch Ionenbeschuß vor der Beschichtung kann durch Erzeugung hoher Konzentrationen von Punktdefekten und von inneren Spannungen die Löslichkeit in der Übergangsschicht und damit die Diffusion während der anschließenden Beschichtung erhöhen. Ein anderes Beispiel liegt vor, wenn zwei geschmolzene Metalle in Kontakt kommen, sich mischen und dann schockartig abgekühlt werden, wie z.B. beim Plasma-Auftragsschweißen.

In der Praxis sind die erhaltenen Interfacezonen oft eine Kombination der verschiedenen Übergangstypen. Hinsichtlich der Haftfestigkeit sind solche am besten,

welche die inneren Spannungen - verursacht durch unterschiedliche Werte der thermischen Ausdehnungskoeffizienten, der Gitterparameter und der Löslichkeiten der Materialpartner - über eine Zone von großem Volumen gleichmäßig verteilen, ohne dass dabei Stellen verminderter Festigkeit oder erhöhter Brüchigkeit auftreten. Diese Bedingungen werden am besten durch Diffusionsübergänge erfüllt, die vielfach durch Erhitzen des Substrates während und/oder nach der Abscheidung erzielt werden.

Mikroskopische Bindungstypen

Wie oben bereits dargestellt, hängt die Haftfestigkeit einer Schicht von der zwischen Schicht- und Substratmaterial bestehenden Bindungsenergie und der Struktur der Interfacezone ab. Dabei kann es sich um eine chemische, eine elektrostatische, eine van der Waals-Bindung oder eine Kombination aus diesen handeln. Die chemische Bindung, die mit einer Bindungsenergie von 0,5 bis etwa 10eV die stärkste unter ihnen ist, kommt dadurch zustande, dass zwischen den Schicht- und Substratatomen Elektronen transferiert oder neu verteilt werden. Dies setzt einen innigen Kontakt der Partneratome voraus. Im Fall der kovalenten, der Ionen- und der metallischen Bindung hängen die Bindungskräfte vom Grad des Elektronentransfers ab. In den ersten beiden dieser Fälle entstehen Verbindungen, die im allgemeinen spröde und brüchig sind, im letzten Fall meistens jedoch duktile Legierungen.

Die van der Waals-Bindung beruht auf einer Polarisationswechselwirkung, die keinen besonders innigen Kontakt der Partneratome verlangt, aber wesentlich schwächer (0,1 bis 0,3eV) als die chemische Bindung ist und mit wachsender Distanz rasch abnimmt.

Es lässt sich zeigen, dass eine typische chemische Bindung (4eV) einer mechanischen Spannung von maximal 10^4Nmm^{-2} widerstehen sollte, und eine typische van der Waals-Bindung (0,2eV) einer Spannung von maximal $5 \cdot 10^2 \text{Nmm}^{-2}$. Messungen der Haftfestigkeit ergeben aber oft geringere Werte als theoretisch zu erwarten wären. Dafür gibt es zwei Gründe: Einmal müssen auch die inneren Spannungen berücksichtigt werden, die sich zu außen angebrachten Spannungen (tensoriell) addieren und dem theoretischen Wert der Haftfestigkeit nahe kommen können. Zum anderen ist die Festigkeit der Interfaceschichten durch Defekte in der Mikrostruktur bestimmt und daher geringer als es der atomaren Bindung entspricht. Und schließlich kann eine geringere Haftfestigkeit ihre Ursache auch in der Struktur der Interfaceschicht haben; z.B. wenn die Kontaktfläche wegen einer geringen Keimdichte bedeutend kleiner als die geometrische Interfacefläche ist.

Da sich durch Ionenbombardement die Keimdichte und damit die Kontaktfläche auf dem Substrat erhöht und die Zonengrenzen im Thornton-Modell zu tieferen T/T_m -Werten hin verschoben werden, ist die Haftfestigkeit von Schichten, die durch Bias-Sputtern erzeugt werden, höher als die von Aufdampfschichten. Noch höhere Haftfestigkeitswerte besitzen Schichten mit einer durch Diffusion und/oder chemischen Reaktion erzeugten Interfacezone, wie z.B. durch Ionenplattieren und CVD-Verfahren hergestellte Schichten.

4.3.5.3. Vorbehandlung der Substrate

Allgemeines

Die Reinheit der Substratoberfläche hat einen entscheidenden Einfluss auf die Mikrostruktur und Haftfestigkeit der Schicht. Die Merkmale, die eine "reine" oder besser gesagt, geeignete Substratoberfläche definieren, hängen von der jeweiligen Anwendung der Schicht ab. Wenn Gasadsorption und/oder -desorption, Oberflächendiffusion oder ähnliche Grenzflächenerscheinungen untersucht werden sollen, wird eine Oberfläche gefordert, die nur mit einem geringen Bruchteil einer Monoschicht von Fremdatomen belegt ist. Hier wird also eine atomar reine Oberfläche verlangt, wie sie nur mit UHV-Techniken herstellbar ist.

Soll andererseits die Stahlkonstruktion einer Brücke durch eine Zinkschicht gegen Korrosion geschützt werden, so genügt Sandstrahlen als Vorbehandlung vollauf. Dadurch werden Rost- und Zunderschichten entfernt, und eine mechanische Übergangszone durch Aufrauen vorbereitet.

Im folgenden sollen spezielle Reinigungsverfahren besprochen werden, die bei den PVD-Prozessen angewendet werden. Hier besteht der Reinigungsprozess im allgemeinen aus einer Vorreinigung außerhalb der Vakuumanlage und einer Reinigung innerhalb derselben als Teilprozeß der Beschichtung.

Die Vakuumanlage wird vielfach so aufgestellt, dass ihre Prozesskammer von einem staubfreien Raum aus zugänglich ist, während der übrige Teil sich in einem Service-Raum befindet.

Vorreinigung

Prinzipiell muss außerhalb der Beschichtungsanlage eine Grobreinigung der Substrate vorgenommen werden. Oftmals sind die zu beschichtenden Substrate stark verschmutzt (z.B. Korrosionsschutzöl bei Gleitlagern) und müssen durch manuelles Putzen, Glasperlen, Sandstrahlen, Polieren, Anätzen in verdünnten Säuren oder Laugen u.ä. gereinigt werden. Viele verschiedene Verfahren für eine weiterführende Reinigung sind für die unterschiedlichsten Typen von Substraten in der Literatur angegeben. Vielfach wird eine Kombination von mehreren Lösungsmitteln und Ultraschallbehandlung gewählt. Die Hervorragende Reinigungswirkung der Ultraschallbehandlung beruht auf "Mikroschruppen" durch Implosion von Gasbläschen in der Flüssigkeit.

Glimmentladungsreinigung

Nachdem in der Bedampfungsanlage ein ausreichendes Hochvakuum hergestellt ist, wird bei gedrosseltem Saugvermögen der Pumpe ein inertes Gas, z.B. Argon von 10Pa eingelassen und eine Glimmentladung zwischen einer isoliert eingeführten stab- oder ringförmigen Al-Kathode und dem geerdeten Substrathalter als Anode gezündet. Die Daten betragen 1mAcm^{-2} am Substrathalter und 2kV Spannung.

Ist das Substrat isolierend, so lädt es sich aufgrund der gegenüber den Ionen höheren Beweglichkeit der Elektronen negativ auf. Dadurch entsteht ein Ionenstrom von Plasma zum Substrat, der letzteres durch Zerstäuben reinigt. Elektrisch leitende Substrate werden von Elektronen aus dem Plasma getroffen, die Desorption von Verunreinigungen sowie ein gewisses Aufheizen bewirken. Ferner schlägt sich auf den Substraten zerstäubtes Glimmelektrodenmaterial nieder und bildet hier Kondensationskeime.

Sputterreinigung

Um ein isolierendes Substrat vor dem Beschichten durch Zerstäuben in der Sputteranlage zu reinigen, muß mit einer HF-Spannung gearbeitet werden. Beim Beschichten durch Ionenplattieren ist das kontinuierliche Sputterreinigen ein integraler Bestandteil des Prozesses. Isolierende Substrate müssen auch hier an eine HF-Spannung angeschlossen sein, um zu vermeiden, dass sie sich positiv aufladen und damit das Ionenbombardement verhindern.

DC-Sputterreinigung ("Sputter-Ätzen") ist die adäquate Methode zur Reinigung von leitenden Substraten durch Ionenbeschuß.

Verbesserung der Haftfestigkeit

Wie bereits erwähnt, liefern das Bias-Sputtern und das Ionenplattieren besser haftende Schichten als einfaches Aufdampfen. Neben den in-situ-Reinigungsverfahren durch Ionen- und/oder Elektronenbeschuß ergeben sich weitere Möglichkeiten zur Verbesserung der Haftfestigkeit:

... Haftvermittler-Schichten: Eine dünne Schicht aus einem reaktiven Metall, wie Ti, Cr etc. wird zunächst auf dem Glas- oder Keramiksubstrat niedergeschlagen, und dann die gewünschte Metallschicht, z.B. Au oder Ag. Die haftvermittelnde Schicht reagiert chemisch oder durch Diffusion mit dem Substrat und legiert sich mit der Metallschicht.

... Gradiertes Oxid-Übergang: Das Aufstäuben einer Metallschicht erfolgt unter Zugabe von Sauerstoff, dessen Partialdruck anfänglich relativ hoch ist und dann langsam auf 0 reduziert wird.

... Implantation oder Diffusion von haftvermittelnden Atomen: Um z.B. die Haftung von W auf SiO_2 -Flächen zu verbessern, kann man Al, das eine geringere Valenzzahl (+3) als Si (+4) hat, in die Oberfläche eindringen lassen. Dadurch entsteht eine unabgesättigte Sauerstoffbindung, die dann zur Bindung von W zur Verfügung steht.

4.3.5.4. Messung der Haftfestigkeit

Allgemeines

Unter idealen Bedingungen sollte ein Gerät für die Bestimmung der Haftfestigkeit quantitative und reproduzierbare Ergebnisse liefern, eine zerstörungsfreie und schnelle Messung an Proben unterschiedlicher Form und Größe ermöglichen, einen weiten Meßbereich haben, standardisierbar und unabhängig von der Erfahrung des jeweiligen Operators sein. In der Praxis ist man jedoch noch sehr weit davon entfernt.

Erschwerend kommt beim Durchführen von Adhäsionsmessungen an dünnen Schichten hinzu, dass an den Grenzflächen durch Inhomogenitäten und Unzulänglichkeiten vor oder während der Beschichtung Abweichungen auftreten, die örtlich sehr verschieden sein können. Weiters ist in Betracht zu ziehen, dass bei den meisten Untersuchungsmethoden die Verbindung zwischen Probe und Prüfapparatur unvollkommen ist, was den Messbereich der jeweiligen Meßmethode begrenzt und die Ergebnisse in nicht kontrollierbarer Weise beeinflusst.

Aus den oben genannten Schwierigkeiten ergibt sich die Tatsache, dass hauptsächlich empirisch gefundene Ergebnisse vorliegen. Trotzdem sollen hier einige der unzähligen Prüfmethoden vorgestellt werden.

Klebebandtest

Dabei wird ein Klebeband auf die zu prüfende Schicht aufgebracht und wieder abgezogen. Der Klebebandtest ist ein subjektiver Test, der nicht nur von der Art des Klebebandes, sondern auch von der Schnelligkeit und vom Winkel beim Abziehen abhängig ist. Löst sich die Schicht beim Abziehen teilweise oder ganz ab, so liegt auf jeden Fall eine schlechte Haftung vor (siehe Abb.4.34.). Bleibt die Schicht am Substrat, ist trotzdem keine weitergehende Aussage darüber möglich, ob die Haftfestigkeit z.B. mäßig, gut oder sehr gut ist.

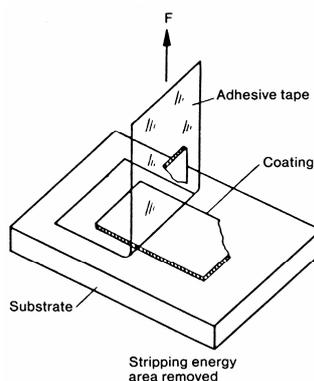


Abb. 4.34.: Schematische Darstellung des Klebebandtests [4, p. 80]

Gitterschnitttest

Eine etwas bessere Aussagekraft als der Klebebandtest hat der Gitterschnitttest. Mit einem geeigneten Schneidegerät (DIN-Norm) wird ein bis auf den Untergrund durchgehendes Schnittband mit sechs Schnitten und im rechten Winkel dazu ebenfalls ein solches gezogen, so dass ein Gitter mit 25 Quadraten entsteht. Nachfolgend wird auf dem Gitternetz ein Klebebandtest durchgeführt und je nach prozentualen Anteil der abgelösten Schichtfläche die Haftfestigkeit in sechs Stufen unterteilt (siehe Abb. 4.35).

