

61

Vorlesungsmanskripte des
21. IFF-Ferienkurses
vom 5. bis 16. März 1990
im Forschungszentrum Jülich

Festkörperforschung für die Informationstechnik

Dieser Kurs wurde vom Institut für Festkörperforschung
in Zusammenarbeit mit dem Institut für Schicht- und
Ionen-technik des Forschungszentrums Jülich sowie
der RWTH Aachen und den Universitäten Düsseldorf
und Köln veranstaltet.

Herausgegeben vom Forschungszentrum Jülich GmbH
Institut für Festkörperforschung
Postfach 1913 · D-5170 Jülich
Telefon (02461) 616048 · Telex 833556 kf d
Zusammenstellung R. Holzle, W. Schmitz
Satz und Lithos KFA
Druck WEKA-Druck GmbH, Linnich
Copyright 1990 KFA Jülich

Vertrieb KFA Jülich GmbH, Zentralbibliothek
Postfach 1913 · D-5170 Jülich
Telefon 02461/61-5367 · Telefax: 02461/61-4646

Forschungszentrum Jülich GmbH

16. Industrielle Entwicklung und Fertigung von Halbleiterbauelementen

H. Föll und F.S. Becker

SIEMENS, Bereich Halbleiter

Teil 1 Technologieentwicklung im wirtschaftlichem Umfeld der Halbleiterspeicher

16.1 Einleitung

16.1.1 Allgemeiner Hintergrund zur Fertigung von Halbleiterbauelementen

In den letzten 20 Jahren hat die Halbleiterentwicklung einen früher für unvorstellbar gehaltenen Aufschwung erlebt. Trotz intensiver Erforschung von Alternativen wie z.B. GaAs hat Silizium dabei seine dominierende Stellung nicht eingebüßt, da es nicht nur in Form von Quarzsand als Rohmaterial reichlich zur Verfügung steht und gute elektrische Eigenschaften aufweist, sondern im Gegensatz zu seinen Konkurrenten auch ein Oxid von hoher elektrischer und mechanischer Qualität bildet. Durch die Einführung der auf einem Siliziumkristall integrierten Schaltung entfiel der mühsame Zusammenbau einzelner Bauelemente wie Transistoren, Widerstände oder Kondensatoren und die so erzielbare hohe Integrationsdichte führte zu einer erheblichen Reduktion des Platzbedarfs und der Kosten. Die Verfügbarkeit dieser kompakten und billigen Bauelemente ermöglichte völlig neue Anwendungen, die dann ihrerseits wieder zu erhöhter Nachfrage und dementsprechend günstiger Massenproduktion führten. Die Speerspitze des technologischen Fortschritts bilden dabei die Speicherbausteine; sie besitzen die höchste Integrationsdichte der jeweiligen Technologiegeneration.

Im Sommer des Jahres 1989 wurde weltweit die Produktion des 4Megabit DRAMs aufgenommen (DRAM = dynamic random access memory; zu deutsch = dynamischer, wahlfreier Speicher /1/); ein Ereignis das in Deutschland sogar vom Bundeskanzler persönlich gefeiert wurde /2/. Weltweit heißt in diesem Falle, daß 9 japanische Firmen die Produktion aufgenommen haben oder im Begriff waren dies demnächst zu tun, daß einige Firmen in Südkorea und Taiwan heftig daran arbeiten und daß ganze drei nichtasiatische Firmen (IBM, SIEMENS und Texas Instruments) sich aus eigener Kraft an dieser speziellen Schlacht um die Beherrschung des Weltmarkts für Elektronik beteiligen (Tab. 1).

Firma	Chipfl. (mm ²)	Zellkonden- satorotyp	Zellgröße (µm ²)	Verdrahtungsebenen		Litho- graphie
				Poly-Si	AL	
Siemens	91	Graben	10.6	3	1	g-Linie
IBM	86	Graben	10.6	1	2	i-Linie
Texas Ins.	85	Graben	8.9	2	2	g-Linie
Fujitsu	64	gestapelt	7.6	4	1	g-Linie
Matsushita	84.3	Graben	10.6	3	1	g-Linie
Mitsubishi	72	gestapelt	13.3	3	1	g-Linie
NEC	99	Graben	10.6	3	1	g-Linie
Hitachi	89.9	gestapelt	11.4	3	2	g-Linie
Toshiba	96.2	Graben	11.6	3	1	g-Linie
Sony	119	Graben	16.2	2	2	?
Oki	99.3	gestapelt	10.5	3	1	i-Linie
Sharp	73.0	Graben	8.74	2	2	g-Linie

Tabelle 1: Hersteller von 4M-DRAMs; IBM ist am Weltmarkt allerdings nicht vertreten. Die Bedeutung der technischen Angaben wird in Kapitel 16.2 bis 16.4 behandelt; Quelle /7/.

Was versteckt sich hinter der einfachen Aussage: "Die Produktion aufnehmen"? Offenbar braucht man eine Fabrik (eine sogenannte Waferfab /3/) in die vorne Siliciumscheiben (Wafer) einlaufen, und aus der hinten verkaufsfähige Chips herauskommen. Damit das so einfach funktioniert muß noch einiges zusammenkommen, man braucht nämlich noch:

- Menschen mit sehr verschiedenem und teilweise hochspezialisiertem Wissen, die Maschinen bedienen, Maschinen am Laufen halten, messen und kontrollieren, Abläufe steuern, Prozesse verbessern, Chips verkaufen, planen, verwalten, entscheiden.
- Chemikalien, Photomasken, Gehäusematerialien, Ersatzteile.
- Das Wissen wie's gemacht wird in Form von Arbeitsanweisungen, Computerprogrammen und der nicht konkret greifbaren "Fertigungskultur".
- Am allerwichtigsten: Viel Geld !

Dies ist, grob schematisch, der "Input" des Systems Waferfab. Der "Output" besteht auch nicht nur aus fertigen (d.h. montierten, getesteten und verpackten) Chips, sondern auch aus:

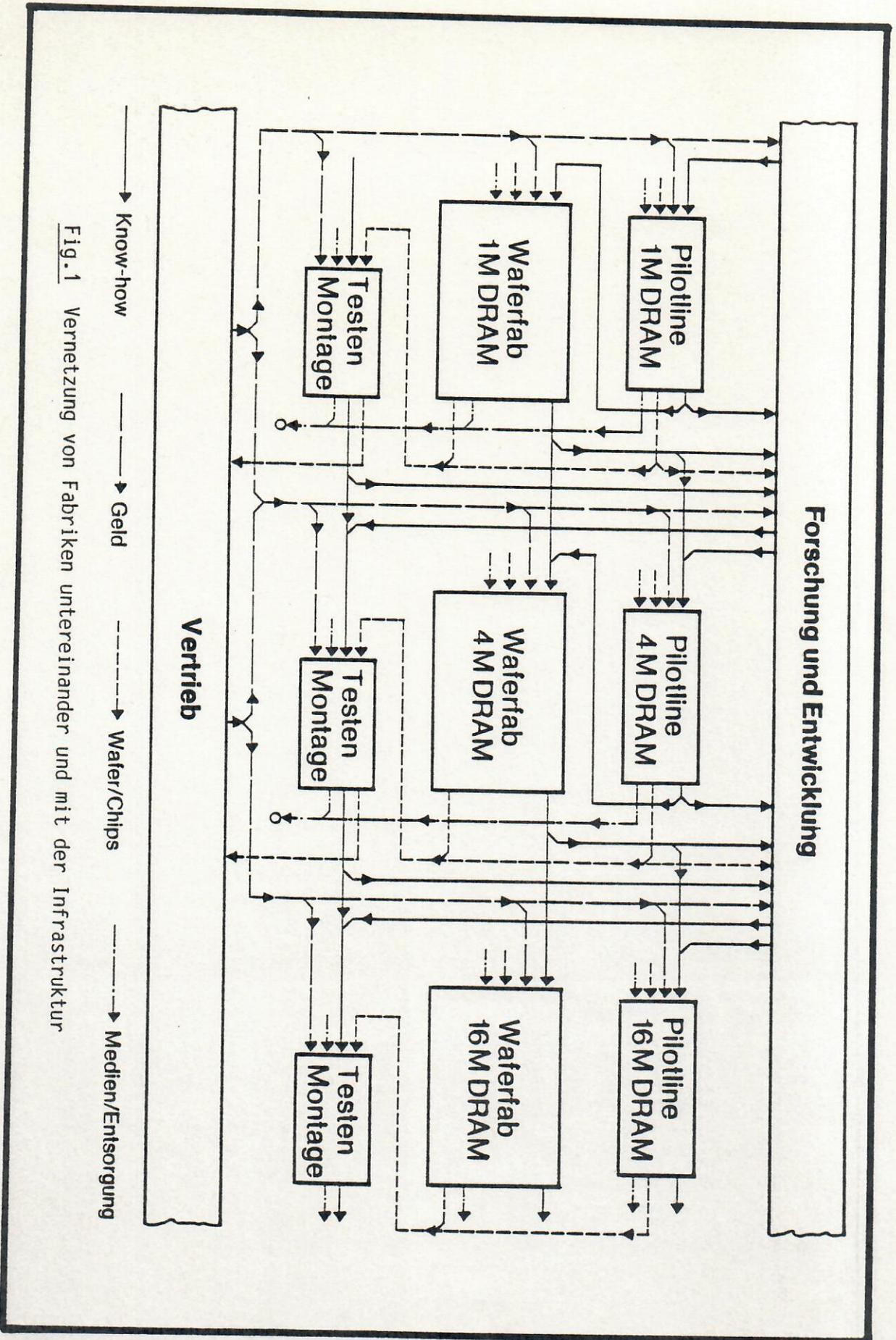


Fig. 1 Vernetzung von Fabriken untereinander und mit der Infrastruktur

- Geld, entstanden durch den Verkauf der fertigen Chips durch den Vertrieb; manchmal auch aus dem Verkauf des "know-hows" in Form einer Lizenz.
- Nicht funktionierenden Chips. Das Verhältnis zwischen den vorne einlaufenden potentiell möglichen Chips (ca. 150 4M DRAM Chips pro 150 mm Si-Scheibe) und den verkauften Chips, die herauskommen, heißt Ausbeute γ .
- Abfall in Form von Gasen, Flüssigkeiten und festen Körpern. Know-how über die Effizienz der Prozesse und Verfahrenstechniken sowie über die zu stellenden Ansprüche an Personal, Hard- und Software. Hierbei spielt die Analyse der defekten Chips und die Optimierung der Materialflüsse eine wichtige Rolle.

