



Markus Reinl (Autor)

Untersuchungen der Kontakteigenschaft beim Testen von integrierten Schaltungen



<https://cuvillier.de/de/shop/publications/1207>

Copyright:

Cuvillier Verlag, Inhaberin Annette Jentsch-Cuvillier, Nonnenstieg 8, 37075 Göttingen,
Germany

Telefon: +49 (0)551 54724-0, E-Mail: info@cuvillier.de, Website: <https://cuvillier.de>

Kapitel 1

Einleitung

Immer mehr elektronische Geräte kommen im täglichen Leben zum Einsatz. Die rasante Entwicklung lässt sich am Beispiel der Telekommunikation gut erkennen. Werden die ersten Generationen an mobilen Telefongeräten mit den heutigen Geräten verglichen, sind nicht nur bei Größe und Gewicht, sondern auch im Funktionsumfang erhebliche Unterschiede zu erkennen. Die Geräte werden mit jeder Generation immer leistungsfähiger. Diese Entwicklung ist auf dem gesamten Bereich der Anwenderelektronik zu erkennen. Die Ursache dafür ist, dass die integrierten Schaltkreise¹ immer komplexer werden und mehr Transistoren pro Schaltung zum Einsatz kommen. Um dabei die voranschreitende Verkleinerung der Geräte erreichen zu können, müssen die Strukturgrößen der einzelnen Transistoren kontinuierlich reduziert werden.

Bereits im Jahre 1965 machte *Gordon E. Moore*, Mitbegründer von Intel, die Beobachtung, dass sich die Anzahl der Transistoren pro Fläche seit Erfindung des integrierten Schaltkreises jährlich verdoppelt hat [1]. In seiner Veröffentlichung, die später als "Moore'sche Gesetz" bekannt wurde, machte er die Aussage, dass dieser Trend die nächsten zehn Jahre anhalten wird. Dieser Trend hat sich jedoch abgeschwächt und wurde 1975 auf eine Verdoppelung der Transistoren alle 24 Monate korrigiert. Das "Moore'sche Gesetz" wird aktuell in leicht abgeänderter Form ausgelegt. Heute wird von einer Verdoppelung der Integrationsdichte alle 18 Monate gesprochen. Diese Aussage stammt jedoch nicht von *Moore*, sondern von einem Intel-Mitarbeiter. Das "Moore'sche Gesetz" besitzt heute noch seine Gültigkeit und diese wird auch für die kommenden Jahre prognostiziert.

Basierend auf dem "Moore'schen Gesetz" wird die *International Technology Roadmap for Semiconductors*² von einem Gremium aus Experten der Halbleiterindustrie erstellt [2]. Diese gibt vor, welche Maßnahmen ergriffen werden müssen, um mit den kommenden Produkttechnologien das "Moore'sche Gesetz" einzuhalten. Sowohl Halbleiterhersteller, als auch die Hersteller von Produktionsanlagen der Halbleiterindustrie halten sich an diese Roadmap. Speziell bei den Strukturgrößen wird man in absehbarer Zeit an technologische, aber vor allem an physikalische Grenzen stoßen, ab denen eine weitere Strukturverkleinerung mit den jetzigen Transistor-konzepten nicht mehr sinnvoll ist.

¹engl. integrated circuit (IC)

²kurz: ITRS

Eine Verkleinerung der Strukturgrößen ist immer mit einem immensen Investitionsvolumen für neue Fertigungsanlagen, speziell in der Lithographie, verbunden. Aus diesem Grund wird bereits das Design der Chips dahingehend optimiert, dass die Fläche pro Chip möglichst minimiert und effektiv genutzt wird. Bei bisherigen Produkttechnologien wird die Fläche unter den sog. "Probing Pads" nicht für aktive Strukturen benutzt. Als "Probing" wird das Testen der einzelnen Chips mittels einer Testkarte³ bezeichnet. Die "Probing Pads" dienen als Kontaktflächen zur Testkarte und sind an den Rändern der Chipoberfläche angeordnet. Diese Positionierung der Pads bedeutet einen höheren Platzbedarf pro Chip und somit einen geringeren Gewinn, da die Kosten pro Fläche Silizium, unabhängig von der zum Einsatz gebrachten Wafergröße, als nahezu konstant angesehen werden können [3,4]. Um die benötigte Fläche bei aktuellen Technologien zu reduzieren, werden die Probing Pads direkt über aktiven Strukturen angeordnet.