Gitterschnittkennwert	Beschreibung	Bild
Gt 0	Die Schnittländer sind vollkommen glatt, kein Teilstück des Anstriches ist abgeplatzt.	-
Gt 1	An den Schnittpunkten der Gitterlinien sind kleine Splitter des Anstriches abgeplatzt; abgeplatzte Fläche etwa 5% der Teilstücke.	
Gt 2	Der Anstrich ist längs der Schnittländer und/oder an den Schnittpunkten der Gitterlinien abgeplatzt; abgeplatzte Fläche etwa 15% der Teilstücke.	
Gt 3	Der Anstrich ist längs der Schnittländer teilweise oder ganz in breiten Streifen abgeplatzt und/oder der Anstrich ist von einzelnen Teilstücken ganz oder teilweise abgeplatzt; abgeplatzte Fläche etwa 35% der Teilstücke.	
Gt 4	Der Anstrich ist längs der Schnittländer in breiten Streifen und/oder von einzelnen Teilstücken ganz oder teilweise abgeplatzt; abgeplatzte Fläche etwa 65% der Teilstücke.	
Gt 5	Abgeplatzte Fläche mehr als 65% der Teilstücke.	-

Abb. 4.35.: Beurteilungskriterien der Haftfestigkeit nach dem Gitterschnitttest [3, p. 304]

Quantitativer Abziehtest

Bei dieser Prüfmethode wird ein zylindrischer Körper auf einer Stirnfläche mittels einer Verbindungsschicht (Kleber) auf der Schichtseite der Probe befestigt (siehe Abb. 4.36.). Durch eine Meßvorrichtung wird festgestellt, welche Kraft erforderlich ist, um den zylindrischen Körper von der Unterlage zu lösen.

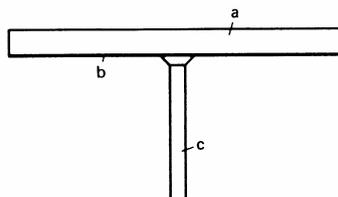


Abb. 4.36.: Beschichtetes Substrat mit aufgeklebtem Prüfkörper [3, p. 305]

Die Aussagekraft dieser Methode wird dadurch eingeschränkt, dass beim Aufbringen des Klebers (meist bei erhöhter Temperatur) durch Diffusionsprozesse oder durch Änderung der inneren Spannung in der Schicht die ursprüngliche Haftfestigkeit zwischen Substrat und Schicht verändert wird.

Naturgemäß sind brauchbare Ergebnisse auch nur dann zu erwarten, wenn die Haftfestigkeit des Klebers (sowohl zur Schicht als auch zum Zylinder) wesentlich höher ist als die der Schicht am Substrat. Bei sorgfältiger Durchführung sind Messungen der Haftfestigkeit bis zu 5 kNcm^{-2} möglich.

Ultrazentrifugentest

Das Hauptproblem bei allen Abziehtests, nämlich die Beeinflussung der Meßergebnisse durch den Verbund zwischen Meßeinrichtung und Prüfling und die dadurch entstehenden Veränderungen, tritt beim Ultrazentrifugentest nicht auf. Die zu prüfende Schicht befindet sich auf einem Rotor, der sich so schnell dreht, dass die Schichtablösung allein durch die Zentrifugalkraft zustande kommt.

Für die Berechnung der Haftfestigkeit ist lediglich die Kenntnis von Dicke und Dichte der auf dem Rotor aufgetragenen Schicht erforderlich. Unter der Voraussetzung, dass die Schichtdicke wesentlich kleiner als der Rotordurchmesser ist, gilt

$$A = 4\pi^2 N^2 R \rho d \quad (4.35)$$

mit

A = Haftfestigkeit [Nm^{-2}]
 R = Rotorradius [m]
 d = Schichtdicke [m]
 N = Umdrehungen [s^{-1}]
 ρ = Dichte der Schicht [kgm^{-3}]

Damit können Haftfestigkeitswerte bis zu $50kNcm^{-2}$ gemessen werden. Neben dem Problem der hohen Drehzahlen bei dünnen Schichten (geringe Massenbelegung) ergibt sich der schwerwiegende Nachteil, dass die Haftung auf technischen Bauteilen nicht gemessen werden kann, sondern ein Rotor (aus dem selben Substratmaterial!) eigens beschichtet werden muss.

Kratz- bzw. Ritztest

Ein abgerundeter Kegel aus Stahl, Wolframkarbid oder Diamant wird in senkrechter Stellung über den Prüfling geführt und dabei schrittweise mit immer höheren Gewichten L belastet, bis bei einer kritischen Last L_c die Schicht den Deformationen des Substrates nicht mehr folgen kann und bricht bzw. sich von der Unterlage abhebt. Diese kritische Last L_c hängt sowohl von der Schicht als auch vom Substratmaterial ab und ist ein Maß für die Haftfestigkeit.

Die Auswertung erfolgt zumeist im Raster-Elektronenmikroskop bzw. bei automatisch arbeitenden Geräten durch einen Schallemissionsdetektor, der die Geräusche erfaßt, die durch den Bruch der Schicht erzeugt werden (siehe Abb. 4.37.).

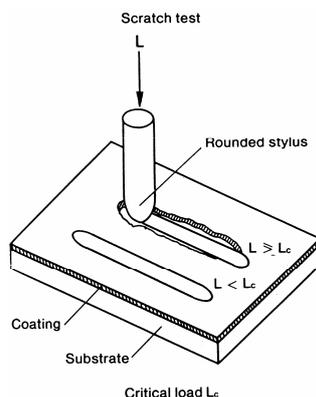


Abb. 4.37.: Kratz- bzw. Ritztest [4, p. 83]

Nach der Theorie ergibt sich die Haftfestigkeit σ_A aus

$$\sigma_A = K \sqrt{\frac{L_c H}{\pi R^2}} \quad (4.36)$$

wobei gilt:

R = Krümmungsradius der Spitze [mm]

H = Härte des Substrates [Nmm⁻²]

L_c = kritische Last [N]

K = experimentell bestimmbarer Koeffizient [dimensionslos]

Im Prinzip wäre der Kratz- bzw. Ritztest für eine quantitative Messung der Haftfestigkeit geeignet und findet auch zunehmend Verwendung. Der Prozess ist jedoch von so komplexer Natur, dass es noch keine allgemein anerkannte Theorie gibt, die alle eine Absolutmessung beeinflussenden Größen erfasst und berücksichtigt. Für Relativmessungen zeichnet sich diese Methode aber durch gute Reproduzierbarkeit aus und ist auch für außergewöhnlich gut haftende Schichten anwendbar.

4.3.6. Reibung und Verschleiß

Der Reibungskoeffizient und die Verschleißrate ($\Delta d/\Delta s$ = Schichtdickenabnahme Δd pro Reibungsweg Δs) werden meist mit einer Stift-Scheibe (pin-on-disk)-Apparatur unter definierten Bedingungen (Belastung und Material des Stiftes, Umfangsgeschwindigkeit u.a.) in einer genau definierten Atmosphäre (trockener Stickstoff; Wasserdampf; UHV etc.) gemessen.

Für manche Anwendungen kommen auch Prüfmaschinen zum Einsatz, die mit speziellen Testprogrammen die Verschleißfestigkeit von tribologischen Beschichtungen überprüfen.

4.3.7. Duktilität

Die Duktilität ist ein wichtiges Kriterium für die plastische Verformbarkeit einer dünnen Schicht und dient zu deren Beurteilung, wenn das Werkstück nach der Beschichtung noch einer Verformung unterworfen werden soll. Mit der Zunahme von dekorativen Beschichtungen und Hartstoffbeschichtungen mit PVD- und CVD-Methoden gewinnt die Duktilitätsmessung auch für diesen Bereich an Bedeutung.

Die Duktilität massiver Metalle wird mit Hilfe der üblichen Zugfestigkeitsprüfungen bestimmt. Als Maß der Duktilität gilt die prozentuale Bruchdehnung einer Probe. Duktile Metalle haben eine hohe, spröde Materialien eine relativ geringe Bruchdehnung. Es gilt

$$\varepsilon_B = \frac{l_Z - l_0}{l_0} \quad (4.37)$$

mit

l_Z = Länge der Probe nach dem Bruch

l_0 = Länge der unbelasteten Probe

ε_B = Bruchdehnung in %

Die Bruchdehnung ε_B von dünnen Schichten kann durch eine Dornbiegeprüfung bestimmt werden. Die auf einem duktilen Metallstreifen abgeschiedene Schicht wird um zylindrische Dorne abnehmender Durchmesser gebogen und derjenige Biegeradius R festgestellt, bei dem in der Beschichtung die ersten Risse auftreten. Es gilt dann mit der Nebenbedingung $d \ll R$ die Beziehung

$$\varepsilon_B = \frac{100d}{2R + d} \quad (4.38)$$

wobei d die Dicke der Schicht darstellt. Eine Variation dieser Technik stellt die drei-Punkt Biegemethode dar. Der Duktile Teststreifen liegt in der Mitte auf und wird an beiden Enden definiert belastet, bis sich erste Risse bilden. Das kann z. B. in-situ in einem REM ausgeführt werden

4.3.8. Härte

4.3.8.1. Allgemeines

Die Härte ist definiert als der mechanische Widerstand, den ein Körper beim Eindringen einem anderen, härteren Körper entgegensetzt. Die Messwerte hängen nicht nur von den elastischen und plastischen Eigenschaften des Materials ab, sondern auch von der Form und Beschaffenheit des Prüfkörpers und von der Art der Untersuchungen. Man unterscheidet daher z.B. die Härte nach Vickers, nach Rockwell, nach Brinell usw.

4.3.8.2. Mikrohärtemessung

Bei der Härtmessung an dünnen Schichten wendet man üblicherweise die Mikrohärteverfahren nach Vickers oder Knoop an. Dabei handelt es sich nicht, wie nach dem Ausdruck "Mikrohärte" angenommen werden könnte, um sehr kleine Härten, sondern um Härteprüfungen mit kleinen Prüfeindrücken. Nach einer zwar nicht standardisierten, aber allgemein akzeptierten Definition des Begriffes Mikrohärteliegen die Prüfgewichte zwischen etwa 0,02 und 2N, in der Praxis meist zwischen 0,05 und 0,5N. Bei Ultramikrohärtemessungen werden Prüfgewichte von 10^{-2} bis herab zu ca. 10^{-5} N verwendet.

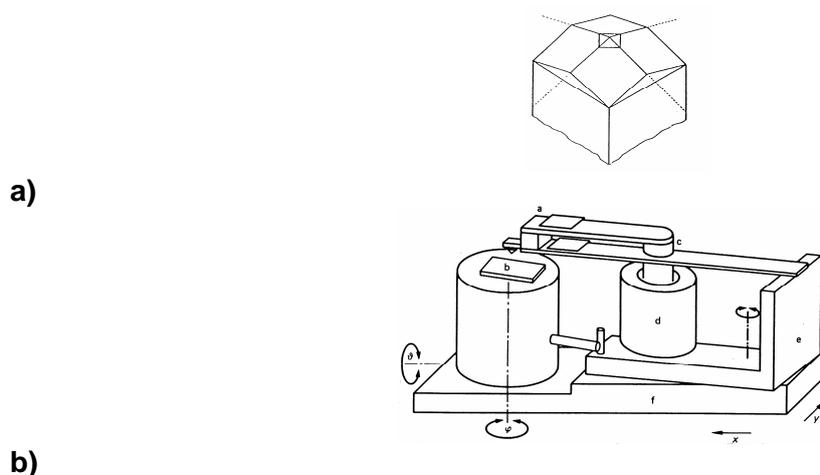


Abb. 4.38.: a) Vickers-Pyramide b) Schematischer Aufbau eines Ultramikrohärteprüfers [3, p.301]
a Dehnungsmeßstreifen; b Probe; c Doppelfeder; d Spule;
e Kupplung; f Grundplatte

Bei der Mikro- und auch bei der Ultramikrohärtemessung wird durch eine Diamantpyramide ein Eindruck in der zu messenden Schicht erzeugt und optisch vermessen. Die Grundfläche der Vickerspyramide ist quadratisch, der Spitzenwinkel zwischen den gegenüberliegenden Pyramidenflächen beträgt jeweils 136° (siehe Abb. 4.38.a).

Die Mikrohärtete kann bei bekannter Prüfkraft p in N und bekannter Länge d der Eindruckdiagonalen in μm aus

$$HV = \frac{18192p}{d^2} \quad (4.39)$$

berechnet werden.

Es ist leicht verständlich, dass eine Härtemessung an den sehr dünnen Schichten für optische oder elektronische Anwendungen mit handelsüblichen Härteprüfgeräten nicht möglich ist. Wegen der geringen Schichtdicke ($<1\mu\text{m}$) muss mit entsprechend niedrigen Prüfgewichten gearbeitet werden; und wegen der damit verbundenen kleinen Eindrücke scheidet eine Auswertung im Lichtmikroskop aus.

Das von Bangert und Wagendristel entwickelte Ultramikrohärtemeßgerät für den Einsatz im Raster-Elektronenmikroskop ist kommerziell erhältlich und ermöglicht Messungen mit extrem kleinen Prüflasten bis unter 10^{-4} N (siehe Abb. 4.38.b) Den letzten Entwicklungsstand stellt der sogenannte "Nano-Indenter" dar, welcher es auch erlaubt, die Abhängigkeit der Eindringtiefe des Prüfkörpers von der Last zu erfassen. Die Eindruckgeometrie kann z. B. mittels eines AFM bestimmt werden.

4.4. Elektrische Eigenschaften

4.4.1. Einleitung

Die globale wirtschaftliche und technologische Bedeutung der Mikroelektronik ist wohl unumstritten. Die Entwicklung, die Erzeugung und die Untersuchung von Bauteilen für die Mikroelektronik stellt somit ein sehr großes Teilgebiet der Dünnschichttechnologie dar. Das Verständnis der elektrischen Leitfähigkeit ist letztlich die Grundlage für die Entwicklungen der vergangenen drei Jahrzehnte, die von der diskret aufgebauten Diode bis zum VLSI-Schaltkreis oder bis zum Gigabit-Speicherchip reicht.