Ein Teil des Outputs ist rückgekoppelt zum Input. Dies gilt in geringem Umfang für die Medien und Materialien (z.B. Luftumwälzung im Reinarum; Recycling von Si-Scheiben), in großem Umfang aber für die Outputs Geld und Information. Eine wirtschaftlich sinnvolle, und damit in einer funktionierenden Marktwirtschaft einzig denkbare Fabrik, muß letztendlich mehr Geld verdienen als sie verbraucht (Anlauf- und Nebenkosten eingeschlossen). Die Rückkopplung der Geld- und Informationsflüsse ist aber nicht direkt, sondern läuft beim Geld zeitlich verschoben, bei der Information auch über Zwischenstufen und Umwege.

Zeitliche Verschiebung heißt, daß zu der Zeit des größten Geldbedarfs (also während der Aufbauphase), hinten noch kein Geld herauskommt. Wenn die Fabrik dann voll läuft, wird es (hoffentlich) umgekehrt sein.

Indirekte Rückkopplung heißt, daß die Erzeugung des grundsätzlichen know-hows für das neue Produkt weitgehend außerhalb der Fabrik erfolgt (in Forschungslabors, Designzentren und Entwicklungsbereichen), aber durch das in der Vorgängertabrik angesammelte Know-how gesteuert wird. Vor Anlauf der Fabrik muß es abgeschlossen sein; es wird (zeitverschoben) durch den Output der Fabrik bezahlt.

Unsere Waterfab ist also kein geschlossenes System, sondern eingebettet in eine komplexe Infrastruktur die sich darüberhinaus zeitlich rasch verändert. Das Gesamtsystem in das die Waterfab eingegliedert ist, besteht mindestens aus:

- Den Vorgängertabrik, die während der Aufbauphase das notwendige Geld erzeugen müssen.
- Der Nachfolgefabrik, die von unserer Waterfab mitfinanziert werden muß.
- Den Forschungs- und Entwicklungseinrichtungen (F+E), die das zu fertigende Produkt entwickeln und eine F+E- oder Pilot-Linie betreiben.
- Der Peripherie (Zulieferer, Vertrieb, Verwaltung, Kantine, Qualitätskontrolle, Betriebsrat, Umweltschutz- und Sicherheitsbeauftragte, ...).

Die Beziehungen zwischen diesen Komplexen sind in Fig. 1 schematisch dargestellt. In der Realität muß nicht jeder

"Kasten" ein individuelles Gebäude sein sondern kann auch durch Ausbau einer schon vorhandenen Struktur entstehen; dies ändert jedoch nichts an der Systematik.

Alle bisherigen Ausführungen gelten in schöner Allgemeinheit für alle Fabriken - nicht nur für die Produzenten von Siliziumchips sondern auch für Kartoffelchiphersteller. Im Gegensatz zu allen anderen Industriezweigen erfolgt die Si-Chipherstellung jedoch unter einigen Randbedingungen, die in ihrer Kombination einzigartig sind. Als Konsequenz dieser speziellen Situation beobachten wir derzeit einen - durchaus heißen - "Krieg" um die Weltherrschaft auf wirtschaftlichem Gebiet, der den heutzutage doch stark abgeflauten kalten Krieg um die Weltherrschaft der Ideologien als Episode erscheinen lassen könnte. Die Frage, ob es sich Europa im allgemeinen und die BRD im Besonderen leisten kann, dabei im wesentlichen tatsächlich nur als Beobachter mitzutun, soll hier nicht näher verfolgt werden (siehe dazu /4/).

16.1.2 Die speziellen Randbedingungen der Halbleiterindustrie

Was sind die besonderen Randbedingungen der Chipherstellung? Allgemein gültige Aussagen (insbesondere für Speicher-Chips, aber auch für andere Chips) sind:

1. Ungebrochenes exponentielles Wachstum des Chipmarktes mit mittleren jährlichen Wachstumsraten von ca. 15% (in DM; Fig. 2). Bezogen auf die Stückzahlen von Speichern liegt die Wachstumsrate eher höher, weil zum einen Speicher einen immer größeren Anteil am gesamten Chipmarkt bilden, und zum anderen die Stückzahlen, bedingt durch den Preisverfall, schneller wachsen als das zugehörige Geldvolumen

Unmittelbare Konsequenzen daraus sind: Die Fertigungskapazitäten müssen mindestens genauso schnell wachsen. Da die Chips einer neuen Speichergeneration immer größer werden, (Fig. 3) heißt dies größere Wafer und/oder mehr Wafer pro Zeiteinheit und/oder längere Arbeitszeiten - also größere Fabriken mit teurerem Equipment und mehr Personalbedarf. Alles in allem wird die neue Fabrik ungefähr doppelt so viel kosten wie die Vorgängerin.

2. Neue Generationen mit vierfacher Leistung oder Komplexität erscheinen alle 3 Jahre (Fig. 4); ein seit 17 Jahren ungebrochener Trend, der sich aller Wahrscheinlichkeit nach noch 5-10 Jahre lang fortsetzen dürfte.

Unmittelbare Konsequenzen daraus sind: Alle 3 Jahre wird eine neue Fabrik benötigt; diese muß gebaut werden, bevor die Forschung und Entwicklung des neuen Produkts an einem Punkt angelangt ist bei dem die rechtzeitige Verfügbarkeit des Know-hows zur Produktion des neuen Produktes halbwegs sichergestellt ist. Darüber hinaus heißt dies, daß sowohl die Zeitdauer für die F+E Phase des neuen Produkts als auch die Herstellungszeit für die Test- und Produktchips (die sogenannte Durchlaufzeit (DLZ)) sich nicht erhöhen darf. Dies führt in letzter Konsequenz unvermeidlich zur Kontischicht mit Samstag/Sonntag Arbeit und zu stark steigenden F+E Kosten.

Handelsvolumen Halbleiter (\$ x 10⁹)

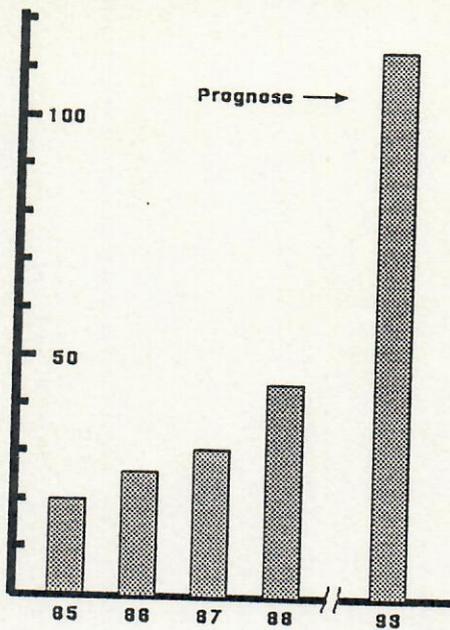


Fig.2: Weltweiter Umsatz mit Halbleitern (Quelle: Semicond. Int.)

Zellengröße [μm^2] Chipgröße [mm^2]

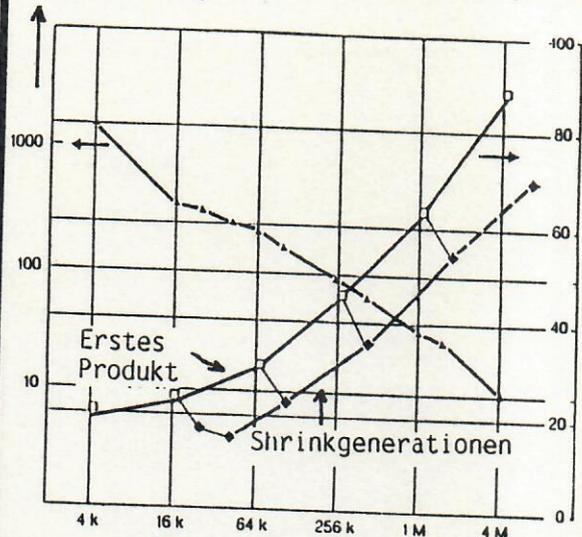


Fig.3: Chipgröße und kleinste Strukturen bei Speichern (Quelle: /6/)

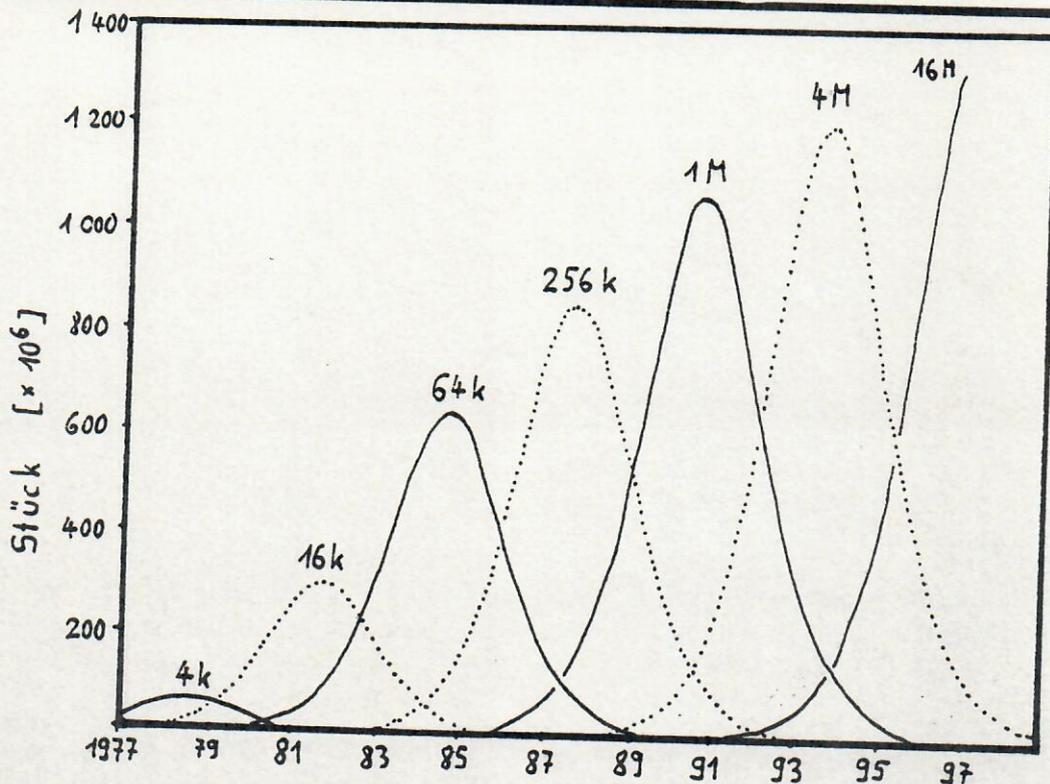


Fig.4: Generationenfolge bei DRAMs in Stück. Die Kurven bis einschließlich 1M entsprechen recht genau der Wirklichkeit; die Folgekurven (nach 1988) sind Prognosen.