Das Hauptproblem beim Testen über aktivem Schaltungsgebiet⁴ ist der Krafteintrag in die Schichten unter den Probing Pads. Durch diesen Krafteintrag können die Schichten darunter beschädigt werden, was zum sofortigen oder verzögerten Ausfall des ganzen Chips führen kann. Die Nadelkarten stellen die am häufigsten verwendete Testkartenart dar. Bei diesem Typ wird der elektrische Kontakt zur Padoberfläche mittels Testnadeln hergestellt. Dabei lässt sich eine starke mechanische Deformation der Oberfläche feststellen. In Abbildung 1.1 ist eine REM⁵-Aufnahme eines Abdrucks einer Nadelkarte, verursacht durch das Probing, zu sehen.

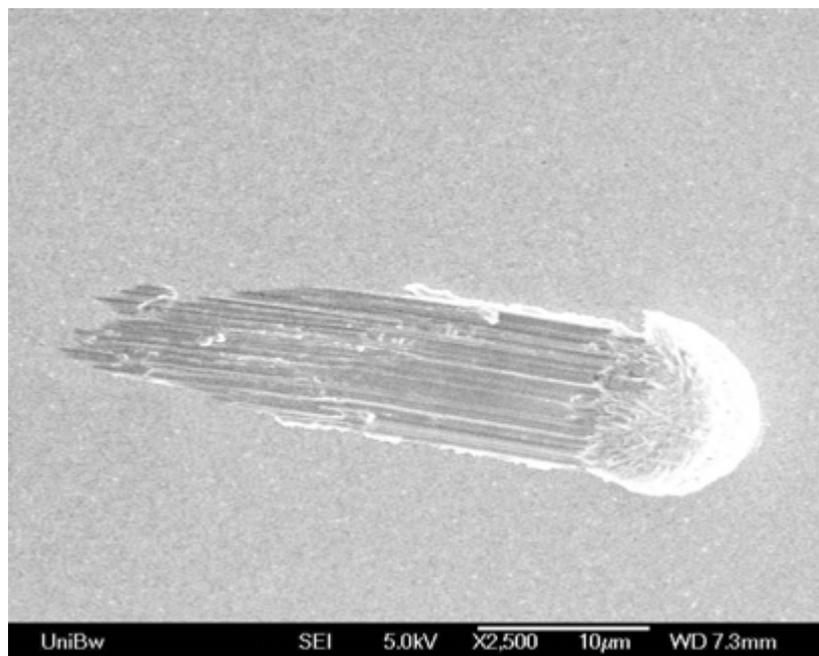


Abbildung 1.1: REM-Aufnahme des Abdrucks einer Nadelkarte auf einer Aluminiumoberfläche nach dem Probing.

Um den Mechanismus, der zum Ausfall der Chips beim POAA führt, zu verstehen, muss zuerst die Kontaktphysik zwischen Testnadel und Padoberfläche ver-

³engl. probe card

⁴engl. probing over active area (POAA)

⁵Abk. für Rasterelektronenmikroskop, engl. scanning electron microscopy (SEM)

standen werden. Zu diesem Zweck werden die einwirkenden mechanischen Kräfte zwischen Nadel und Padoberfläche in Abhängigkeit vom elektrischen Widerstand genauer untersucht.

Im Rahmen der vorgelegten Arbeit wird ein entsprechender Messplatz entwickelt, mit dessen Hilfe die relevanten Parameter, wie die Kontaktkräfte – lateral und normal – und der elektrische Widerstand, während des Probing mittels einer einzelnen Nadel gemessen werden können. Mit Hilfe dieses Messplatzes werden Messungen auf unterschiedlichen Materialsystemen angefertigt. Aus den Ergebnissen können Schlussfolgerungen bezüglich der Probleme beim Probing gezogen werden. Zusätzlich werden Reinigungsverfahren vorgestellt, die die Performance des Probing verbessern. Ein besonderes Augenmerk liegt dabei auf der industriellen Nutzbarkeit dieser Reinigungsprozesse.

Kapitel 2 gibt zunächst einen Überblick über den Aufbau von Probing Pads und geht auf die möglichen Probleme und Ausfallmechanismen beim POAA ein und erläutert diese. Dabei wird die Notwendigkeit für einen eigenständigen Messplatz zur Messung der einwirkenden Kräfte und des elektrischen Widerstandes aufgezeigt.