Die elektrische Leitfähigkeit dünner Schichten weicht aufgrund des Größeneffektes (Schichtdicken $< 1\mu\text{m}$) und wegen vorhandener Schichtdefekte von dem Wert des massiven ("Bulk")-Materials ab. Sie kann aus der Messung des Flächenwiderstandes R_N bestimmt werden.

4.4.2. Definition und Messung

Der Flächenwiderstand ist derjenige Widerstand, den man zwischen zwei Elektroden misst, die an gegenüberliegenden Seiten einer Schicht mit quadratischer Fläche angebracht sind. Der Flächenwiderstand R_N ist unabhängig von der Größe der quadratischen Fläche, denn für den elektrischen Widerstand eines Quaders mit der Dicke d , Breite b und Länge l gilt (siehe auch Abb. 4.39.):

$$R = \rho \frac{l}{db} \quad (4.40)$$

Bei quadratischer Oberfläche ($l=b$) ist

$$R = R_{\square} = \frac{\rho}{d} \quad (4.41)$$

unabhängig von der Größe des Quadrates.

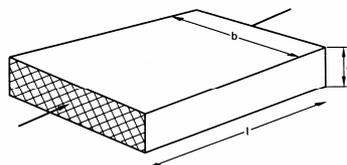


Abb. 4.39.: Definition des Flächenwiderstandes [3, p. 262]

Der Flächenwiderstand wird mit der Vierpunktmethode gemessen. Dazu werden vier punktförmige Elektroden auf die Schichtoberfläche gesetzt. Die Elektroden können linear mit jeweils gleichem Abstand voneinander oder in den vier Ecken eines Quadrates angeordnet sein (siehe Abb. 4.40.). Bei der linearen Anordnung wird durch die äußeren beiden Elektroden ein Stromfluß I (mit einer Konstantstromquelle im mA- bzw. μA -Bereich) erzeugt und der Spannungsabfall U zwischen den inneren beiden Elektroden gemessen.

Es gilt für den Flächenwiderstand

$$R_{\square} = 4,532 \frac{U}{I} \quad (4.42)$$

Im Falle der quadratischen Elektrodenanordnung wird der Strom durch die benachbarten Elektroden geschickt und der Spannungsabfall zwischen den beiden gegenüberliegenden Elektroden gemessen:

$$R_{\square} = 9,06 \frac{U}{I} \quad (4.43)$$

Die Beziehungen (4.42) und (4.43) gelten nur für den Fall, dass der Durchmesser der zu prüfenden Schichtfläche groß gegen den Elektrodenabstand ist, anderenfalls müssen geometrische Korrekturfaktoren für die Messwerte berücksichtigt werden, die in der Literatur angegeben sind.

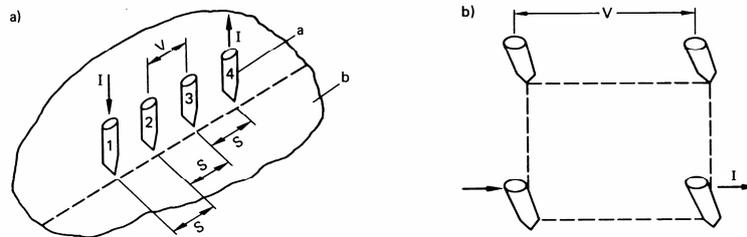


Abb. 4.40.: Bestimmung des Flächenwiderstandes [3, p.263]

- a:** Lineare Vier-Punkt-Methode
- b:** Quadratische Vier-Punkt-Methode
- I** eingepprägter Strom; **U** Messspannung;
- s** Sondenabstand; **a** Sonden; **b** Probe

4.4.3. Theoretische Grundlagen

4.4.3.1. Leitfähigkeit durchgehender dünner Schichten

Die Beschreibung der elektrischen Leitfähigkeit in dünnen Schichten geht auf Fuchs und Sondheimer (ca. 1940) zurück. Ausgangspunkt dafür ist die Theorie von Drude für die elektrische Leitfähigkeit der Metalle.

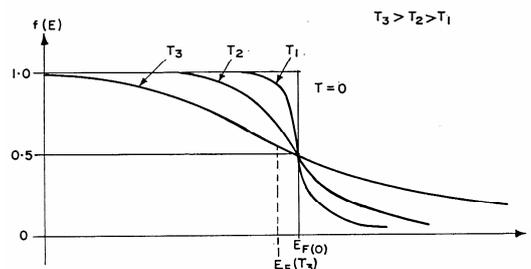


Abb. 4.41.: Fermi-Dirac-Statistik: Verteilungsfunktion $f(E)$ [10, p. 30]

Die Leitungselektronen werden als freies Elektronengas angenommen, welches der Fermi Dirac Statistik gehorcht. Diese gibt die Verteilung im Gleichgewicht an:

$$dN = 2 f_0(E) (V / h^3) m^3 dv_x dv_y dv_z \quad (4.44)$$

mit

$$f_0(E) = \frac{1}{1 + e^{(E-E_f)/kT}} \quad (4.45)$$

Ohne äußeres Feld ergibt sich aufgrund der Symmetrie kein Nettostromfluss (siehe Anhang)

$$\mathbf{j} = -e \int_N \mathbf{v} dN / V = 2e \int_v f_0(E) \mathbf{v} (m^3 / h^3) dv_x dv_y dv_z = 0 \quad (4.46)$$

Ein extern angelegtes Feld verändert jedoch die Energieverteilung in der Art, dass ihre sphärische Symmetrie verloren geht. Die Änderung der Verteilungsfunktion kann mit Hilfe der Boltzmann Gleichung berechnet werden.

Unter der Annahme kleiner Änderungen gilt für den stationären Fall:

$$\frac{df}{dt} = \frac{e\mathbf{E}}{m} \text{grad}_v f - \mathbf{v} \text{grad}_r f + \left(\frac{df}{dt} \right)_{coll} \quad (4.47)$$

Mit dem Ansatz

$$f(t) - f_0 = C e^{-t/\tau} \quad (4.48)$$

wird der Kollisionsterm

$$\left(\frac{df}{dt} \right)_{coll} = - \frac{f(t) - f_0}{\tau} \quad (4.49)$$

welcher voraussetzt, dass die Verteilung nach einer Störung wieder exponentiell dem ungestörten Zustand zustrebt. Die Zeitkonstante τ ist die mittlere Zeit zwischen zwei Stößen. In einem isotropen Metall gilt im Gleichgewicht $df/dt = 0$; somit ergibt (4.47):

$$\frac{e\mathbf{E}}{m} \text{grad}_v f - \mathbf{v} \text{grad}_r f - \frac{f(t) - f_0}{\tau} = 0 \quad (4.50)$$

Unter der Annahme, dass das E-Feld nur eine Komponente in der x-Richtung aufweist und die Verteilung der Elektronen nur in der z-Richtung (durch die Grenzflächen der Schicht) beeinflusst wird ergibt sich:

$$\frac{eE}{m} \frac{df}{dv_x} - v_z \frac{df}{dz} = \frac{f(t) - f_0}{\tau} \quad (4.51)$$

Um für dünne Schichten die Streuung der Elektronen an den Interfacegrenzen zu berücksichtigen wird die Verteilungsfunktion f geschrieben als

$$f = f_0 + A(E, z) \quad (4.52)$$

wobei A eine kleine Störfunktion darstellt. Gleichung (4.51) ergibt damit

$$v_z \frac{dA}{dz} + \frac{A}{\tau} = \frac{eE}{m} \frac{df_0}{dv_x} \quad (4.53)$$

welche die allgemeine Lösung

$$A = \frac{eE\tau}{m} \frac{df_0}{dv_x} (1 + Ke^{-z/\tau v_z}) \quad (4.54)$$

besitzt, mit der Integrationskonstanten K , die mit Hilfe der Randbedingungen zu bestimmen ist.

Die Stromdichte j ist gegeben durch

$$j = -e \int v_x f dv_x dv_y dv_z \quad (4.55)$$

und die Leitfähigkeit σ durch

$$\sigma = j / E \quad (4.56)$$

Alle weiteren Ergebnisse werden auf die Leitfähigkeit des Bulk-Materials σ_0 bezogen. Die Reflexionen der Elektronen an den Schichtgrenzen kann entweder als vollkommen diffuse Streuung oder als teilweise gerichtete Reflexion behandelt werden.

Diffuse Streuung

Unter der Annahme einer vollständig diffusen Streuung gilt für die Elektronen, dass diese nach einem Stoß an der Grenzfläche wiederum mit einer Gleichgewichtsverteilung starten; d.h. die Felddrift durch das E -Feld geht vollständig verloren. Daraus folgt - nach längerer Rechnung - für die Leitfähigkeit

$$\frac{\sigma}{\sigma_0} = 1 - \frac{3}{8k} + \frac{3k}{4} \left(1 - \frac{k^2}{12} \right) \int_k^\infty \frac{e^{-x}}{x} dx + \left(\frac{3}{8k} - \frac{5}{8} - \frac{k}{16} + \frac{k^2}{16} \right) e^{-k} \quad (4.57)$$

wobei gilt: $k = d/\lambda_0$, mit d = Schichtdicke und λ_0 = mittlere freie Weglänge der Leitungselektronen im Bulk-Material (10 bis 40nm bei Raumtemperatur).

Für Spezialfälle kann der Ausdruck (4.57) vereinfacht werden:

$$\frac{\sigma}{\sigma_0} = 1 - \frac{3}{8k} \quad \text{für } k \gg 1 \quad (4.58)$$

und

$$\frac{\sigma}{\sigma_0} = \frac{3k}{4} \ln \frac{1}{k} \quad \text{für } 0 < k \ll 1 \quad (4.59)$$

Teilweise Gerichtete Reflexion

Für den Fall der teilweise gerichteten Reflexion der Elektronen an der Grenzfläche ist einsichtig, dass vollständig elastische Stöße der Elektronen an der Grenzfläche die Leitfähigkeit nicht ändern, da die Drift unverändert bleibt. Es ist daher nur der Fall von Interesse, wo eine teilweise inelastische Streuung angenommen wird. Mit p = Anteil der elastisch gestreuten Elektronen ergibt sich

$$\frac{\sigma}{\sigma_0} = 1 - \frac{3(1-p)}{8k} + \frac{3}{4k}(1-p)^2 \sum_{n=0}^{\infty} p^{n-1} \left[E_1(kn) \left(k^2 n^2 - \frac{k^4 n^4}{12} \right) + e^{kn} \left(\frac{1}{2} - \frac{5kn}{6} - \frac{k^2 n^2}{12} + \frac{k^3 n^3}{12} \right) \right] \quad (4.60)$$

mit

$$E_1(k) = \int_k^{\infty} \frac{e^{-x}}{x} dx \quad (4.61)$$

Für die Vereinfachung $k > 1$ gilt

$$\frac{\sigma}{\sigma_0} = 1 - \frac{3}{8k}(1-p) \quad (4.62)$$

In Abbildung 4.42. ist die Abhängigkeit des spezifischen Widerstandsverhältnisses ρ/ρ_0 von k zu sehen.

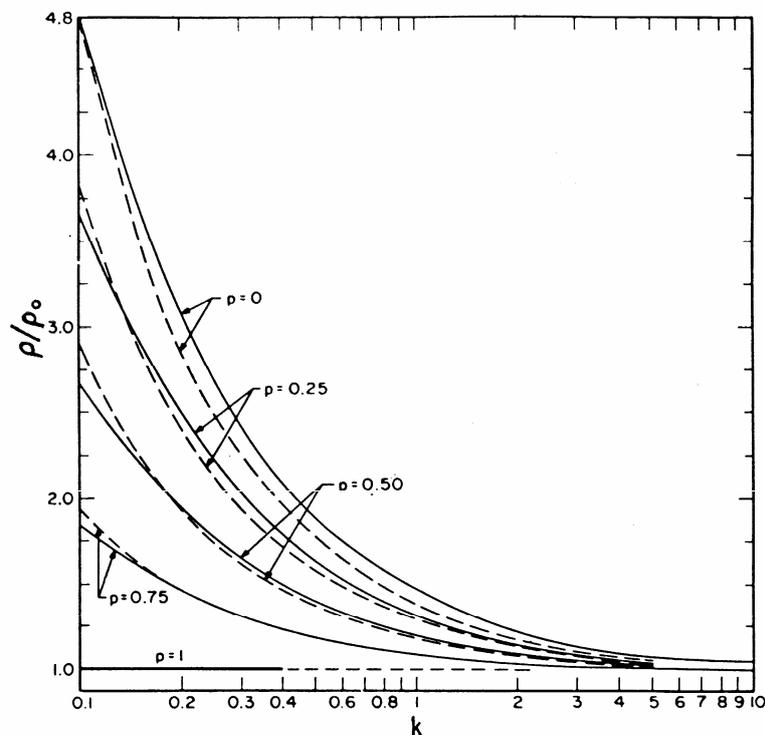


Abb. 4.42.: Spez. Widerstandsverhältnis in Abhängigkeit von k [10, p. 155]:
 — für Glg. (4.60)
 - - - für Glg. (4.62)

Obwohl obige Betrachtung bereits teilweise gerichtete Reflexion berücksichtigt und eine Verbesserung gegenüber dem einfachen Modell der diffusen Streuung bringt, muss darauf hingewiesen werden, dass eine genaue Beschreibung damit nicht möglich ist.