Fig.6: Erzielbare Auflösung und Schärfentiefe der Lithographie und Anforderungen bei der Produktion als Funktion der DRAM Generationen (Quelle: /5/).

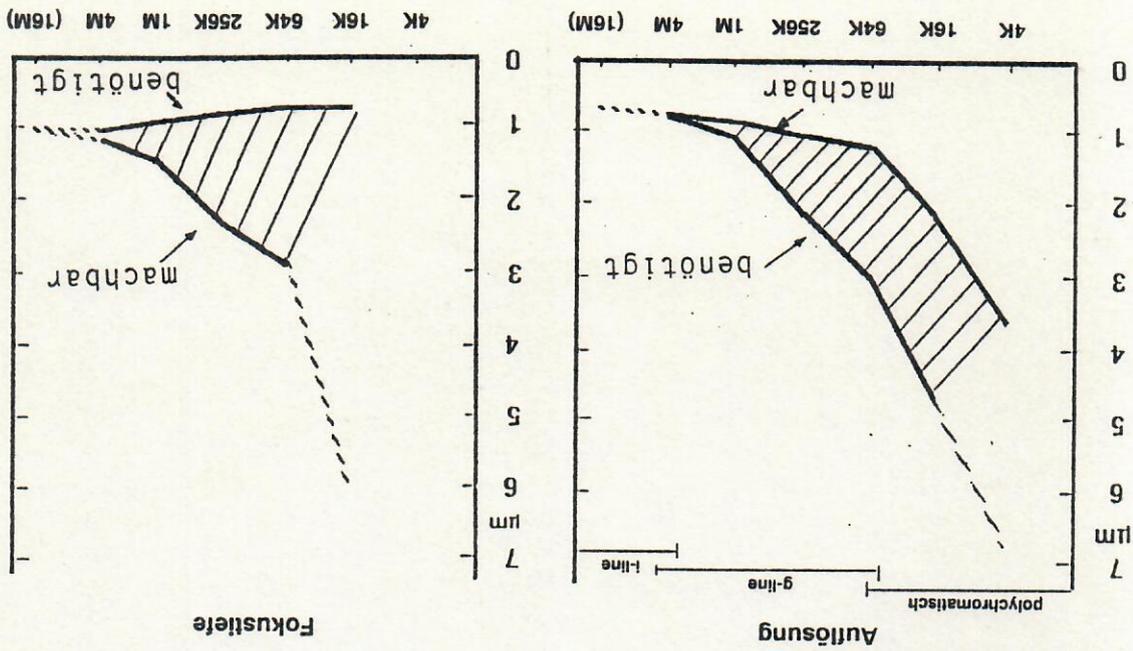
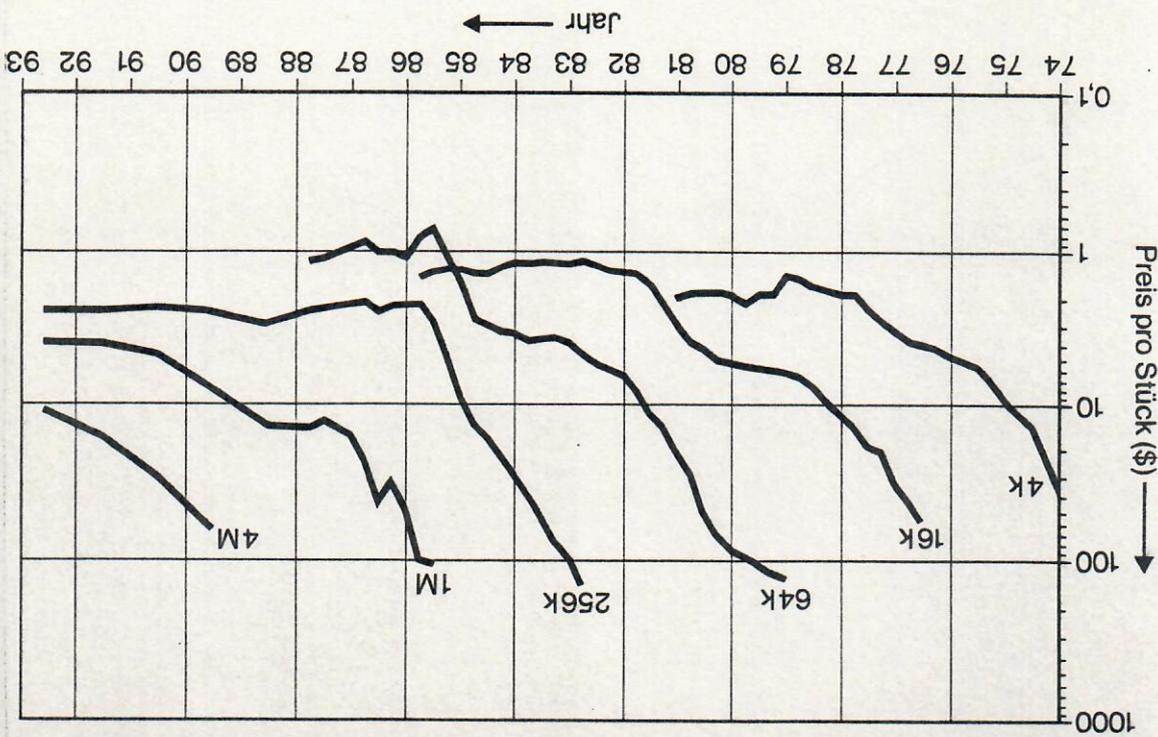


Fig.5: Preisentwicklung bei DRAMs (Ist-Werte bis 1987; danach Prognose; Quelle: Dataquest). Man beachte den logarithmischen Maßstab!



3. Rapider Preisverfall innerhalb kurzer Zeit für Speicherchips (Fig. 5).

Unmittelbare Konsequenzen daraus sind: Nur Produzenten die erstens von Anfang an dabei sind und zweitens ihre Ausbeute mindestens genauso schnell steigern, wie der Preis verfällt, haben eine Chance, genügend Geld zu verdienen. Schon ein um ein halbes Jahr verzögerter Markteintritt kann zur finanziellen Katastrophe führen.

4. Produktion eines Massenartikels an den Grenzen der Physik/Chemie/Ingenieurskunst. Fig. 6 illustriert dies für die Lithographie. Im Gegensatz zu praktisch allen sonstigen Produktionsprozessen in "normalen" Industriezweigen ist zwischen den für die Fertigung benötigten spezifischen Leistungen und den im Labor erreichbaren praktisch kein Unterschied mehr da.

Unmittelbare Konsequenz daraus ist, daß während der F+E Phase mit Prozessen und Anlagen gearbeitet werden muß, die den Ansprüchen an die späteren Fertigungsprozesse nicht genügen. Ob das in der Entwicklung befindliche Produkt rechtzeitig fertig ist und dann kostengünstig gefertigt werden kann, ist daher nicht mit ausreichender Sicherheit planbar. Im Endresultat heißt dies, daß Vorleistungen in Milliardenhöhe bei extrem hohem Risikoerbracht werden müssen.

Als spezifisch europäische/deutsche Randbedingungen kommen noch hinzu:

5. Im Vergleich zu Japan/USA eine nahezu nicht existierende Infrastruktur bei Geräteherstellern (Fig. 7) und eine deutlich schwächere Basis bei der Ausbildung (Mikroelektronik war bis vor kurzem kein relevantes Thema an deutschen Schulen und Universitäten).
6. Hohe und damit zeitraubende arbeitsrechtliche/gesellschaftspolitische Hürden die den erforderlichen schnellen Wandel dieses Industriezweigs behindern.
7. Erheblich geringere Akzeptanz in der BRD von "High-Tech" im allgemeinen und Mikroelektronik im Besonderen, dokumentiert in dem vergleichsweise kleinen Chipverbrauch pro Kopf der Bevölkerung. (Japan: 170 DM an Chips/Kopf und Jahr; BRD: 60 DM an Chips pro Kopf und Jahr).

Um Mißverständnisse auszuschließen: Die Punkte 5-7 sollen keine wie auch immer geartete Wertung der Verhältnisse in der BRD darstellen, sondern lediglich die im Vergleich mit Ostasien und USA vorliegende Situation beleuchten.

16.1.3 Forschungs- und Entwicklungsphase

Ungefähr 4-5 Jahre vor Aufnahme der Produktion muß die F+E Phase der neuen Speichergeneration einsetzen. Im Höhepunkt dieser Phase, d.h. den letzten beiden Jahren vor Produktionsbeginn, arbeiten größenordnungsmäßig 100 Physiker, Ingenieure

und andere hochqualifizierte Mitarbeiter an dem Projekt. Diese Experten arbeiten überwiegend nicht in Labors, sondern in großen Reinräumen (sog. F+E- oder Pilot-Linien), die sich von einer Fertigungslinie höchstens noch dadurch unterscheiden, daß sie etwas kleiner sind - dafür in der Regel teurer pro m². Legt man die dabei entstehenden Kosten auf die direkten Projektmitarbeiter um, kostet ein Arbeitsplatz ca. 10⁶ DM/a.

Schon in der F+E Phase schlagen deshalb betriebswirtschaftliche Sachzwänge durch. Aus der Sicht der Fertigung müssen die vom Vorgänger her vorhandenen Prozesse, Designs, Verfahrensweisen etc. weitestgehend übernommen werden. In anderen Worten: Das neue Produkt soll sich möglichst wenig vom alten unterscheiden. Das "alte" Produkt ist übrigens zu diesem Zeitpunkt noch gar nicht in der Fertigung. Die F+E Leute haben aber zwischendurch viel dazugelernt; in der einschlägigen Literatur sind einige Zentner Papier zu neuen Prozessen und Verfahren erschienen; es gibt ganz neue Vorschläge für Prozesse und Prototypen von neuen Anlagen - natürlich will man jetzt alles besser (und damit anders) machen als zuvor. Damit ist ein permanenter Konflikt zwischen Forschern, Entwicklern und Fertigern angelegt, der, wenn es schlecht läuft, zur Lähmung und Frustration der Betroffenen führen kann. Wenn es gut geht, wird in bester dialektische Manier sich das aus These (nichts ändern) und Antithese (neu ist besser) aufgespannte Spannungsfeld in einer Synthese (dem neuen Produkt) entladen, die allen Anforderungen gerecht wird (besser, aber nicht teurer). Daß es im Mittel gut geht, zeigt die kurze Historie der integrierten Schaltungen: Atemberaubender Zuwachs an Komplexität bei gleichzeitigem starken Rückgang des Preises per Funktion (Fig. 8).