Gründe dafür sind:

... Streuung der Elektronen tritt an beiden Grenzflächen der Schicht auf, was die Notwendigkeit nach sich zieht, zwei verschiedene Streuparameter anzunehmen.

... Streuung der Elektronen am Kristallgitter (τ_G), an Korngrenzen (τ_K) und an Verunreinigungen (τ_V).

Damit kann die in Glg. 4.25 definierte Relaxationszeit τ durch

$$\frac{1}{\tau} = \frac{1}{\tau_G} + \frac{1}{\tau_K} + \frac{1}{\tau_V} \quad (4.63)$$

dargestellt werden. („Matthiessen-Regel“).

4.4.3.2. Elektrische Leitfähigkeit in diskontinuierlichen Schichten

Wenn noch keine durchgehende Beschichtung vorhanden ist, wird die elektrische Leitfähigkeit einer metallischen Schicht durch deren mittlere Dicke bestimmt (siehe auch Abb. 4.43.).

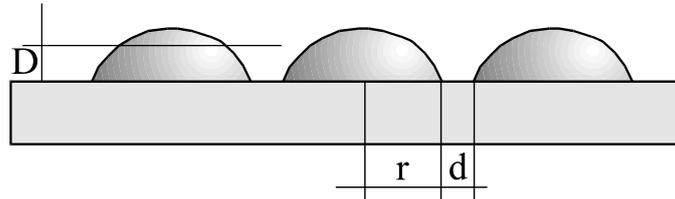


Abb. 4.43.: Schematische Darstellung eines diskontinuierlichen Films

Die experimentellen Befunde:

1. $\sigma_{dis} \ll \sigma_{kont}$
2. $\sigma \text{ prop. } e^{-A/kT}$
3. $\sigma = \sigma(E)$
4. $\sigma = \sigma(r, d)$

können durch zwei verantwortliche Mechanismen erklärt werden.

1. Thermische Emission von Elektronen entweder ins Vakuum oder in das Leitungsband des isolierenden Substrates. Dieser Mechanismus überwiegt für $d > 10\text{nm}$ und $T > 300\text{K}$
2. Tunneleffekt: wiederum über Substrat und Vakuum möglich

4.4.4. Praktische Aspekte

4.4.4.1. Allgemeines

Die Anforderungen an Dünnschichtwiderstände sind vielfältig. Eine der wesentlichsten ist die große Breite des abzudeckenden Widerstandsbereiches, der sich von etwa $100\Omega/\square$ bis zu $100M\Omega/\square$ erstreckt. Eine Variation des Flächenwiderstandes über so viele Größenordnungen lässt sich bei begrenzter Substratgröße nicht allein über geometrische Änderungen an den Schichten realisieren, sondern es sind verschiedene Materialien mit unterschiedlichem spezifischen elektrischen Widerstand erforderlich. Abb. 4.44. zeigt für verschiedene gebräuchliche Materialien für Widerstandsschichten eine Übersicht über die Abhängigkeit des elektrischen Temperaturkoeffizienten vom Flächenwiderstand.

Für Schichten mit relativ niedrigem elektrischen Widerstand kommen Metalle oder Legierungen in Frage, für solche mit höherem Widerstand sogenannte Cermets (Keramik-Metall-Kombinationen) und für Isolatoren natürlich die klassischen Isolationsmaterialien wie beispielsweise Quarz.

Weitere wichtige Forderungen sind ein geringer elektrischer Temperaturkoeffizient, gute Stabilität der elektrischen Eigenschaften und kostengünstige Herstellung. bei der Schichterstellung spielen daher nicht nur die verwendeten Materialien eine Rolle, sondern ebenso wichtig sind die gewählten Beschichtungsverfahren und Parameter, die ebenfalls die Schichteigenschaften beeinflussen.

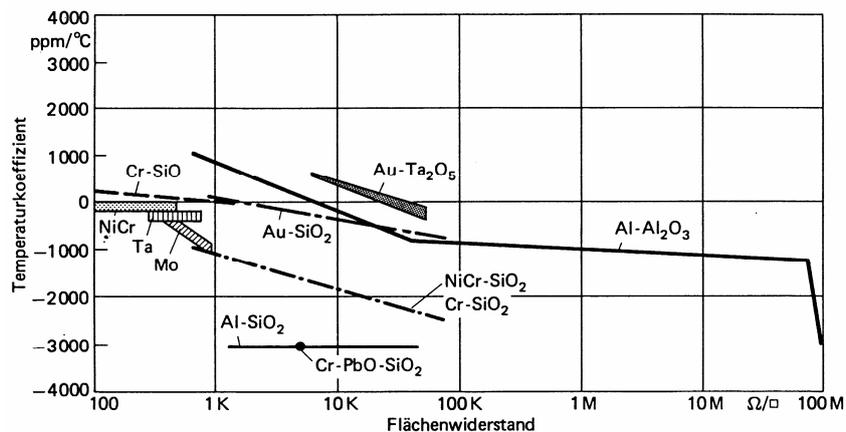


Abb. 4.44.: Elektrischer Temperaturkoeffizient verschiedener Materialien in Abhängigkeit vom Flächenwiderstand [3, p. 403]

4.4.4.2. NiCr-Widerstandsschichten

Für niedrige Widerstandswerte werden wegen ihrer hervorragenden Eigenschaften (Rauscharmut, Stabilität, hohe Belastbarkeit) bevorzugt NiCr-Schichten angewendet. Sowohl Flächenwiderstand als auch Temperaturkoeffizient hängen von der jeweiligen Legierungszusammensetzung ab. Die elektrischen Eigenschaften werden bei den PVD-Verfahren besonders kritisch vom Restgasdruck und der Restgaszusammensetzung beeinflusst. In Abb. 4.45a. ist der Temperaturkoeffizient von NiCr-Schichten in Abhängigkeit von der Substrattemperatur und der Zusammensetzung zu sehen.

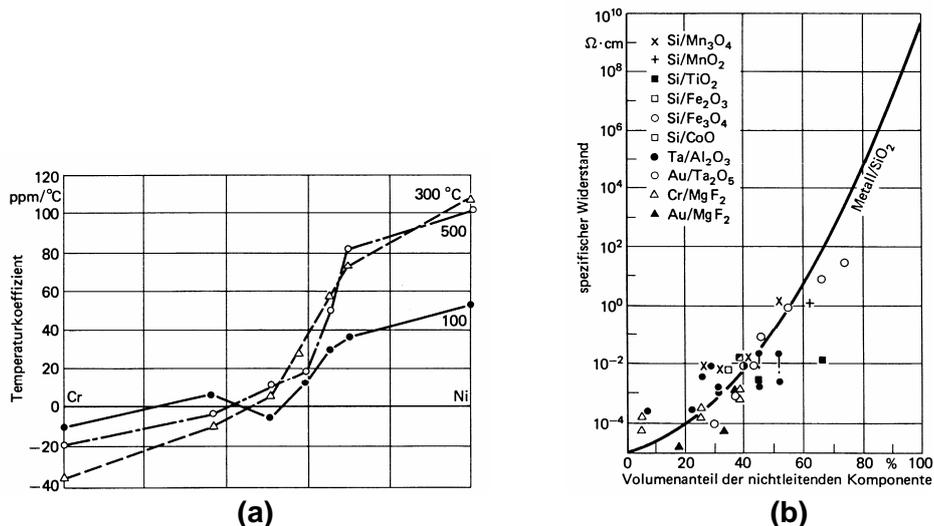


Abb. 4.45.: (a) Temperaturkoeffizient von NiCr-Schichten (PVD) in Abhängigkeit von ihrer Zusammensetzung bei verschiedenen Substrattemperaturen
 (b) Widerstandswerte typischer Cermets
 [3, pp. 406, 409]

4.4.4.3. Cermet-Widerstandsschichten

Der Anwendungsbereich von Dünnschichtwiderständen aus NiCr oder auch anderen Legierungen erstreckt sich bis zu etwa $1000\Omega/\square$. Diese Grenze ergibt sich dadurch, dass bei hohen Widerständen und damit dünnen Schichten die Anforderungen an die Güte der Substratoberflächen sehr hoch sind, die Reproduzierbarkeit der Schichteigenschaften schlechter wird und außerdem der Temperaturkoeffizient für die meisten Anwendungsfälle einen zu hohen Wert annimmt.

Für höhere Widerstandswerte eignen sich Cermets, Mischungen aus Metallen und Metalloxyden oder anderen chemischen Verbindungen. Damit lassen sich Flächenwiderstände von weniger als $1\Omega/\square$ bis zu $1M\Omega/\square$ und niedrigem Temperaturkoeffizienten realisieren. Zahlreiche Materialkombinationen wurden bereits untersucht und angewendet (siehe auch Abb. 4.45 b).

4.4.4.4. Halbleiter-Schichten

In Forschung und Entwicklung wird weltweit derzeit intensiv auf dem Gebiet der Halbleitermaterialien und Halbleitertechnologien gearbeitet, um den Anforderungen an moderne elektronische und optoelektronische Bauelemente gerecht zu werden. Dies gilt insbesondere für die Dünnschichttechnologien, ohne deren Einsatz die moderne Halbleitertechnologie nicht denkbar wäre.

Unter den verwendeten Halbleitermaterialien dominiert dabei Silizium (90% aller Bauteile), dessen Technologie am weitesten entwickelt und erforscht ist. In zunehmendem Maße gewinnen Verbindungshalbleiter (GaAs, AlAs, InP) an Bedeutung für die Elektronik und Optoelektronik.

Als Beschichtungsverfahren kommen hauptsächlich "hochgezüchtete" Methoden zur Anwendung: hochreine CVD-Verfahren; MBE (Molecular Beam Epitaxy); MOCVD (Metallo Organic CVD).

4.4.4.5. Magnetische Schichten

Natürgemäß ist auch Magnetismus eine direkte Folge der elektronischen Struktur eines Materiales. Im Zusammenhang mit der Dünnschichttechnologie sei hier nur erwähnt, dass sich aufgrund des Einsatzes von PVD-Verfahren wie z. B. Magnetron-Sputtern die Speicherdichte von magnetischen Datenträgern (z. B. Festplatten) aufgrund der Herstellung extrem feinkörniger Schichten von einem MByte/dm² (3 1/4 "-Diskette) auf einige 10 GByte/dm² steigern ließ.

4.5. Optische Eigenschaften

4.5.1. Einleitung

Eine der ersten großtechnisch verwertbaren Anwendungen der Dünnschichttechnologie waren die optischen Schichten. Wegen ihrer großen Bedeutung hat man die Erfindung und technische Realisierung der reflexionsmindernden Schichten in der ersten Hälfte des 20. Jahrhunderts als einen der wenigen fundamentalen Fortschritte der praktischen Optik bezeichnet. Die Anwendungsbreite der Interferenzoptik hat sich in den letzten Jahrzehnten durch die Entwicklung neuer Beschichtungsverfahren, durch die Verbesserung der Mess- und Anlagentechnik und auch durch Erschließung neuer Fachgebiete stark vergrößert.

4.5.2. Theorie

4.5.2.1. Fresnelsche Formeln

Für optisch wirksame Schichten gilt die Energiegleichung

$$T + R + A + S = I \quad (4.64)$$

Dabei bedeutet

T = Transmission

R = Reflexion

A = Absorption

S = Streuung

Da die Streuung im allgemeinen vernachlässigbar ist, kann man bei absorptionsfreien Schichten, die bei optischen Anwendungen eine dominierende Rolle spielen, die Transmission berechnen, wenn die Reflexion bekannt ist und umgekehrt. Die folgenden Berechnungen können daher auf eine dieser beiden Größen beschränkt werden.

Eine Aussage über den reflektierten und durchgelassenen Anteil sowie über den Polarisationszustand von einer auf eine Grenzfläche auftreffende Welle, von der die Schwingungsrichtung bekannt ist, erhält man mit den erstmals von A.J. Fresnel abgeleiteten Gleichungen:

$$r_k^p = \frac{n_{k-1} \cos \varphi_k - n_k \cos \varphi_{k-1}}{n_{k-1} \cos \varphi_k + n_k \cos \varphi_{k-1}} \quad (4.65)$$

$$r_k^n = \frac{n_{k-1} \cos \varphi_{k-1} - n_k \cos \varphi_k}{n_{k-1} \cos \varphi_{k-1} + n_k \cos \varphi_k} \quad (4.66)$$

$$t_k^p = \frac{2n_{k-1} \cos \varphi_{k-1}}{n_{k-1} \cos \varphi_k + n_k \cos \varphi_{k-1}} \quad (4.67)$$

$$t_k^n = \frac{2n_{k-1} \cos \varphi_{k-1}}{n_k \cos \varphi_k + n_{k-1} \cos \varphi_{k-1}} \quad (4.68)$$

Dabei sind r und t die Amplitudenvektoren (Fresnel-Koeffizienten) für die Reflexion und Transmission, wobei die p -Vektoren senkrecht zur Zeichnungsebene liegen und die

positive Richtung nach oben zeigt. φ_k sind die Winkel, die der Strahl im k -ten Medium mit der Brechzahl n_k gegen die Einfallsnormale bildet

Für senkrechten Lichteinfall ist $r_k^p = r_k^n$ und $\varphi_k = 0$. Für eine Festkörperoberfläche wird dann

$$r = \frac{n_0 - n_1}{n_0 + n_1} \quad (4.69)$$

und

$$t = \frac{2n_1}{n_0 + n_1} \quad (4.70)$$

wobei n_0 die Brechzahl des Umgebungsmediums und n_1 die des Festkörpers ist. Für den Energiefluß R , der das Verhältnis der reflektierten zur einfallenden Energie angibt, ist bei senkrechtem Lichteinfall

$$R_k^p = R_k^n = (r_k^p)^2 = (r_k^n)^2 \quad (4.71)$$

und

$$R_k = (r_k)^2 = \left(\frac{n_{k-1} - n_k}{n_{k-1} + n_k} \right)^2 \quad (4.72)$$

Für die Transmission ergibt sich bei senkrechtem Lichteinfall unter Berücksichtigung, dass der Lichtstrahl Medien mit unterschiedlichen Brechzahlen durchläuft

$$T_k = \frac{n_{k-1}}{n_k} (t_k)^2 = \frac{4n_{k-1}n_k}{(n_{k-1} + n_k)^2} \quad (4.73)$$

Obige Zusammenhänge gelten nur für absorptionsfreie Materialien und wenn die Streuung vernachlässigbar ist.