Die F+E Mannschaft ist also gezwungen, die ihr gesetzten Ziele mit kleinstmöglichem Aufwand zu erreichen. Für den größten Teil der Mitarbeiter heißt die Devise: Probleme lösen statt erklären; wissen worauf es ankommt, nicht warum. Dies mag aus der Sicht des "reinen" Forschers (eine insbesondere in Deutschland Artenschutz genießende Spezies) unannehmbar sein - es sei jedoch daran erinnert, daß es heute keine subum Technologie gäbe, wenn Detailverständnis vor Weiterarbeit gesetzt wird. Manche Schlüsselfragen, z.B. nach den Diffusionsmechanismen vom Dotieratomen, sind bis heute nicht völlig verstanden. Wer hier Zusammenhänge vermutet zwischen der schwachen Stellung der Mikroelektronik in Deutschland und der hierzulande besonders gepflegten, porentiefen Reinheit der Forschung, sei auf Referenz /4/ verwiesen.

Selbstverständlich ist ein weitgehendes Verständnis der physikalisch-technischen Grundlagen notwendig, und selbstverständlich muß es laufend weiterentwickelt werden. Ganz ohne neue Prozesse und Verfahren entsteht keine neue Generation - es ist die Kunst des Managements bei Großprojekten wie der Entwicklung einer neuen Speichergeneration, die diversen Interessen ständig neu auszubalancieren und die richtige Mischung aus Innovationen und Althergebrachtem zu finden - und dabei alle Mitarbeiter motiviert zu halten.

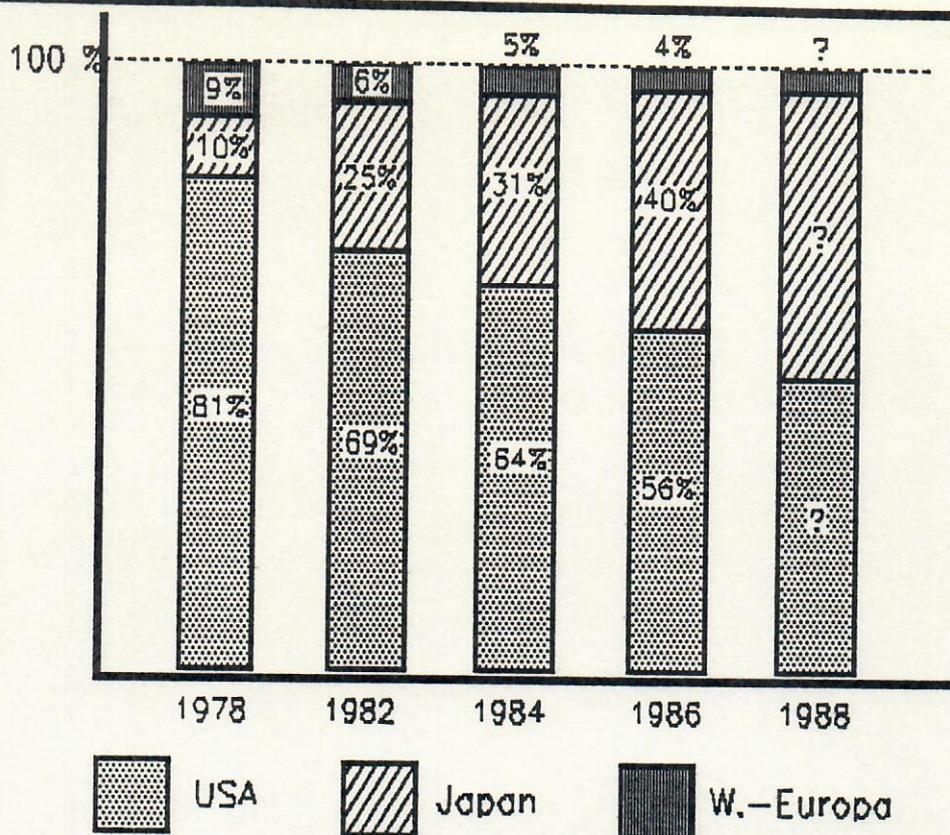


Fig. 7: Aufteilung des Gerätemarktes auf die Erzeugerländer USA, Japan und W.-Europa (Quelle: Dataquest).

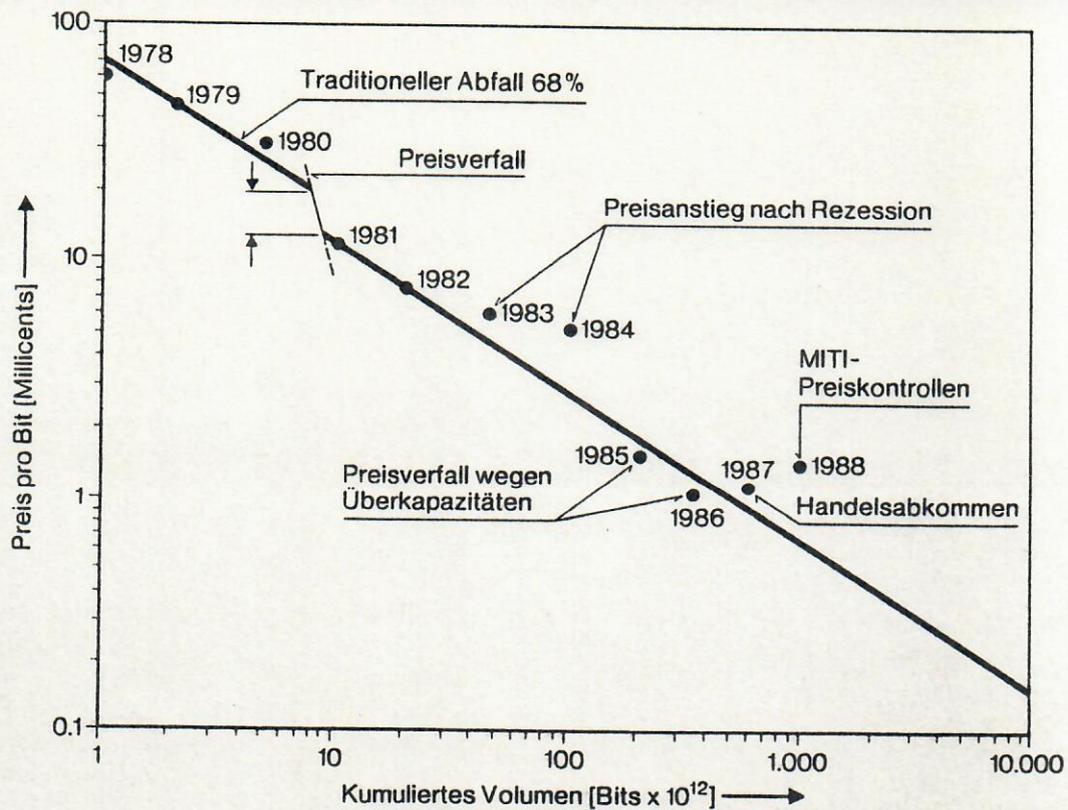


Fig. 8: Entwicklung des Preises pro Bit (Quelle: 141).

Eine detaillierte Ausarbeitung der angeklungenen Themen würde zu einer Buchserie führen; hier können nur einige Teilaspekte etwas vertieft werden. Geschlossene Darstellungen des Gesamtthemas sind den Autoren nicht bekannt; zur weiteren Vertiefung ist insbesondere die Zeitschrift: "mikro elektronik" (Hrsgb.: VDE/VDI) empfohlen, aber auch: "Semiconductor International" (Hrsgb.: D. J. Levinthal; Cahner Publ.) und "Solid State Technology" (Hrsgb.: Penwell Publ. Co).

16.1.4 Hintergrund und Gliederung dieser Arbeit

Im Gesamtfeld der Mikroelektronik gibt es einige kontrovers diskutierte Problemkreise; dazu relevante Stichworte sind schon angeklungen: Sontagsarbeit, Wirtschaftskrieg, Umweltgefährdung, etc.. Wenn in dieser Arbeit direkt oder indirekt auf diese Themen eingegangen wird, dann aus der Überzeugung heraus, daß alle diese Themen sachlich diskutiert werden können und daß ein Ausklammern auf Dauer mehr Schaden als Nutzen stiftet. Sofern die Aussagen dazu eine Wertung enthalten oder zu enthalten scheinen, handelt es sich um die Privatmeinung der Autoren. Im folgenden wird aus Platzmangel (und aus mangelnder direkter Kompetenz der Autoren) nur noch auf den konkreten Fall der Entwicklung und Fertigung von Speichern, speziell DRAMs, Stellung genommen. Im großen und ganzen gilt das hier Ausgeführte aber für die gesamte Mikroelektronik; im Detail kann die Lage bei anderen Produkten jedoch recht verschieden sein.

In den folgenden Kapiteln wird zunächst der wissenschaftlich - technische Hintergrund der DRAM Herstellung näher beleuchtet, wobei weniger die unmittelbare Physik einzelner Komponenten und Prozeßschritte betrachtet werden soll, sondern die Einbindung in das Gesamtthema Entwicklung und Fertigung. Schaltungstechnik, Design, Maskenerstellung, Einzelprozesse wie Lithographie, Ätzen oder Schichtabscheidung, Prüftechnik, Montage und Zuverlässigkeit sind die zu nennenden Stichwörter.

In den weiteren Kapiteln wird der Aufbau, die Organisation und die Infrastruktur des Reinraums - der Linie im Insider-Sprachgebrauch - beschrieben. Zum Schluß fügen wir die einzelnen Bausteine zusammen und betrachten die Struktur und den Ablauf des gesamten Projektes: Entwicklung und Fertigung einer neuen Speichergeneration.

Diese Arbeit ist nicht "sine ira et studio" entstanden. Soweit möglich, wurde versucht, alle Aussagen durch Zahlen und Fakten zu belegen. Wo dies aus Vertraulichkeitsgründen nicht direkt möglich war, wurden Zahlen soweit verändert, daß sie einerseits keine "nützlichen" Rückschlüsse auf die konkrete Situation bei Siemens zulassen, andererseits aber zur Belegung der Aussage noch vollständig korrekt sind.

16.2 Schaltungstechnik und Design

Wie Tabelle 2 zeigt, ist mit jeder Speichergeneration nicht nur die Kapazität der Chips, sondern auch die Komplexität der Schaltung gewachsen. Im folgenden Abschnitt wird beschrieben,

wie man von der prinzipiellen Idee für einen neuen Chip zu den fertigen Masken für die Technologielinie gelangt.

16.2.1 Prinzip und Realisierung eines Halbleiterspeichers

Gemeinsame Grundlage aller Halbleiterspeicher ist die binäre Digitaltechnik. Diese beruht darauf, daß zuerst die zu verarbeitenden Informationen digitalisiert, d.h. in Zahlen überführt werden, gleichgültig, ob es sich um Text, Bilder oder z.B. Musik handelt. Diese Zahlen werden dann im Binär-system dargestellt, das im Gegensatz zu unserem vertrauten Dezimalsystem nicht 10, sondern nur 2 Zahlen kennt. Da es sich hierbei um "Null" und "Eins" handelt, braucht die verarbeitende Schaltung nur zwei Möglichkeiten (z.B. Spannung oder Ladung ja/nein) zu erkennen, wodurch neben einer beträchtlichen Vereinfachung auch eine Reduzierung der Störanfälligkeit erreicht wird. Die Aufgabe des Halbleiterspeichers besteht nun darin, die ihm eingeschriebenen, als "Bits" bezeichneten Null/Eins Informationen bis zu ihrer Wiederverwendung bereit zu halten. Je nach Verwendung unterscheidet man dabei zwischen nichtflüchtigen und flüchtigen Speichern. Während erstere ihre Informationen auch bei Abschalten der Spannungsversorgung behalten (z.B. "Read only Memory", ROM), gehen diese bei flüchtigen Speichern verloren (z.B. Random Access Memory-RAM) /42/. Letztere finden deswegen ihr Einsatzgebiet z.B. als "Kurzzeitgedächtnis" in Computern, in denen sie die bei Abarbeitung eines Programmes anfallenden Informationen bis zur Weiterverarbeitung bereithalten.

Bei dem im Mittelpunkt dieser Betrachtungen stehenden "Dynamic Random Access Memory (DRAM)" handelt es sich um einen flüchtigen Speicher mit wahlfreier Möglichkeit, Bits einzuschreiben bzw. auszulesen. Die Information ist in Form einer Ladung in einem Kondensator gespeichert, der über den als Schalter wirkenden Auswahltransistor angesteuert werden kann. Fig. 9 zeigt diese für das DRAM typische "Ein-Transistor-Zelle", die durch Adressierung der entsprechenden Wort- und Bitleitungen angesprochen werden kann.

Da die im auch "Varaktor" genannten Kondensator gespeicherte Ladung mit der Zeit durch Leckströme verloren geht, ist eine periodische Auffrischung nötig. Diese erfolgt durch eine in der "Peripherie", d.h. außerhalb des Zellenfeldes angebrachte Logikschaltung, die beim 4 Megabit-DRAM z.B. alle 16 Millisekunden das Vorhandensein einer Ladung im Kondensator überprüft und diese ggf. erneuert ("Refresh") /24/. Eine hohe Kapazität des Zellenkondensators ist dabei nicht nur wichtig, um die Refreshfrequenz möglichst niedrig halten zu können. Auch die Störsicherheit gegenüber von Alphateilchen verursachten momentanen Kurzschlüssen und das Verhältnis zur Kapazität der die einzelne Zelle ansteuernden Leiterbahnen sind zu berücksichtigen. Wegen der durch das Gehäuse beschränkten Chipfläche führen die Forderungen nach hoher Kondensatorkapazität (= große Kondensatorfläche) bei kurzen Leiterbahnen (= geringe parasitäre Kapazität) z.B. beim Siemens 16Mbit-DRAM zur Verwendung sowohl einer dreidimensionalen Grabenzelle als auch zweier Lagen von Aluminium-Leiterbahnen. Aufgrund des hohen Anteils des Zellenfelds (> 50%) an der

Eigenschaften	16 K	64 K	256 K	1 M	4 M	16 M
Speicherbare DIN A4 Schreibmaschinenseiten	1	4	16	64	250	1000
Durchmesser Si-Scheiben (Produktion) in mm	100	100	125	150	150	200
Chips pro Scheibe	390	250	220	240	140	ca. 130
Chipfläche in mm ²	16	25	45	55	91	ca. 140
Anzahl der Bauelemente	36500	150000	565000	2,2 Mio	8,9 Mio	ca. 36 Mio
Zahl der Transistoren	ca. 18000	27000	290000	1,1 Mio	4,7 Mio	ca. 19 Mio
Organisation	16000 x 1	64000 x 1	256000 x 1	1 M x 1	4 M x 1	16 M x 1
Technologie	NMOS	NMOS	NMOS	CMOS	CMOS	CMOS
Zahl der Prozessschritte	80	80	120	250	ca. 400	ca. 450
Fertigungsdurchlauf	ca. 2 Monate					> 2 Mon
Produktionszeitraum	80/81 85/86	81/82 86/87	seit 85/86	seit 87/88	ab 1989	ab 1992
Kleinste Strukturgröße	4.0 µm	2.0 µm	1,5 µm	1,2 µm	0,8 µm	0,6 µm
Kritische Defektgröße	1,3 µm	0,8 µm	0,6 µm	0,4 µm	0,2 µm	0,15 (?)

Tabelle 2 Vergleichsdaten der verschiedenen DRAM Generation

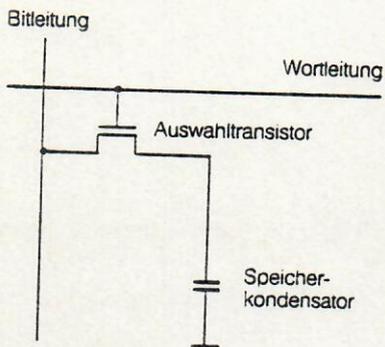


Fig. 9 Schaltbild der Speicherzelle eines DRAMs. Der 4M-DRAM enthält diese Kombination aus Speicherkondensator und Auswahltransistor 4, 19 Millionen mal.

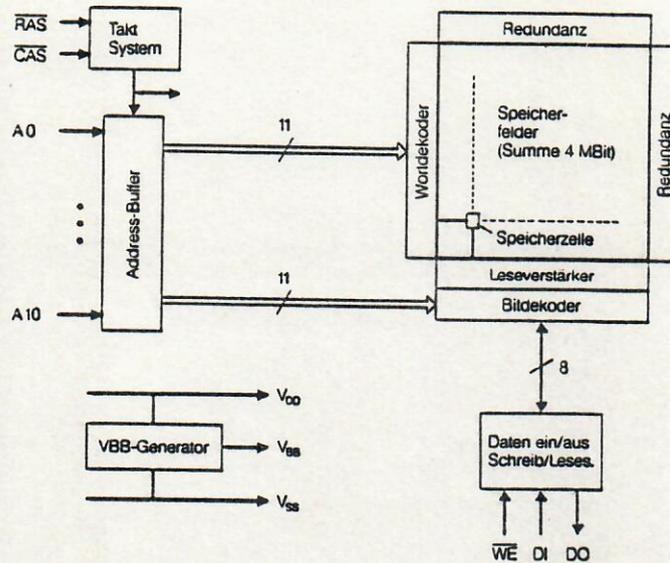


Fig. 10 Schematische Anordnung der einzelnen Komponenten eines DRAMs.

Gesamtfläche des Chips ist der Druck, durch Minimierung der Strukturen Platz zu sparen, dort besonders groß. Dennoch wird das Zellenfeld z.B. bei 4Mbit-DRAM in 8 512kbit-"Blöcke" aufgegliedert, da sonst zu manchen Zellen die Leitungen zu lang wären, wodurch sowohl deren Eigenkapazität als auch die Signalaufzeiten zu groß würden.

Wie in Fig. 10 gezeigt ist das Zellenfeld von den sog. "Rasterschaltungen" umgeben. Diese bestehen z.B. aus den Dekoderschaltungen in Wortleitungs- und Bitleitungsrichtung, Leseverstärkern und Schaltungen zur Vorladung der Bitleitung. Die sich daran anschließende "Peripherie" des Speichers ist unregelmäßig. Sie enthält z.B. Generatoren für verschiedene Spannungspegel, die für den internen Betrieb benötigt werden, I/O Schaltungen sowie Schutzvorrichtungen gegen Spannungsspitzen, die durch elektrostatische Entladungen bei der späteren Weiterverarbeitung der Chips verursacht werden können (ESD-Schutzschaltungen).

Aus Gründen der Ausbeutesteigerung und damit der Wirtschaftlichkeit im Fertigungsstadium benutzen alle Speicher der Megagenerationen redundante, d.h. zusätzliche Wort- und Bitleitungen. Im Falle unseres 4M-bit-Speichers sind je 4 redundante Wort- und Bitleitungen pro 512k-bit Block vorgesehen. Nach Beendigung des technologischen Herstellungsprozesses auf Scheibenebene wird in einem ersten Funktionstest beurteilt, ob der Baustein reparierbar ist und wenn ja, welche Bits zu reparieren sind. Dieses Testergebnis steuert dann über einen Redundanzalgorithmus eine Laserapparatur, welche mit gezielten Laserschüssen etwa 2 μm breite Bahnen so auftrennt, daß die fehlerhaften Wort- und Bitleitungen abgetrennt und dafür die redundanten, funktionstüchtigen Leitungen hinzugeschaltet werden. Die Zahl der maximal reparierbaren Ausfälle hängt von der Verteilung der Fehler ab. Erfahrungsgemäß kann diese Laserredundanz speziell in der Anfangsphase einen Ausbeutegewinn von mehr als 100% erbringen.

Einer besonderen Feinoptimierung bedarf das Lesen und Bewerten der gespeicherten Information. Dieses erfolgt nach dem Prinzip der Differenzverstärkung: Eine Bitleitung, welche nach der Vorladung auf 2,5V durch Auslesen des Zellinhalts potentialmäßig geringfügig abgesenkt oder angehoben wurde, wird mit einer benachbarten Bitleitung verglichen. Die Potentialunterschiede dieser beiden Bitleitungen liegen bei nur ca. 0.1 Volt, die Bewertung muß innerhalb von ca. 10 Nanosekunden sicher erfolgen. Die Sicherheit in der Bewertung wird erhöht durch eine möglichst große Zellkapazität und eine möglichst kleine Bitleitungskapazität (d.h. starke Unterteilung des Chips in Bitleitungsrichtung). Ferner muß die Empfindlichkeit auf störende, statistisch auftretende radioaktive Alphastrahlung minimiert werden. Eine Verfälschung des Informationsgehaltes kann sowohl durch Treffen einer Zelle als auch durch Treffen einer Bitleitung von einem Alphateilchen geschehen. Diese natürliche Alphastrahlung wird aus der Preßmasse, die den Chip umhüllt, aber auch aus Materialien, welche im Herstellungsprozess verwendet werden (z.B. Aluminium oder Metallsiliziumlegierungen) ausgesandt. Die Verwendung hochreiner Materialien (z.B. strahlungsarme Preßmassen) sowie optimiertes Design der

Zelle sind die wesentlichen Ansätze zur Beherrschung dieses wichtigsten Störeffektes bei dynamischen Speichern.