4.5.3. Reflexionsmindernde Schichten

4.5.3.1. Einfachschicht

Bringt man auf ein absorptionsfreies Substrat mit der Brechzahl n_2 eine planparallele, isotrope, absorptionsfreie dünne Schicht mit der Brechzahl n_1 auf, deren Dicke d_1 in der Größenordnung der eingestrahlten Lichtwellenlänge liegt, und die auf der anderen Seite an ein absorptionsfreies Medium mit der Brechzahl n_0 grenzt, so treten Mehrfachreflexionen an den Grenzflächen n_2/n_1 und n_1/n_0 auf (siehe Abb.4.46.).

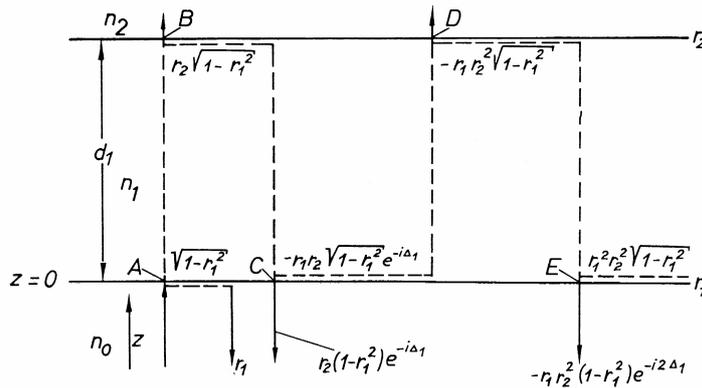


Abb. 4.46.: Zur Ableitung der Reflexion an einer absorptionsfreien, homogenen Einzelschicht bei senkrechtem Lichteinfall [11, p.13]

Für senkrechte Incidenz liefert das Aufsummieren aller Teilstrahlen

$$r = \frac{r_1 + r_2 e^{-i2\delta_1}}{1 + r_1 r_2 e^{-i2\delta_1}} \quad (4.74)$$

und

$$t = \frac{t_1 t_2 e^{-i2\delta_1}}{1 + r_1 r_2 e^{-i2\delta_1}} \quad (4.75)$$

wobei r_1, r_2 , (siehe Gleichungen 4.81 und 4.82!) t_1 und t_2 die Fresnel-Koeffizienten für die beiden Grenzflächen sind und $\delta_1 = (2\pi/\lambda)n_1 d_1 \cos\varphi_1$ die Phasendicke der Schicht ist. n_1 ist deren Brechzahl, d_1 die Schichtdicke, λ die Wellenlänge des eingestrahlten Lichts und φ_1 der Winkel, den der Strahl in der Schicht mit dem Einfallslot bildet.

Weiter ergibt sich dann, unter Berücksichtigung von Gln. (4.72), (4.73) sowie der Beziehung $T_k + R_k = 1$,

$$R = \frac{r_1^2 + 2r_1 r_2 \cos 2\delta_1 + r_2^2}{1 + 2r_1 r_2 \cos 2\delta_1 + r_1^2 r_2^2} \quad (4.76)$$

und

$$T = \frac{n_0}{n_2} \frac{t_1^2 t_2^2}{1 + 2r_1 r_2 \cos 2\delta_1 + r_1^2 r_2^2} \quad (4.77)$$

Nach Gleichung (4.74) verschwindet die Reflexion entweder bei unendlich großem Nenner - was physikalisch nicht möglich ist - oder dann, wenn der Zähler Null ist. In diesem Fall müssen sowohl Real- als auch Imaginärteil Null sein.

Wegen

$$e^{+/-i2\delta_1} = \cos 2\delta_1 + /- i \sin 2\delta_1 \quad (4.78)$$

muss

$$r_1 + r_2 \cos 2\delta_1 = 0 \quad (4.79)$$

und

$$r_2 \sin 2\delta_1 = 0 \quad (4.80)$$

erfüllt sein, damit $R = 0$. Mit

$$r_1 = \frac{n_0 - n_1}{n_0 + n_1} \quad (4.81)$$

und

$$r_2 = \frac{n_1 - n_2}{n_1 + n_2} \quad (4.82)$$

sind die Gleichungen (4.79) und (4.80) dann erfüllt, wenn $2\delta_1$ ein geradzahliges Vielfaches von π ist. Diese Lösung mit $n_0 = n_2$ hat in der Praxis Bedeutung für „unsichtbare“ Schutzschichten („ $\lambda/2$ -Schichten“). Bei der zweiten Lösung ($n_0 \neq n_2$) ist $2\delta_1$ ein ungeradzahliges Vielfaches von π . Es ergibt sich dann aus den Gleichung (4.79) bis (4.82)

$$n_1 = \sqrt{n_0 n_2} \quad (4.83)$$

mit

$$\delta_1 = (2m - 1) \frac{\pi}{2} \quad \text{mit } m=1,2,3,\dots \quad (4.84)$$

Um also durch eine Einzelschicht mit der Brechzahl n_1 auf einem Substrat mit der Brechzahl n_2 eine Nullstelle der Reflexion zu erhalten, muß sowohl die Amplitudenbedingung (4.83), als auch die Phasenbedingung (4.84) erfüllt sein.

Bei senkrechtem Lichteinfall ergeben sich durch Einsetzen von (4.81) und (4.82) in (4.76) und (4.77) für

$$R_{\max} = \left(\frac{n_0 - n_2}{n_0 + n_2} \right)^2 \quad \text{für} \quad \lambda_{\max} = \frac{2n_1 d_1}{m} \quad (4.85)$$

und

$$R_{\min} = \left(\frac{n_1^2 - n_0 n_2}{n_1^2 + n_0 n_2} \right)^2 \quad \text{für} \quad \lambda_{\min} = \frac{4n_1 d_1}{2m - 1} \quad (4.86)$$

In Abbildung 4.47. ist für senkrechten Lichteinfall der Reflexionsgrad R von Einzelschichten mit der Brechzahl n_1 dargestellt, die sich auf einer Glasunterlage mit der Brechzahl $n_2 = 1,52$ befinden und auf der anderen Seite an Luft ($n_0 = 1$) grenzen.

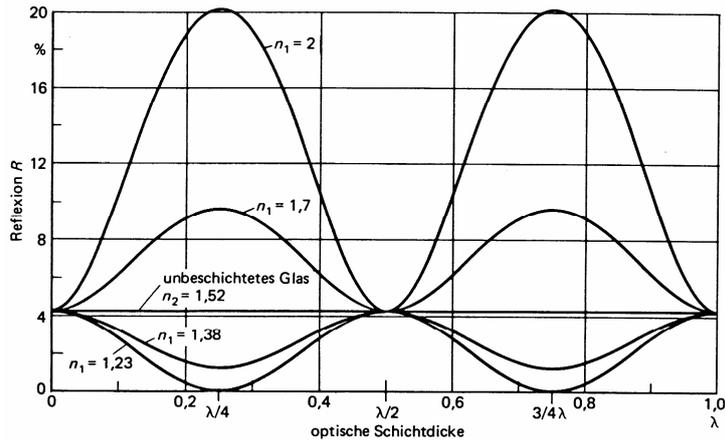


Abb. 4.47.: Reflexion an absorptionsfreien Einzelschichten mit verschiedenen Brechzahlen n_1 auf Glassubstraten ($n_2 = 1,52$) [3, p. 459]

Für $n_1 = n_2$ ergibt sich keine Änderung von R , $n_1 < n_2$ bedingt eine Verminderung und $n_1 > n_2$ eine Erhöhung der Reflexion. Eine Nullstelle der Reflexion erhält man nur dann, wenn Amplituden- und Phasenbedingung (4.83) und (4.84) gleichzeitig erfüllt sind. Wenn nur die Phasenbedingung erfüllt ist, ergibt sich ein Reflexionsminimum, das dem Nullwert um so näher kommt, je kleiner $(n_1^2 - n_0 n_2)$ ist.

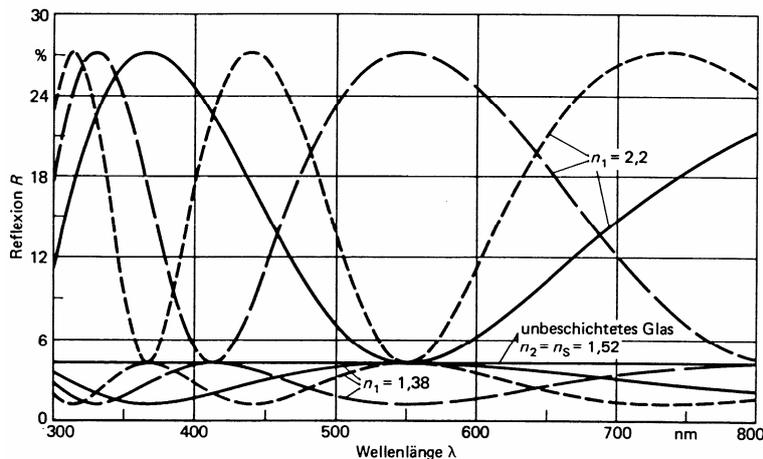


Abb. 4.48.: Reflexion an absorptionsfreien Einzelschichten mit verschiedenen Brechzahlen und verschiedenen Dicken auf Glas in Abhängigkeit von der Lichtwellenlänge; Bezugswellenlänge: 550nm [3, p. 459]

Bei nicht senkrechtem Lichteinfall sind r^p und r^s verschieden. Bei der Berechnung der Reflexion ist der Winkel φ_1 zu berücksichtigen. Die Nullstellen der Reflexion treten dann bei $n_1 d_1 \cos \varphi_1 = (2m - 1)\lambda/4$ auf. Die Schichtdicke muss bei schrägem Lichteinfall um $1/\cos \varphi_1$ größer sein als bei senkrechtem Lichteinfall. Bei schrägem Lichteinfall verschieben sich also die Maxima oder Minima der Reflexion mit zunehmendem Neigungswinkel zu kürzeren Wellenlängen (siehe Abb. 4.49.).

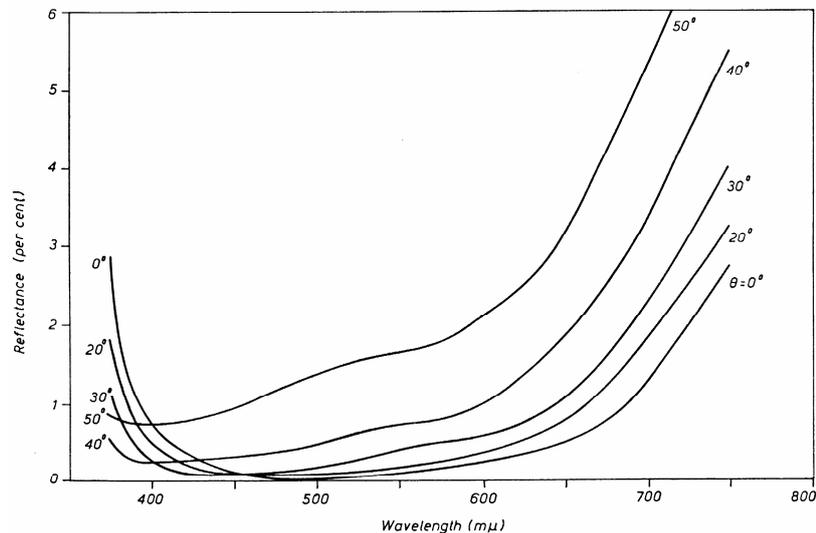


Abb. 4.49.: Reflexion eines reflexionsmindernden Schichtsystems bei verschiedenen Einfallswinkeln [12, p. 57]

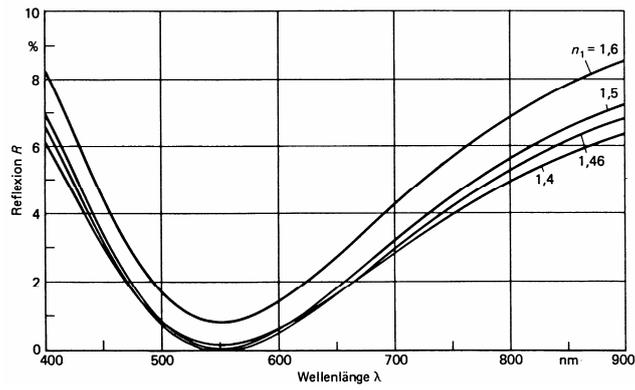
Man sieht aus den obigen Betrachtungen, dass es unmöglich ist, vollständige Transmission über einen breiten Wellenlängenbereich für verschiedene Einfallswinkel zu bekommen. Trotzdem gelingt es - mit Hilfe moderner Berechnungs-, Beschichtungs- und Meßmethoden - optische Systeme mit hohen Transmissionswerten zu erzeugen. Dies ist jedoch nur mit Vielschichtsystemen möglich.