16.2.2 Schaltungsentwurf und Simulation

Komplexe Schaltungen mit Hunderttausenden oder Millionen von Einzelelementen lassen sich nur noch computerunterstützt herstellen. Ausgehend von den vorgegebenen Zielfunktionen wird die Schaltung zunächst in Funktionsblöcke zerlegt, deren Gesamtverhalten untersucht wird. Zunehmend wird beim Design auch mit "Logikzellen" gearbeitet, wobei es sich um häufig gebrauchte Teilfunktionsblöcke handelt, deren Daten in Rechnerbibliotheken zur Verfügung stehen.

Fig. 11 zeigt ein Beispiel für den schrittweisen Aufbau einer integrierten Schaltung mit Hilfe eines "Computer Aided Engineering" (CAE) Systems. Man sieht, daß auf allen Ebenen die Simulation eine wichtige Rolle spielt.

Bei der sogenannten Logik-Simulation geht es z.B. darum, die logischen Abläufe auf dem Chip zu verifizieren. Diese Logiksimulation ist im wesentlichen digital. Ein Beispiel dafür ist die Überprüfung der Dekoderschaltungen, die den außen angelegten Adressen entsprechende Wort- und Bitleitung auswählt. Von zentraler Bedeutung ist die sogenannte Schaltkreissimulation, bei der der zeitliche Ablauf (timing) auf einem Datenpfad simuliert wird. Bevor eine solche Simulation durchgeführt werden kann, müssen sämtliche an einem solchen Datenpfad beteiligten Elemente wie Transistoren, Widerstände und Kapazitäten als Datenbasis für diese Simulation im Rechner vorhanden sein. Auf Basis der so entstandenen Netzliste, die am Beispiel des 4M-bit-Speichers bis zu 5.000 Transistoren umfaßt, kann dann das gesamte Zeitverhalten vom Beginn z.B. eines Lesezyklus bis zum Anliegen des ausgelesenen Datums am Ausgangspin simuliert werden. Derartige Simulationen sind äußerst rechnerintensiv. So verbraucht z.B. eine Simulation des kritischen Pfades des 4M-Speichers an einem modernen Vektorprozessor 4,5CPU Stunden. Die Abwicklung einer solchen Simulation im Vergleich zu später gemessenen Werten liegt erfahrungsgemäß bei 10-20%. Mit Hilfe der Simulation kann nicht nur die Zugriffszeit eines Speichers im vorhinein ermittelt, sondern auch die Empfindlichkeit des Bausteins auf verschiedene Einflußgrößen untersucht werden, indem man z.B. die nicht vermeidbaren Prozeßschwankungen vorgibt und als Eingangsparameter für Simulationsläufe verwendet.

Die auf diese Weise entworfenen logischen Schaltungen müssen jetzt in ein geometrisches Layout umgesetzt werden. Hier werden die Informationen für alle im Herstellprozeß des Speichers verwendeten Strukturierungs- und Implantationsschritte aufgebaut. Mit umfangreichen Verifikationsläufen wird festgestellt, daß das geometrische Layout die geforderten Designregeln erfüllt, elektrisch funktionsfähig ist und die geometrisch beschriebenen Schaltungen identisch mit den entworfenen logischen Schaltungen sind. Erst nach diesen Kontrollen werden die Daten für jede Prozeßebene getrennt in sogenannte Kenndaten für den Elektronenstrahlschreiber

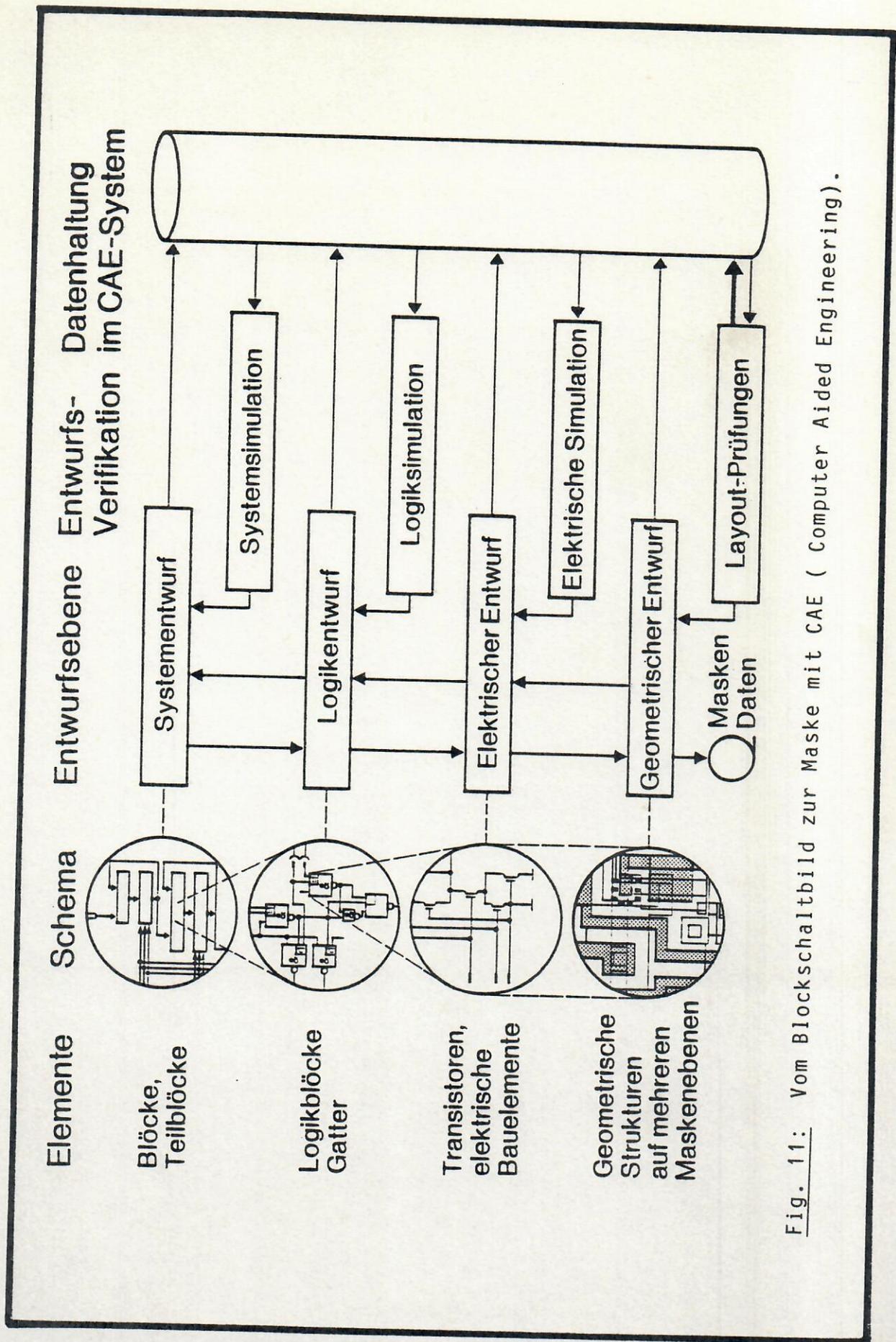


Fig. 11: Vom Blockschaltbild zur Maske mit CAE (Computer Aided Engineering).

umgesetzt, der dann die Masken für die einzelne Ebenen (ca. 19 beim 4M bit DRAM) erzeugt.

16.3 Einzelprozesse

Die Herstellung einer integrierten Schaltung geschieht durch häufige Wiederholung einer Reihe von Einzelprozessen. Diese lassen sich grundsätzlich in 5 Kategorien einteilen:

1. Erzeugung neuer Schichten durch Abscheidung, Aufstäuben, oder Oxidation.
2. Veränderung existierender Materialien (Diffusion, Implantation).
3. Photolithographische Übertragung der Maskenstrukturen in eine Lackschicht auf dem Chip.
4. Ätztechnische Übertragung der Lackstrukturen in die Chipoberfläche.
5. Naßchemische Prozesse wie z.B. Reinigung.

Die Entwicklungsfortschritte bei der Erhöhung der Chip-Integrationsdichte haben die Anforderungen an die Einzelprozesse ständig verschärft. Besonders in der Schicht- und Ätztechnik hat dies zu einer breiten Vielfalt von Prozessen geführt, die für jeweils spezifische Anforderungen optimiert wurden. Die folgende Charakterisierung kann sich deshalb nur auf einige wesentliche Aspekte des jeweiligen Gebiets beziehen. Details findet man z.B. in /8,9/.

Schichtherstellung

Der klassische Prozeß zur Erzeugung isolierender Siliziumdioxidschichten ist die Oxidation freiliegenden Siliziums bei Temperaturen von 900°-1100°C. Als oxidierende Atmosphäre wird O₂ oder H₂O verwendet, wobei in manchen Fällen auch HCl zur Verbesserung der Oxidqualität zugesetzt wird. Hauptanwendungen sind bei DRAMs die Herstellung des dicken Feldoxids (ca. 800 nm) und der dünnen "Gateoxide" (< 20nm). Ersteres dient zur elektrischen Isolation der aktiven Gebiete, während dünne Oxide sowohl für die Transistorgates als auch für das Dielektrikum des Speicherkondensators benötigt werden. Mit der Zunahme der Anzahl von Schichten in integrierten Schaltungen stieg die Bedeutung der Schichtabscheidung durch "Chemical Vapor Deposition" (CVD). Dabei werden in einem Reaktionsraum Gase durch Energiezufuhr zur Reaktion gebracht. Die Prozesse können sowohl bei Atmosphärendruck (Atmospheric Pressure CVD) als auch bei Unterdruck (Low Pressure CVD) ablaufen. Die festen Reaktionsprodukte schlagen sich als Schichten auf den Siliziumscheiben, z.T. auch auf den Reaktorwänden nieder, während die flüchtigen Verbindungen abgepumpt werden. Die Energiezufuhr geschieht üblicherweise thermisch, d.h. durch Wärme, oder über eine elektrische Entladung (Plasma-CVD). Tabelle 3 gibt einen Überblick über die wichtigsten Standard-CVD Verfahren, die grundlegenden Reaktionsparameter und die hauptsächlichen Schichtenanwendungen.