4.5.3.2. Doppelschicht

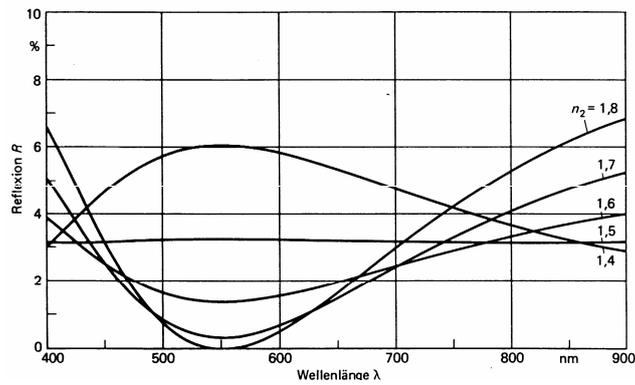
Nullstellen der Reflexion lassen sich mit Einzelschichten nur bei Substraten realisieren, deren Brechzahl größer ist als die des Schichtmaterials. Von den zur Verfügung stehenden Beschichtungsmaterialien, die hinsichtlich Beständigkeit gegenüber Umwelteinflüssen und mechanischen Einwirkungen den heute gestellten Anforderungen genügen, hat MgF_2 mit 1,38 die kleinste Brechzahl. Nach (4.83) lässt sich, wenn gleichzeitig die Phasenbedingung erfüllt ist (4.84), ein Verschwinden der Reflexion bei einer bestimmten Wellenlänge nur bei einem Substratmaterial mit der Brechzahl $n_s = 1,9$ erreichen.

Die üblicherweise in der Optik verwendeten Gläser haben Brechzahlen, die mit 1,5 bis 1,7 weit unter 1,9 liegen. Nullreflexion ist daher mit einer MgF_2 -Schicht nicht realisierbar. Mit Doppelschichten und insbesondere mit Mehrschichten kann auch bei Substraten mit kleiner Brechzahl Nullreflexion erreicht werden. Bei deren Berechnung ist die gleiche Methode anwendbar wie bei der Herleitung der Reflexions- und Transmissionsgleichungen von Einzelschichten.

In Abb. 4.50. ist die Abhängigkeit der Reflexion von $\lambda/4$ -Doppelschichten auf Glas ($n=1,52$) von der Wellenlänge bei verschiedenen Brechungsindizes zu sehen.



a



b

Abb. 4.50.: Abhängigkeit der Reflexion einer $\lambda/4$ -Doppelschicht von der Wellenlänge bei variabler Brechzahl von **a** n_1 ($n_2=1,8$, $n_3=1,52$) bzw. **b** n_2 ($n_1=1,46$, $n_3=1,52$) [3, p. 462]

In der Praxis hat man meist keine Beschichtungsmaterialien mit der rechnerisch ermittelten Brechzahl zur Verfügung, um die Amplitudenbedingung exakt zu erfüllen und damit eine Nullstelle der Reflexion zu erreichen. Um bei einem vorgegebenen Anwendungsfall eine möglichst gute Annäherung des tatsächlichen Reflexionsvermögens an den gewünschten Wert zu erreichen, werden die Schichtdicken oft abweichend von $\lambda/4$ gewählt. So besteht die Möglichkeit, sowohl die Höhe des Reflexionsminimums als auch die Breite des Bereiches mit niedrigem Reflexionsvermögen zu ändern.

4.5.3.3. Drei- und Mehrfachschichten

Im Vergleich zu Doppelschichten ist die Reflexionsminderung in einem breiteren Wellenlängenbereich ungleich größer (siehe Abb. 4.51.). Allein schon daher machen sich geringe Abweichungen von der Amplituden- und Phasenbedingung durch Änderung der Reflexionsfarbe - auch mit freiem Auge - deutlich bemerkbar.

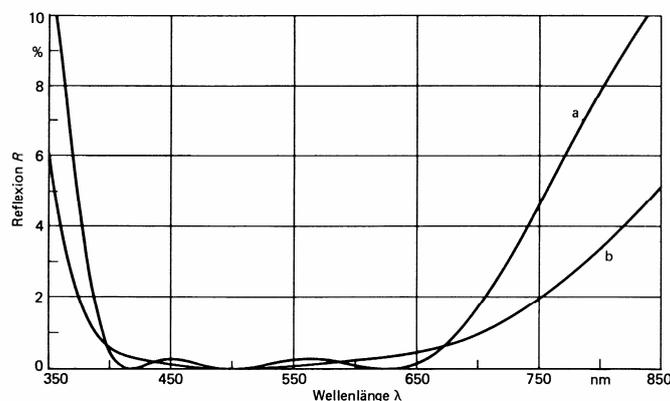


Abb. 4.51.: Reflexionsminderung von Glas durch Dreifachschichten unterschiedlicher Dickenkombination: **a** $\lambda/4 - \lambda/2 - 3\lambda/4$; **b** $\lambda/4 - \lambda/2 - \lambda/4$
 $n_0 = 1$; $n_1 = 1,38$; $n_2 = 2,1$; $n_3 = 1,7$; $n_4 = 1,52$ [3, p. 464]

Noch höhere Transmissionsgrade können mit Vielschichtsystemen erreicht werden. Ohne diese Beschichtungen wären moderne optische Systeme mit bis zu 100 Glas-Luft-Flächen undenkbar. Begriffe wie "Fully Coated", "Multi Coated" usw. sind an nahezu allen optischen Bauteilen zu finden.

4.5.4. Reflexionserhöhende Schichten

4.5.4.1. Metallspiegel

Für hochreflektierende Schichten kommen im einfachsten Fall Metallspiegel in Frage, deren Reflexionsvermögen durch zusätzliche absorptionsfreie Interferenzschichten, die oft gleichzeitig auch als Schutzschichten dienen, erhöht werden kann (siehe Abb. 4.53.). Das Reflexionsvermögen von metallischen Schichten ist nicht nur vom Material, sondern auch vom Beschichtungsverfahren, von den Parametern und vom Substrat abhängig, da eine Voraussetzung für höchstmögliche Reflexion eine Oberfläche mit geringer Rauigkeit ist.

Für Spiegelschichten wird am häufigsten Aluminium benutzt, das vom UV- bis zum IR-Bereich relativ gute Reflexionseigenschaften besitzt (siehe Abb. 4.52.). Bei Lagerung an Luft bildet sich eine dünne Oxidschicht aus, die die Reflexion reduziert, aber die chemische Beständigkeit verbessert.

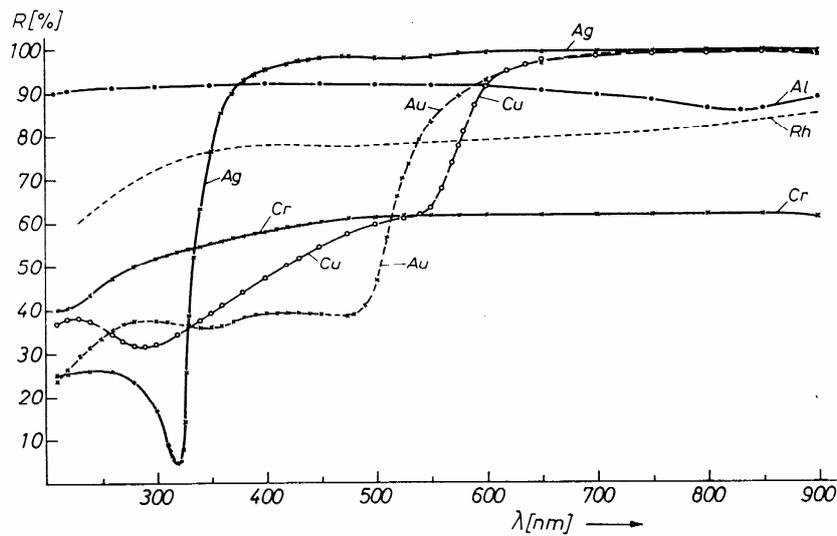


Abb. 4.52.: Reflexionsgrad von frisch aufgetragenen Spiegelschichten [11, p. 93]

Silberspiegel haben eine höhere Reflexion als Aluminiumspiegel. Sie sind jedoch mit dem Nachteil behaftet, dass sie sich schon nach kurzer Zeit wegen der Sulfidbildung verfärben und stark an Reflexionsvermögen einbüßen. Gold und Kupfer haben ein sehr hohes Reflexionsvermögen im Infraroten, während die optischen Eigenschaften im sichtbaren Bereich stark wellenlängenabhängig sind.

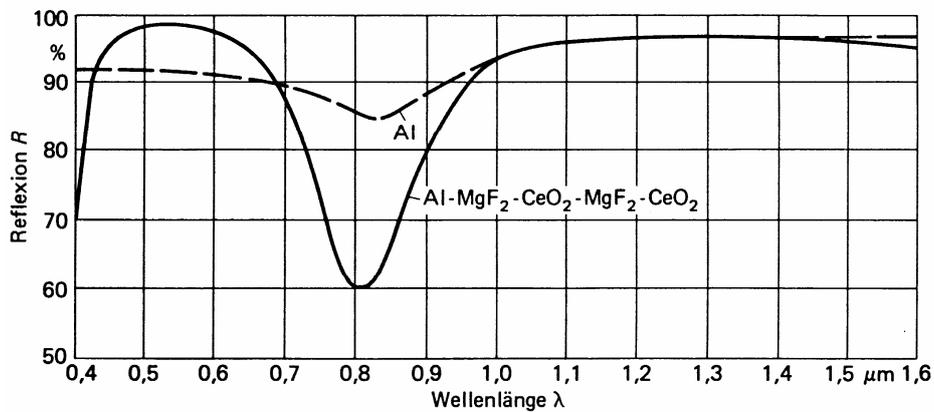


Abb. 4.53.: Reflexion einer aufgedampften Al-Schicht mit und ohne reflexionserhöhenden Schichten [3, p. 476]

4.5.4.2. Vielschichtsysteme

Eine weitere Möglichkeit zur Reflexionserhöhung bieten Vielschichtsysteme: Nach (4.84) tritt dann eine Reflexionserhöhung auf, wenn auf ein Substrat ein Material aufgebracht wird, dessen Brechzahl größer ist als die des Substrates. Die Reflexionserhöhung ist bei einem bestimmten Substratmaterial um so größer, je größer die Brechzahl der Schicht ist. Weiters geht aus (4.84) hervor, dass bei gegebener Brechzahl die Reflexion ein Maximum erreicht, wenn die optische Dicke ein ungeradzahliges Vielfaches von $\lambda/4$ ist.

Eine Verstärkung dieses Effektes kann erreicht werden, indem alternierend $\lambda/4$ -Schichten aus niedriger- bzw. höherbrechenden Materialien aufgebracht werden. Maximale Reflexion wird dann erreicht, wenn die Anzahl der Schicht ungerade ist und sich am Anfang und am Ende des Pakets eine Schicht mit höherbrechenden Material befindet.

In der Abbildung. 4.54. sind für zwei verschiedene Materialkombinationen die Reflexion für verschiedene Schichtzahlen berechnet, wobei Streuung und Absorption vernachlässigt wurden.

Anzahl der Schichten	Reflexion in %	
	$n_L = 1,38$ $n_H = 2,3$ $n_s = 1,51$	$n_L = 1,47$ $n_H = 2,3$ $n_s = 1,51$
3	53,89	53,23
5	85,20	80,84
7	94,67	92,15
9	98,08	96,79
11	99,31	98,68
13	99,75	99,46
15	99,91	99,78
17	99,97	99,91
19	99,99	99,63

Abb. 4.54.: Reflexion von Vielschichtsystemen; H bedeutet hochbrechend ("High"), L niedrig brechend ("Low") [3, p. 467]

Die Anwendung von hochreflektierenden Schichtsystemen liegt einerseits bei Laserspiegeln und andererseits bei den sogenannten Kaltlichtspiegeln. Laserspiegel mit 19 und mehr Schichten höchster Qualität weisen Verluste unter 0,01% auf. Kaltlichtspiegel weisen im sichtbaren Bereich des Spektrums hohe Reflexionswerte und im IR-Bereich hohe Transmissionswerte auf. So wird hohe Lichtausbeute bei gleichzeitiger geringer Erwärmung erreicht (z.B. für Projektionsapparate). Auch für Infraroptiken und Röntgenspiegel werden üblicherweise Vielschichtsysteme als Reflektorelemente eingesetzt, da Einzelschichten in diesen Spektralbereichen zu geringe Reflektivitäten aufweisen.