Vorteile des vielseitigen CVD-Verfahrens sind gute Schichtqualität (Kantenbedeckung, Gleichförmigkeit), z.T. auch hoher Scheibendurchsatz. Nachteilig sind die z.T. hohen

Name	Schicht Formel	Abscheideparameter		Prozeßgase	Schichtanwendung	Bemerkungen
		Temp. °C	Druck			
Poly-Silizium	Si	630	LP	SiH ₄	Gate- bzw. Kondensator- elektrode, leitende Verbindungen	Langjähriger Standardprozeß
Silizium- nitrid	Si ₃ N ₄	<800	LP	SiH ₂ Cl ₂ + NH ₃	LOCOS-Maske, Ätzmaske, ONO-Dielektrikum	Langjähriger Standardprozeß
TEOS- Oxid	SiO ₂	720	LP	Si(OC ₂ H ₅) ₄	Ätzmaske, Isolations- Oxid	Neuer Standardprozeß (Lit. 43)
TEOS- Arsenglas	SiO ₂ + As ₂ O ₃	700	LP	Si(OC ₂ H ₅) ₄ + O ₂ + AsO(OC ₂ H ₅) ₃	Grabendotierung	Spezieller 4M-DRAM Prozeß (Lit. 44)
BPSG	SiO ₂ +B ₂ O ₃ +P ₂ O ₅	450	AP	SiH ₄ /N ₂ +O ₂ + B ₂ H ₆ /N ₂ + PH ₃ /N ₂	Planarisierendes Zwischenoxid	Viele verschiedene Verfahren entwickelt (Lit. 40)
PSG	SiO ₂ +P ₂ O ₅	430	AP	SiH ₄ /N ₂ +O ₂ +PH ₃ /N ₂	Passivierung	Problem: Kantenbedeckung
Plasma- Nitrid	Si ₃ N ₄ (+H ₂ , N ₂)	400	LP	SiH ₄ + NH ₃	Passivierung	Streß zu optimieren
Wolfram	W	400	LP	WF ₆ + SiH ₄	Kontaktlochauffüllung Metallisierung	Neuer, teurer Prozeß (Lit. 22)
Plasma Oxid	SiO ₂	ca. 350	LP	Si(OC ₂ H ₅) ₄ + O ₂	Intermetallidi- elektrikum	Alter Prozeß auf neuen Anlagen (Lit. 40)

LP = Low Pressure (<10 Torr), AP = Atmospheric Pressure

Tabelle 3: Gebräuchliche Chemical Vapor Deposition (CVD) Verfahren

Abscheidetemperaturen, die Partikelproblematik und die Verwendung gefährlicher Gase.

Speziell für Elemente, die nicht für CVD reaktionsgünstige Verbindungen bilden, wurde das Verfahren des Aufstäubens, engl. "Sputtern", entwickelt. Bei diesem rein physikalischen Prozeß wird das auf einem "Target" befindliche Ausgangsmaterial mit energiereichen Ionen beschossen und auf die gegenüber angebrachte Scheibe aufgestäubt. Auch wenn dabei einige der CVD-Probleme umgangen werden können, so sind doch geringer Durchsatz und hohe Partikelanfälligkeit gewichtige Nachteile. In der Praxis werden vor allem Antireflexschichten und leitende Elemente (Mo, Ta, W, Al) gesputtert, wobei die prozeßinhärente schlechte Kantenbedeckung besonders beim Aluminium inzwischen zu einem grundlegenden Hindernis bei der Erhöhung der Integrationsdichte geworden ist. Die Entwicklung von CVD Prozessen (selektiv, bzw. mit guter Kantenbedeckung) auch für Metalle bildet deswegen einen aktuellen Entwicklungsschwerpunkt der Schichttechnik.

Veränderung von Materialien

Diese Überschrift faßt Prozesse zusammen, die nur die Eigenschaften bereits vorhandener Materialien verändern. Bei der Implantation werden hochenergetische Ionen (B, P, As) zumeist in Silizium hineingeschossen, um dessen Leitfähigkeit zu verändern. In einer anschließenden Hochtemperatur-Diffusion werden diese Elemente dann durch Einbau in das Kristallgefüge elektrisch aktiviert und räumlich weiter, bzw. gleichmäßiger verteilt, wobei zugleich die bei der Implantation entstandenen Kristallschäden ausgeheilt werden. Unter der Bezeichnung "Temperung" dient ein ähnlicher Temperaturschritt auch zur Verbesserung der Qualität gesputterter Schichten bzw. der Bildung von Verbindungen (z.B. MoSi). Durch eine bei erhöhter Temperatur durchgeführte "Verfließung" schließlich läßt sich eine bei der Abscheidung mit Bor, bzw. Phosphor dotierte SiO₂-Schicht ("Borphosphorsilikatglas", BPSG) erweichen, so daß die Oberflächenspannung die Konturen glätten kann /10/.

Photolithographie

Da Schichten in der Regel nur ganzflächig abgeschieden werden, aber nicht überall erwünscht sind, müssen sie in einer extra Prozeßfolge strukturiert werden. Als erster Schritt wird dazu die Scheibe mit einem lichtempfindlichen, auch "Resist" genannten Fotolack überzogen. In diesen wird dann unter Verwendung der entsprechenden Maske die Struktur hineinbelichtet. Während an den vom Licht getroffenen Stellen der Lack durch eine chemische Reaktion so verändert wird, daß er durch organische Lösungsmittel leicht entfernt werden kann, bleibt er in den unbelichteten Bereichen stehen. Die entstandenen Lackstrukturen werden gegebenenfalls noch durch eine spezielle Bestrahlung mit ultraviolettem Licht nachgehärtet, um bei dem anschließenden Ätzschritt die darunterliegende Schicht besser vor dem Abtrag zu schützen. Bei einer in den letzten 15 Jahren prinzipiell ähnlich gebliebenen Prozeßfolge haben sich die Anforderungen an die Photolithographie mit jeder Speichergeneration weiter

verschärft. Grund ist einmal die durch fortschreitende Integrationsdichte erzwungene Tendenz zu immer feineren Strukturen (vgl. Fig. 6,23). Darüberhinaus nehmen jedoch zugleich die Höhenunterschiede auf dem Chip zu, da die Bauelemente zwar enger gepackt werden, die Zahl der Schichten aber zunimmt und ihre Dicken kaum verringerbar sind. Dies hat zur Folge, daß die Tiefenschärfe der die chipweise Belichtung durchführenden "Stepper" ein immer kritischer Parameter wird. In Fig. 6 ist die Entwicklung der mit den jeweils modernsten Steppern erzielbaren Fokustiefen und der minimalen lateralen Auflösung im Vergleich zu den geforderten Strukturbreiten über den verschiedenen Speichergenerationen dargestellt. Man erkennt, daß eine Verbesserung der Strukturfeinheit mit einer Verringerung der Fokustiefe, d.h. des auf dem Chip tolerierbaren Höhenunterschieds, verbunden ist. Dieser Zusammenhang liegt in der Physik des Abbildungsvorganges begründet, bei der die Maskenstrukturen durch Linsenobjektive verkleinert auf den Chip übertragen werden, wie es schematisch in Fig. 12 dargestellt ist. Die minimale Bahnbreite R (= Auflösung) berechnet sich gemäß der Formel

$$R = a \cdot \frac{\lambda}{NA}$$

Dabei ist λ die Lichtwellenlänge und a ein empirisch ermittelter Wert, der in der Praxis typischerweise ca. 0.8 beträgt. Die numerische Apertur NA stellt eine für das Linsensystem charakteristische Zahl dar, deren Vergrößerung einen erheblichen Entwicklungsaufwand für neue Objektive erfordert. Außerdem muß eine größere numerische Apertur mit einem kleineren Bildfeld und einem reduzierten Tiefenschärfenbereich F gemäß der Beziehung

$$F = \lambda / (NA)^2$$

erkaufte werden.

Aber auch der zweite Ansatzpunkt zur Verringerung von R , der Übergang zu kürzeren Lichtwellenlängen, ist mit erheblichen Problemen verbunden, da er den Einsatz neuer Lichtquellen, anderer Linsen sowie speziell entwickelter Lacke erfordert.

Derzeit werden als Lichtquellen für die Herstellung hochintegrierter Schaltkreise hauptsächlich zwei Emissionslinien von Hochdruck-Quecksilberlampen, nämlich die g Linie (436 nm) und die i Linie (365 nm) eingesetzt (vgl. Tabelle 1).

Für die immer feineren Strukturen zukünftiger Chipgenerationen wird eine Reduzierung der Wellenlänge der für die Lithographie verwendeten Strahlen nötig sein. Untersucht werden das sogenannte "Tiefe Ultraviolett" (200-250 nm) /11/ und sogar Röntgenstrahlung (ca. 1 nm) /12/. Die bei einem solchen Wechsel notwendigen technologischen Umstellungen sind jedoch so gravierend /7/, daß aus Fertigungsgründen so lange wie möglich mit der bisherigen Lithographie gearbeitet werden wird. Eng verbunden mit der Fototechnik ist die Lacktechnik, da eine Reihe von Schwierigkeiten eine intensive Zusammenarbeit erfordern, um Lackstrukturen mit steilen Flanken und gleichmäßiger, von der wechselnden Lackdicke unabhängiger

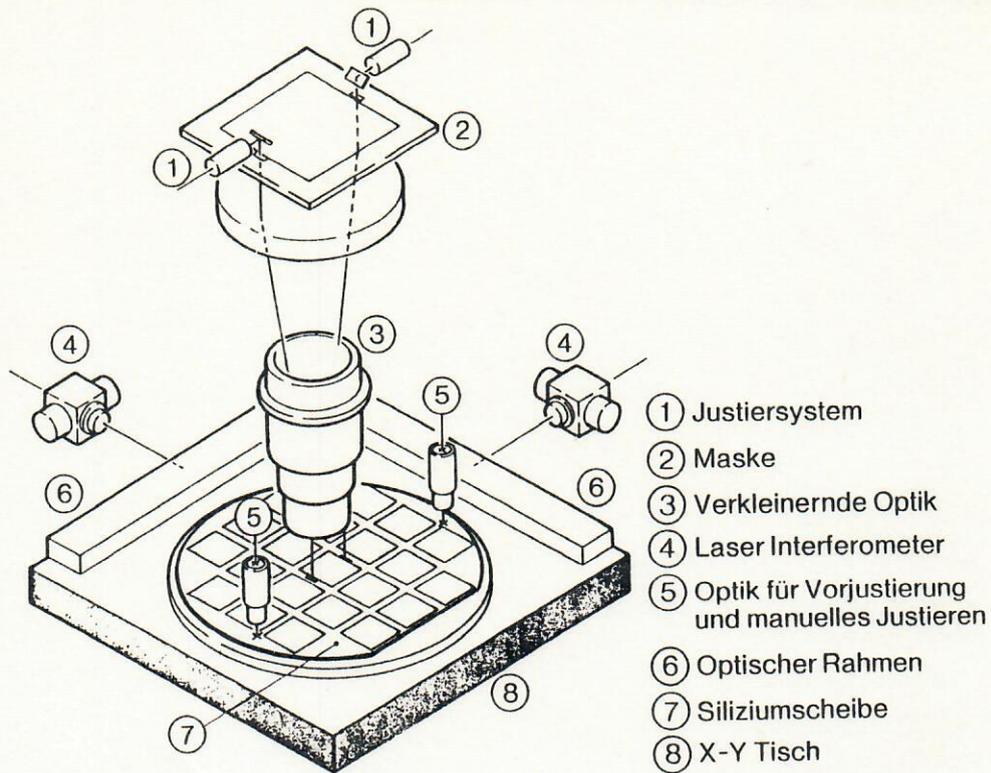


Fig. 12: Prinzipieller Aufbau eines modernen Belichtungssteppers.