4.5.5. Spezialfilter

4.5.5.1. Fabry-Perot-Filter

Im einfachsten Fall besteht ein Fabry-Perot-Filter aus zwei teilreflektierenden Metallschichten (meist Ag), zwischen denen sich eine Distanzschicht aus einem absorptionsfreien Material befindet. Von der optischen Dicke der Distanzschicht hängt die Lage der Durchlässigkeitsbande ab. Es wird nur Licht der Wellenlänge

$$\lambda = \frac{2nd}{k} \quad \text{mit } k=1,2,3,\dots \quad (4.87)$$

durchgelassen, wobei nd die optische Schichtdicke ist. Mit k bezeichnet man die sog. Ordnung des Fabry-Perot-Filters.

Aufgrund der Absorption der Metallschichten ergeben sich Nachteile (Nebenmaxima; Schwierigkeiten bei der Herstellung in Bezug auf die Lage des Durchlassbereiches), die durch Verwendung von dielektrischen Vielschichtsystemen vermieden werden können. Die Wirkungsweise ist in beiden Fällen die gleiche, nur werden die metallischen Reflektoren durch hochreflektierende Schichtsysteme ersetzt. An die Distanzschicht schließt sich an beiden Seiten ein reflexionserhöhender Schichtstapel an, der aus einer ungeraden Anzahl von $\lambda/4$ -Schichten besteht, wobei die beiden äußeren Schichten jeweils aus dem höherbrechenden Material bestehen (siehe Abb. 4.55.).

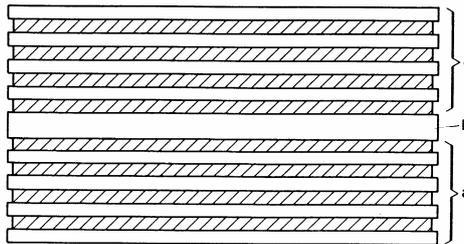


Abb. 4.55.: Aufbau eines Fabry-Perot-Filters aus dielektrischen Schichten [3, p. 480]:
a reflexionserhöhendes $\lambda/4$ -Schichtsystem
b $\lambda/2$ -Distanzschicht

Durch geeignete Wahl der Parameter kann man die Breite, Form und Lage des Durchlassbereiches sowie die Transmission in weiten Bereichen beeinflussen. In der Praxis sind Durchlässigkeiten von mehr als 90% und Halbwertsbreiten unter 0,05nm realisierbar (siehe Abb. 4.56.). Beim Gebrauch von derartigen Filtern ist zu beachten, dass sich bei ungeschützten Systemen der Peak mit der Luftfeuchtigkeit und Temperatur verschiebt.

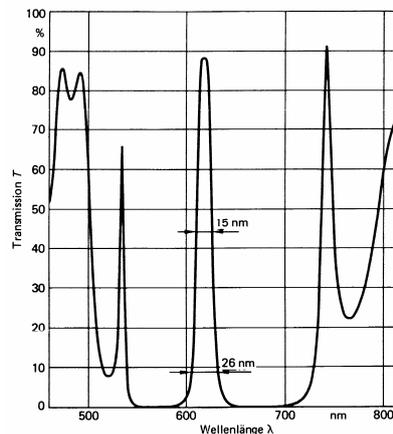


Abb. 4.56.: Transmissionskurve eines Schmalband-Interferenzfilters [3, p. 482]:
 Aufbau: $(HL)_3 H (LH)_3 L (HL)_3 H (LH)_3$
 Mit H bzw. L wird eine hoch- bzw. niedrigbrechende $\lambda/4$ -Schicht bezeichnet.

4.5.5.2. Kantenfilter

Als Kantenfilter bezeichnet man Schichtsysteme, bei denen in einem bestimmten Spektralbereich die Transmission und in einem angrenzenden Spektralbereich die Reflexion sehr groß und die Übergangszone zwischen den beiden Bereichen sehr schmal ist (siehe auch Abb.4.57.). Beispiele sind Wärmeschutzfilter und Block- bzw. Durchlassfilter für bestimmte Spektralbereiche (IR, VIS, UV).

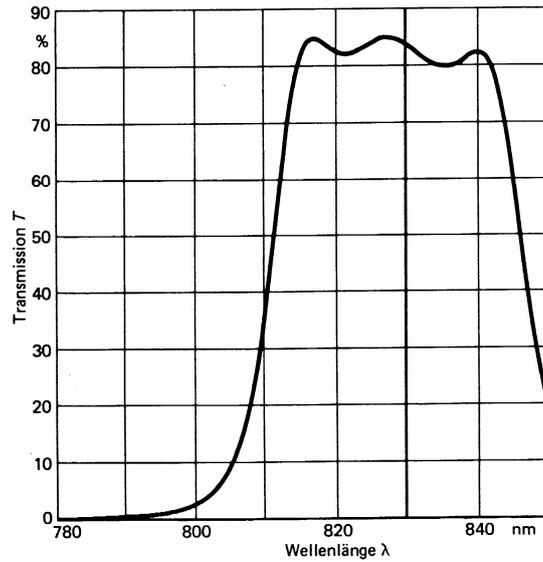


Abb. 4.57.: Transmissionskurve eines Kantenfilters [3, p. 482]:

Aufbau: $(HL)_2 4H (LH)_2 L (HL)_2 4H (LH)_2 4H (LH)_2$

Mit H bzw. L wird eine hoch- bzw. niedrigbrechende $\lambda/4$ -Schicht bezeichnet.

Durch geeignete Wahl der Schichtdicken und deren Aufeinanderfolge, der Schichtmaterialien und der Beschichtungsparameter ist es möglich, nahezu jedes gewünschte Filterdesign (Bandbreite, Kantenlage, Transmission) zu realisieren.

4.5.6. Prüfung optischer Schichten:

4.5.6.1. Allgemeines

Die reproduzierbare Herstellung dünner Schichten für die Optik mit bestimmten Eigenschaften ist mit den heute verfügbaren messtechnischen Hilfsmitteln, der vervollkommenen Anlagentechnik und den Automatisierungsmöglichkeiten sicherer oder mit einem wesentlich höheren Wahrscheinlichkeitsgrad erreichbar als noch vor wenigen Jahren.

Trotzdem ist eine Kontrolle der Fertigprodukte, und wenn es sich auch nur um Stichproben handelt, unerlässlich. Bei optisch wirksamen Schichten nützt das Einhalten der vorgegebenen optischen Werte wenig, wenn nicht auch die mechanischen Eigenschaften, wie z.B. Härte und Haftfestigkeit, oder die chemische Beständigkeit gut genug ist, um eine gebrauchsfähige Schicht garantieren zu können.

4.5.6.2. Prüfverfahren

Die Meßmethoden für Schichtdicke, mechanische Eigenschaften und chemische Zusammensetzung werden in den jeweiligen Kapiteln genauer behandelt. Die Messung der optischen Konstanten erfolgt zumeist mit dem Ellipsometer.

Ellipsometer

Von den vielen möglichen Ellipsometeranordnungen hat für den Einsatz an Beschichtungsanlagen zur Zeit nur die RAE (Rotating Analyser Ellipsometer)-Anordnung praktische Bedeutung (siehe Abb. 4.58.).

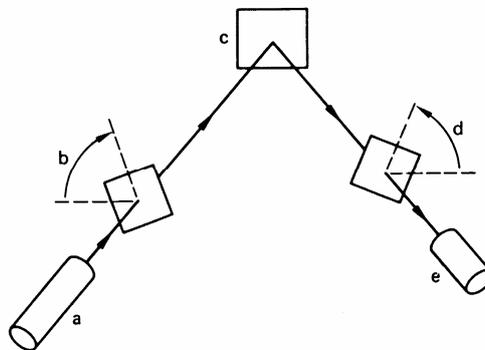


Abb. 4.58.: RAE-Ellipsometer: **a** Lichtquelle; **b** Polarisator; **c** Substrat; **d** Analysator; **e** Photodetektor
[3, p. 394]

Typische Werte für den Winkel zwischen **a** und **e** sind 130 bis 140°. Bezeichnet man mit **P** den Winkel des Polarisators gegenüber der Einfallsebene und mit **A** den entsprechenden Winkel des Analysators, dann ergibt sich als Ausgangssignal des Photodetektors für die ellipsometrischen Größen Ψ und Δ folgender Zusammenhang

$$I(P, A) = C \left[I - \cos 2\Psi (\cos 2A + \cos 2P) + \cos 2A \cos 2P + \sin 2\Psi \cos \Delta \sin 2A \sin 2P \right] \quad (4.88)$$

wobei **C** eine Gerätekonstante, Ψ das Amplitudenverhältnis und Δ die Phasendifferenz ist.

Die Änderung der ellipsometrischen Größen Ψ in Abhängigkeit von Δ für unterschiedliche Schichtmaterialien ist in Abb. 4.59. dargestellt.

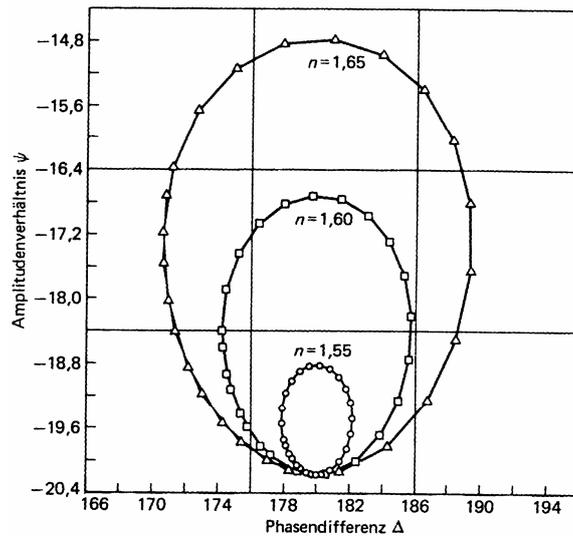


Abb. 4.59.: Ellipsometerwerte Ψ und Δ [3, p.397]

Der große Vorteil der optischen Meßmethoden kommt dann zum Tragen, wenn man an Stelle der indirekten Messung (Testglas) die Substrate mit dem eigentlichen Schichtsystem direkt misst.

Weitere Prüfverfahren

Optische Beschichtungen werden zahlreichen speziellen mechanischen oder chemischen Prüfungen unterzogen, um deren Gebrauchstauglichkeit zu überprüfen. Einige ausgewählte Beispiele sollen genügen:

Feuchtwechselklima: 12 Stunden bei 40°C und 100% Luftfeuchtigkeit, und dann 12 Stunden Abkühlung ergibt einen Zyklus. Die Anzahl der Zyklen wird variiert.

Sprühnebelprüfung: Bei 50°C wird mit einer Düse eine wässrige Lösung mit 5% NaCl aufgesprüht

Sandrieselverfahren: Man läßt 3kg Sand der Kornklasse 0,4 bis 0,8mm aus 1650mm Höhe durch ein Fallrohr auf die Probe rieseln, die sich auf einem Drehteller befindet, der 45° geneigt ist.

Kochen in Salzwasser und deionisiertem Wasser: Der Prüfling verbleibt 5 bis 60 min in der Salz-Lösung

4.6. Chemische Zusammensetzung

4.6.1. Einleitung

Konventionelle chemische Analysemethoden, wie Atomemission, Atomabsorption, Spektralanalyse, Röntgenfluoreszenz und Massenspektrometrie spielen eine große Rolle bei der Herstellung der in den verschiedenen Verfahren benutzten Beschichtungsmaterialien. In den meisten Fällen sind Proben von einigen Milligramm erforderlich, wobei die Nachweisgrenze durchaus im ppm-Bereich liegt.

Andererseits haben die gesteigerten Anforderungen an die Mikroanalyse von Oberflächen und dünnen Schichten hinsichtlich Orts- und Flächenauflösung und Informationstiefe zu einer Vielzahl neuer Methoden geführt. Diese Methoden beruhen darauf, dass man Wechselwirkungsprozesse von Photonen, Elektronen, Ionen oder Teilchen mit der zu analysierenden Schicht untersucht. Dabei werden - je nach Art der Anregung und der zu analysierenden Substanz - Teilchen elastisch oder inelastisch gestreut oder Sekundärteilchen emittiert und diese Teilchen durch ein entsprechendes Detektorsystem analysiert. Eine Zusammenstellung der Methoden, geordnet nach anregenden und nachzuweisenden Teilchen, ist in Abb. 4.60. ersichtlich.

Anregung durch		Nachweis durch				
		Photonen		Elektronen		Ionen
		optisch	Röntgen			
Photonen	optisch	AA UV IR		ESCA	UPS	LIMA
	Röntgen		XRF XRD		XPS	
Elektronen			EPM	SEM TEM STM	AES SAM LEED RHEED	
Ionen		SCANIIR	IIX			SIMS SNMS RBS ISS

Erklärung der Abkürzungen:

AA	Atomic Absorption
AES	Auger Electron Spectroscopy
EPM	Electron Probe Microanalysis
ESCA	Electron Spectroscopy for Chemical Analysis
IIX	Ion Induced X-Rays
IPM	Ion Probe Microanalysis
IR	Infrared Spectroscopy
ISS	Ion Scattering Spectroscopy
LEED	Low Energy Electron Diffraction
LIMA	Laser induced Ion Mass Analyzer
RBS	Rutherford Backscattering Spectroscopy
RHEED	Reflexion High Energy Electron Diffraction
SAM	Scanning Auger Microanalysis
SCANIIR	Surface Composition Analysis by Neutral and Ion Impact Radiation
SEM	Scanning Electron Microscopy
SIMS	Secondary Ion Mass Spectrometry
SNMS	Secondary Neutrals Mass Spectrometry
STM	Scanning Tunnel Microscopy
TEM	Transmission Electron Microscopy
UPS	UV-Photoelectron Spectroscopy
UV	UV-Spectroscopy
XPS	X-Ray Photoelectron Spectroscopy
XRD	X-Ray Diffraction
XRF	X-Ray Fluorescence Spectroscopy

Abb. 4.60.: Methoden zur Bestimmung der chemischen Zusammensetzung und der Struktur von Oberflächen und dünnen Schichten [1, p. 46]

4.6.2. Elektronenstrahl-Mikroanalyse (EPM)

Die Elektronenstrahl-Mikroanalyse (EPM), auch Röntgenmikroanalyse genannt, ist das älteste dieser Verfahren, wobei die Anregung mit Elektronen und der Nachweis durch Röntgenquanten erfolgt. Es wird eingesetzt, wenn Schichten von mindestens $1\mu\text{m}$ Dicke zu analysieren sind. Es existieren zwei prinzipielle Ausführungsformen: die "Mikrosonde" mit wellenlängendispersiven Kristallspektrometer (WDX), welche vielfach mit einem TEM kombiniert wird. Im Raster-Elektronenmikroskop (SEM, REM) zur Anwendung kommt meist das energiedispersive Si(Li)-Halbleiterspektrometer (EDX) zum Einsatz.