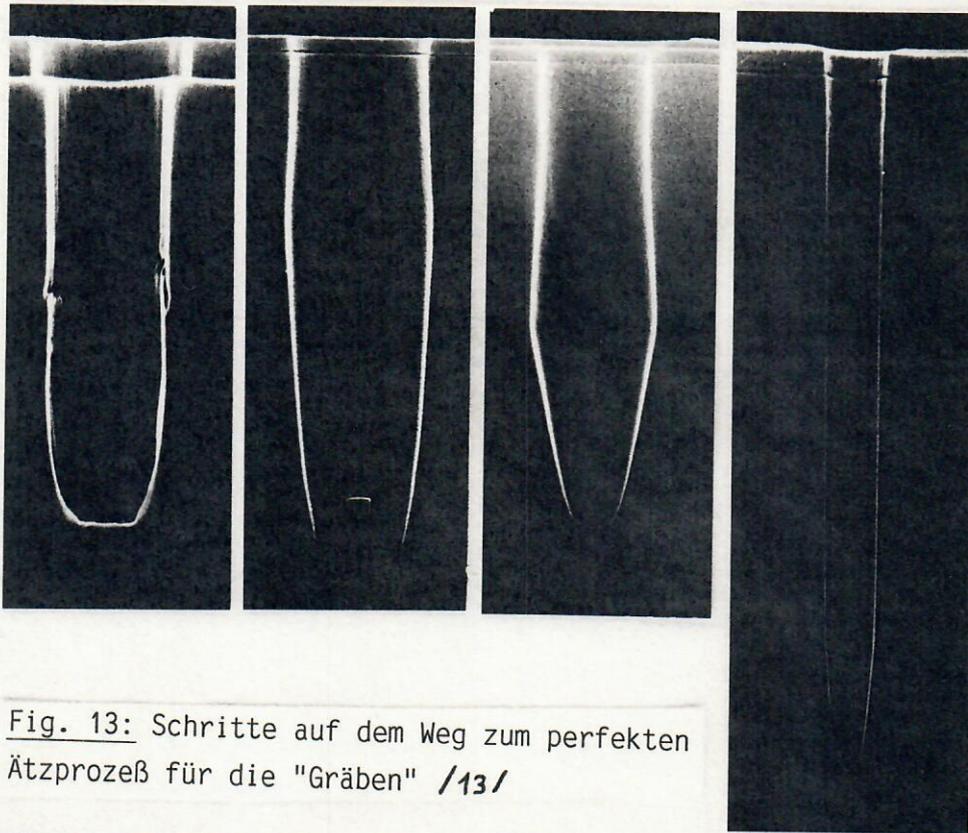


Fig. 13: Schritte auf dem Weg zum perfekten Ätzprozeß für die "Gräben" /13/

Stegbreite zu erreichen. Störende Einflüsse sind z.B. die unterlagenabhängige Lichtreflexion, die wegen der Unterschiede im Schichtaufbau über den Chip variiert. Ebenso bedingen die auf die einebnende Wirkung des Lacks zurückzuführenden Lackdickenunterschiede eine sehr sorgfältige Optimierung der Belichtungsdosis, damit tiefe Lackbereiche voll durchbelichtet werden, ohne zugleich an Stellen geringerer Lackdicke durch Überbelichtung zu schmale Stege zu erzeugen. Voraussetzung für eine Fertigbarkeit höchst-integrierter Schaltungen ist jedoch, daß für alle diese Probleme im Routinebetrieb praktikable Lösungen gefunden werden.

Ätztechnik

Die Ätzung entfernt die nicht durch Photolack geschützten Schichtbereiche und überträgt so die in den Lack belichtete Strukturinformation in die Unterlage. Der Prozeß sollte maßgenau sein und selektiv nur die zu entfernende Schicht, nicht jedoch deren Unterlagen angreifen. Für integrierte Schaltungen mit Strukturen in μm -Bereich hat sich generell die Plasma-Ätzung durchgesetzt, bei der in einer Gasentladung erzeugte Ionen beschleunigt auf die Scheibe auftreffen und so einen bei den früher verwendeten Naßverfahren nicht erzielbaren anisotropen, d.h. gerichteten Abtrag bewirken. Der Erfolg ist dabei von einer Reihe von Randbedingungen abhängig. Art und Behandlung des Photolacks spielen ebenso eine Rolle wie der Schichtaufbau, die Ätzgase, die Plasmaenergie und die Reaktorkonstruktion. In der Praxis führt dies dazu, daß für jede spezifische Anwendung der Ätzprozeß neu entwickelt oder zumindest doch optimiert werden muß. Tabelle 4 gibt einen schematischen Überblick über die am häufigsten verwendeten Reaktortypen; Fig. 13 zeigt Entwicklungsschritte bei der Ätzung des "Grabens" für das 4M-DRAM.

Bei den Prozessen wird sowohl hohe Anisotropie, d.h. Richtungspräferenz, als auch gute Selektivität gefordert, d.h. nur die zu ätzende Schicht und nicht deren Unterlage soll entfernt werden. In der Praxis besteht jedoch ein prinzipieller Widerspruch zwischen beiden Anforderungen. So lassen sich gerichtete Ätzungen am besten mit hochenergetischen, aber chemisch nicht so reaktiven Ionen (z.B. Br-Verbindungen) erreichen, während gute Selektivität eine stärker chemisch dominierte Reaktion erfordert, (z.B. F-haltige Moleküle) /7/. Häufig werden deshalb beide Prozesse in einem Reaktor sequentiell kombiniert, um in einer Überätzphase am Schluß unerwünschte Schichtreste an Stufen zu entfernen.

Naßchemische Prozesse

In früheren integrierten Schaltungen mit einem großen Verhältnis von Strukturbreite zu -Höhe wurde Naßätzprozesse auch zur Strukturierung eingesetzt. Bei höchstintegrierten Schaltungen mit Verhältnissen von fast eins ist jedoch die prozeßimmanente seitliche Unterätzung nicht mehr tolerabel. Daß dennoch die Zahl der Naßätzprozesse beim 4M weiterhin bei ca. 40% liegt /16/ ist auf die verstärkte Notwendigkeit von Reinigungsschritten zurückzuführen /45/. Automatisierung und sichere Entsorgung der verwendeten Säuren (HF, H_3PO_4 , HNO_3 , etc.) sind hier aktuelle Arbeitsschwerpunkte.

Verfahren	Barrel - Ätzen	Down-Stream-Ätzen	Plasmaätzen/ plasma mode (PE)	Reaktives Ionenätzen (RIE)	Magnetfeldunterstütztes Reaktives Ionenätzen (M-RIE)	Trioden - Reaktives Ionenätzen (TRIE)
Reaktorprinzip						
Art der Ätzung	chemisch (batch) im Plasma	chemisch (batch, single-w.) außerhalb des Plasmas	chemisch bis phys.-chem. im Plasma (Anode)	physikalisch - chemisch im Plasma (Kathode)	physikalisch - chemisch im Plasma (Kathode)	physikalisch - chemisch im Plasma (Kathode)
Substratposition	Scheiben	Scheiben				
Druck (Torr)	0.5 - 1	0.1 - 0.5	$10^{-1} - 3 \times 10^{-2}$	$10^{-3} - 10^{-1}$	$10^{-1} - 10^1$	$10^2 - 1$
Flußrate (sccm)	$10^2 - 10^3$	50 - 500	$10 - 10^2$	$1 - 10^2$	$1 - 10^2$	$1 - 10^2$
Ionenenergie (eV)	10	thermisch (keine Ionen)	10 - 100	$10^2 - 10^3$	50 - 200	$10^2 - 10^3$
Beschußrichtung	undefiniert	undefiniert	undef. - bevorzugt senkr.	bevorzugt senkrecht	bevorzugt senkrecht	bevorzugt senkrecht
Ätzspezies	Neutralteilchen, Ionen	Neutralteilchen	Neutralteilchen, Ionen	Ionen, Neutralteilchen	Ionen, Neutralteilchen	Ionen, Neutralteilchen
Frequenz	13.56 MHz	2.45 GHz, 13.56 MHz	< 1 MHz	13.56 MHz, < 1 MHz	13.56 MHz	13.56 MHz, 1 MHz, DC
Anisotropie	isotrop	isotrop	gering - hoch	mittel - hoch	mittel - hoch	mittel - hoch
Selektivität	hoch	hoch	mittel - hoch	gering - mittel	gering	gering - mittel
Ätzrate	niedrig	mittel - hoch	hoch	niedrig	hoch	mittel - hoch
Homogenität	gering	mittel	gering - mittel	mittel - hoch	gering - mittel	mittel - hoch
Schädigung	gering	keine	gering	mittel - hoch	gering - mittel	mittel - hoch
Kontamination	gering - mittel	gering	mittel - hoch	mittel - hoch	gering - mittel	mittel - hoch
Anwendung	isotropes Entschichten von Si und Si_3N_4 auf Vorder- und Rückseite der Scheiben, Lackveraschen	isotropes Strukturieren und Entschichten von Si, Si_3N_4 Lackveraschen	Strukturierung von Poly Si, Si_3N_4, SiO_2	Strukturierung von Poly Si, Mono Si, Si_3N_4, SiO_2 Silizide, Polyzide, AlSiCu	Planarisierung Tri-Level-Maskenätzung Sputterätzen	Strukturierung im Sub-µm-Bereich von Mono-Si (trench), AlSiCu
Anlagenbeispiel	IPC L2101	CDE-8	LAM Autoetch 490, PK 24 PE (Plasmatherm)	HIRRIE 100, Al-E 8100 (Hexcode)	MFC Aries	Tegal 1500

Tabelle 4: Die wichtigsten Trockenätzverfahren und ihre charakteristischen Eigenschaften /14/