Diese Verfahren versagen im Bereich der leichten Elemente ($Z < 8$) und bei Schichten, die dünner sind als die durch die Eindringtiefe des Elektronenstrahls gegebene Informationstiefe (ca. $1\mu\text{m}$).

4.6.3. Auger-Elektronenspektroskopie (AES)

Die Probe wird üblicherweise mit Elektronen von 1 bis 10keV beschossen. Dadurch werden aus den inneren Schalen der Probenatome Elektronen entfernt, die durch Übergänge von Elektronen aus höheren Schalen wieder aufgefüllt werden (siehe Abb. 4.61.). Die dabei frei werdende Energie wird entweder als Röntgenquant mit der Energie $h\nu_F$ abgestrahlt (Röntgenfluoreszenz) oder innerhalb des Atoms auf ein drittes Elektron übertragen (Auger-Effekt), das in Freiheit gesetzt wird.

Die durch die Austrittstiefe der Auger-Elektronen gegebene Informationstiefe beträgt 1 bis 10nm. Zur Analyse wird die Zahl der pro Zeiteinheit emittierten Auger-Elektronen mit einem Sekundärelektronenvervielfacher als Funktion ihrer kinetischen Energie (z.B. mit einem Zylinder-Spiegel-Analysator) gemessen. Die Maxima des so gewonnenen AES-Spektrums geben aufgrund tabelliert vorliegender Werte für die Energieterme Auskunft über die Art der Probenatome und ihre Konzentration. Um den von Streuelektronen herrührenden Untergrund des Spektrums zu unterdrücken, wird dieses mit elektronischen Mitteln differenziert.

Die Ordnungszahl der nachweisbaren Atome ist $Z > 3$ und die Nachweisgrenze der Teilchenanteile beträgt etwa 0,1%. Die Messung von Tiefenprofilen ist nicht zerstörungsfrei möglich, sondern nur durch Abtragen der Schicht mittels Sputtern.

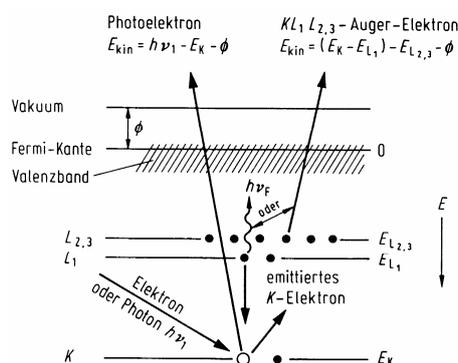


Abb. 4.61.: Termschema für die Auger-Elektronen (AES)- und die Photoelektronen (ESCA)-Spektroskopie. Q Elektronenaustrittsarbeit [1, p. 48]

4.6.4. Photoelektronenspektroskopie (ESCA)

Zur Anregung dient entweder Röntgenstrahlung (dann heißt das Verfahren XPS) oder UV-Licht (dann heißt es UPS). Beim XPS-Verfahren verwendet man z.B. die Aluminium-K-Strahlung mit der Photonenenergie $h\nu_1 = 1486\text{eV}$ (siehe Abb. 4.61.). Die auf das in Freiheit gesetzte Photoelektron übertragene kinetische Energie $E_{kin} = h\nu_1 - E_k - F$ ist gleich $h\nu_1$ vermindert um die für das zu analysierende Element charakteristische Bindungsenergie E_k und die Austrittsarbeit F . Ähnlich wie bei der AES-Analyse wird die Zahl der pro Zeiteinheit emittierten Elektronen in Abhängigkeit von E_{kin} gemessen.

Die Ordnungszahl nachweisbarer Elemente ist ebenfalls $Z > 3$, die Informationstiefe 0,5 bis 10nm und die Nachweisgrenze der Teilchenanteile 0,1%. Ein Vorteil des ESCA-Verfahrens besteht darin, dass außer der analytischen Information eine Aussage über den Bindungszustand der Atome aufgrund der von der Art der Bindung (z.B. metallisch oder oxydisch) abhängigen Linienverschiebung (chemical shift) erhalten werden kann.

4.6.5. Sekundärionen-Massenspektrometrie (SIMS)

Die Probe wird mit positiv geladenen Ionen, meist Cs (Nachweis elektronegativer Elemente) oder O_2 (Nachweis elektropositiver Elemente) 10 keV Energie beschossen (siehe auch Abb 4.62.). Auch Ga, Au- und Bi Ionen kommen zum Einsatz. Diese werden von Quellen geringer Intensität emittiert und erlauben sowohl vertikal als auch lateral Hochauflösendes SIMS. Zwischen den einfallenden Ionen und den Targetatomen kommt es zu elastischen Stoßprozessen, als deren Folge im Target Stoßkaskaden auftreten, von denen einige auch die Oberfläche erreichen und hier zur Emission einzelner Atome oder Molekülbruchstücke führen (Sputter-Effekt).

Von den emittierten, d.h. zerstäubten Teilchen ist ein gewisser Teil positiv oder negativ geladen und somit z.B. in einem Quadrupol-Massenspektrometer nachweisbar. Neuere Anlagen bedienen sich auch der Flugzeit (Time Of Flight)-Massenspektrometrie (TOF-SIMS), welche auch sehr geringe Ionenströme verlässlich nachweisen kann. Das Sekundärionen-Massenspektrum ist charakteristisch für die Verteilung der Elemente an der Probenoberfläche und auch (aufgrund der Molekülbruchstücke) für deren Bindungszustand. Die mittlere Austrittstiefe beträgt einige nm. Da das SIMS-Spektrum vom Prinzip her untergrundfrei ist, besitzt das Verfahren eine sehr niedrige Nachweisgrenze (Teilchenzahlanteile $10^{-4}\%$). Alle Elemente (auch Wasserstoff) nebst Isotopen sind nachweisbar.

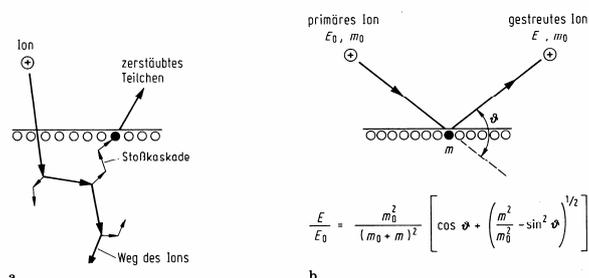


Abb. 4.62.: Prinzipdarstellung der (a) Sekundärionen-Massenspektrometrie (SIMS), der Sekundär-Neutralteilchen-Massenspektrometrie (SNMS) und (b) der Ionenstreuungsspektroskopie (ISS) [1, p. 49]

Während die Verfahren AES und ESCA zerstörungsfrei arbeiten (falls die Probe nicht zusätzlich einem Sputterprozeß ausgesetzt wird), ist das SIMS-Verfahren zerstörender Art. Das bringt aber den Vorteil, dass Tiefenprofile der Atomkonzentration gemessen werden können.

4.6.6. Sekundär-Neutral-Massenspektrometrie (SNMS):

Beim SIMS-Verfahren sind die Vorgänge der Emission und Ionisation der zu analysierenden Teilchen gekoppelt (Matrixeffekt). Daher sind die Intensitäten der einzelnen Linien des Massenspektrums nicht repräsentativ für die Zusammensetzung der untersuchten Oberfläche. Um zu einer quantitativen Analyse zu gelangen, bedarf es einer Kalibrierung mit Proben bekannter Zusammensetzung. Im Interesse der Messgenauigkeit, insbesondere bei der Analyse von Sandwich-Strukturen mit Diffusions- und Implantationsprofilen, ist es jedoch vorzuziehen, die Vorgänge der Emission und Ionisation unabhängig voneinander auszuführen (siehe auch Abb. 4.62.).

Dies geschieht beim SNMS-Verfahren dadurch, dass die durch Ionenbombardement erzeugten neutralen Partikel, die den Hauptteil der Sputteremission bilden, auf geeignete Weise ionisiert und dann im Quadrupol-Massenspektrometer nachgewiesen werden. Die Atomkonzentrationen einer gegebenen Masse ist dann eine Funktion der jeweiligen Sputterausbeute und der Ionisationswahrscheinlichkeit in der verwendeten Apparatur. Als Mittel zur Ionisation der zerstäubten Teilchen hat sich ein HF-Niederdruckplasma (Ar) bewährt. Neuerdings kommen für die Ionisation der abgetragenen Teilchen auch Excimer-Laser zur Anwendung (Husinsky; Inst. f. allg. Physik)

4.6.7. Ionen-Streuspektroskopie (ISS)

Dieses Verfahren ermöglicht die Analyse der obersten Atomlage der Probe, deren Oberfläche mit Ionen bekannter Energie E_0 (einige 100 bis 10^3 eV) und der Masse m_0 beschossen wird (siehe Abb 4.62.). Die an den Oberflächenatomen der Masse m unter einem bestimmten Winkel (z.B. $\Theta = 90^\circ$) elastisch gestreuten Primärionen werden in einem Detektor ihrer Energie nach analysiert. Das so erhaltene ISS-Spektrum besitzt Maxima bei den Werten E / E_0 , die nach der Theorie der elastischen Streuung den Massenverhältnissen m / m_0 entsprechen (siehe Abb. 4.62.).

4.6.8. Rutherford-Rückstreuspektroskopie (RBS)

Die RBS-Analyse ist eine Hochenergieversion der ISS-Methode (siehe Abb. 4.63.). Sie wird mit Ionen von Wasserstoff, Helium oder anderen leichten Elementen im Energiebereich von 0,1 bis 5MeV ausgeführt. Aus den unter Variation der Energie E_0 der Primärionen gemessenen Energiespektren der rückgestreuten Ionen lassen sich Tiefenprofile der Atomkonzentrationen für alle leichten Elemente im Bereich von 0,1 bis $10\mu\text{m}$ zerstörungsfrei bestimmen. Ein besonderer Vorteil der RBS-Methode, die eine Nachweisgrenze von $10^{-3}\%$ besitzt, besteht darin, dass sie eine absolute Analysenmethode ist, während die oben genannten Methoden nur mittels Standardproben kalibriert werden können.

Mit energiereichen Ionen lassen sich, wie in Abb. 4.63. schematisch dargestellt ist, noch zwei weitere, für die Festkörperphysik wichtige Analysenmethoden verwirklichen: Die Methode der Ioneninduzierten Röntgenstrahlung (IIX) und die Methode der durch Ionen ausgelösten Kernreaktionen (NRA)

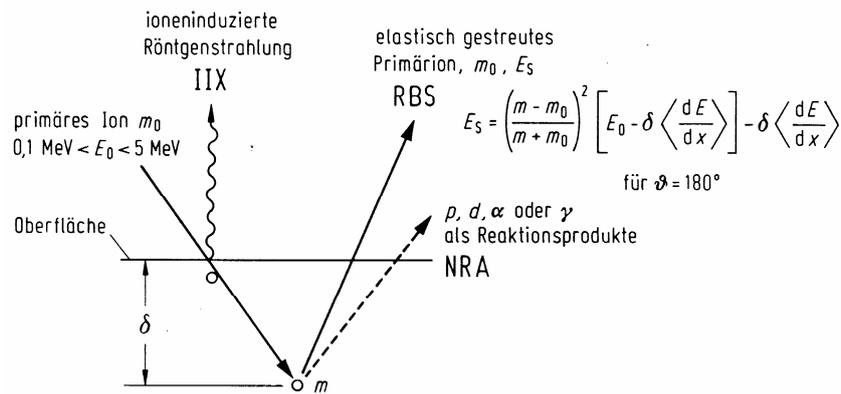


Abb. 4.63.: Schematische Darstellung der Wechselwirkungen von energiereichen leichten Ionen mit Festkörpern und deren Anwendung zur Analyse [1, p. 51]:

- a:** Ioneninduzierte Röntgenstrahlung (Ion induced X-rays: IIX)
- b:** Elastische Rückstreuung der Primärionen (Rutherford Backscattering: RBS)
- c:** Kernreaktionen unter Emission charakteristischer Reaktionsprodukte, wie beispielsweise Protonen, Deuteronen, γ -Quanten oder α -Teilchen (nuclear reaction analysis: NRA)