Mit dem Aufbau des Messplatzes beschäftigt sich **Kapitel 3**. Die einzelnen Vorteile der entwickelten Variante im Vergleich zu bestehenden Messplätzen werden erläutert. Dadurch kann eine genauere Nachbildung des Probing – wie es in der Industrie vollzogen wird – erreicht werden.

In **Kapitel 4** wird auf die Theorie des elektrischen Kontaktwiderstandes eingegangen und dieser näher betrachtet. Dadurch lassen sich die Messergebnisse, vorgestellt in **Kapitel 5**, besser einordnen. Es werden Untersuchungen präsentiert, die den Einfluss des natürlichen Metalloxides auf den elektrischen Kontaktwiderstand verdeutlichen. Reinigungsverfahren mit Wasserstoff und deren Auswirkungen auf die Performance beim Probing werden näher erläutert. Durch die Untersuchungen zeigt sich, dass das natürliche Metalloxid nicht der alleinige Grund für einen schlechten Kontaktwiderstand ist. Deshalb wird in **Kapitel 6** der Einfluss von Verunreinigungen auf der Padoberfläche untersucht. Diese entstehen aufgrund des Prozessablaufs und haben einen höheren elektrischen Kontaktwiderstand und eine höhere mechanische Krafteinwirkung zur Folge.

In **Kapitel 7** werden mehrere mögliche Reinigungsverfahren hinsichtlich der Entfernung der Verunreinigungen auf der Padoberfläche untersucht. Ein besonderes Augenmerk liegt dabei auf der Entfernung von Fluorkontaminationen und der Evaluierung eines möglichst industrietauglichen Reinigungsverfahrens.

Eine Zusammenfassung und ein Ausblick wird in **Kapitel 8** gegeben. Dabei wird auf mögliche weitere Untersuchungen eingegangen.

Kapitel 2

Stand der Technik

Die Produktion von integrierten Schaltungen für den Massenmarkt ist durch den großen Bedarf an Hightech-Fertigungsanlagen und der benötigten Infrastruktur sehr kostenintensiv. Um am Markt wirtschaftlich agieren zu können, versuchen die Halbleiterhersteller permanent die Produktionskosten möglichst gering zu halten bzw. durch Einführung neuer Technologien weiter zu reduzieren. Dies geschieht hauptsächlich im FEOL⁶ bzw. BEOL⁷, da es sich hierbei um die kostenintensivsten Bereiche bei der Herstellung von integrierten Schaltungen handelt. Als FEOL wird der Bereich der Produktion bezeichnet, in dem die unteren Strukturierungsebenen auf dem Wafer entstehen, welche die Funktionsblöcke aus Dioden, Transistoren und passiven Bauelementen der Schaltung enthalten. Im BEOL werden die einzelnen Verdrahtungsebenen der Schaltung erzeugt und somit die Bauelemente zu Funktionsblöcken miteinander verbunden. Das BEOL endet mit der Abscheidung einer Passivierungsschicht auf der Oberfläche des Wafers.

Der darauf folgende Teil der Herstellung wird als Back-end bezeichnet. Dieser beinhaltet das Testen der auf dem gesamten Wafer befindlichen Schaltungen, Sägen der Wafer in die einzelnen Rohchips⁸ und die Aufbau- und Verbindungstechnik⁹. Nach dem Packaging werden die fertigen Chips noch einmal geprüft, um mögliche Beschädigungen während des Prozesses zu entdecken und die fehlerhaften Chips auszusortieren. In diesem Bereich wird ebenfalls verstärkt versucht, Kosten einzusparen.

Einen wichtigen Schritt im Back-end stellt das Testen der vollständig prozessierten Wafer dar, das im Englischen als "Probing" bezeichnet wird. Wie sich an dem Auszug aus der IRTS in Abbildung 2.1 erkennen lässt, wird in den kommenden Jahren die Anzahl der Anschlüsse¹⁰ pro Chip weiter steigen. Die Kosten für das Packaging pro Pin werden demgegenüber im gleichen Maß fallen. Die Kosten pro Chip können folglich als konstant betrachtet werden. Da die Anzahl der Pins pro Chip immer weiter steigt, nimmt automatisch die Ausfallwahrscheinlichkeit aufgrund von Beschädigungen beim Testen zu. Daher ist es um so wichtiger, fehlerhafte Schal-

⁶engl. front-end-of-line

⁷engl. back-end-of-line

⁸engl. die

⁹engl. packaging

¹⁰engl. pin

tungen noch vor dem Vereinzeln der Wafer und dem Packaging zu identifizieren und auszusondern.

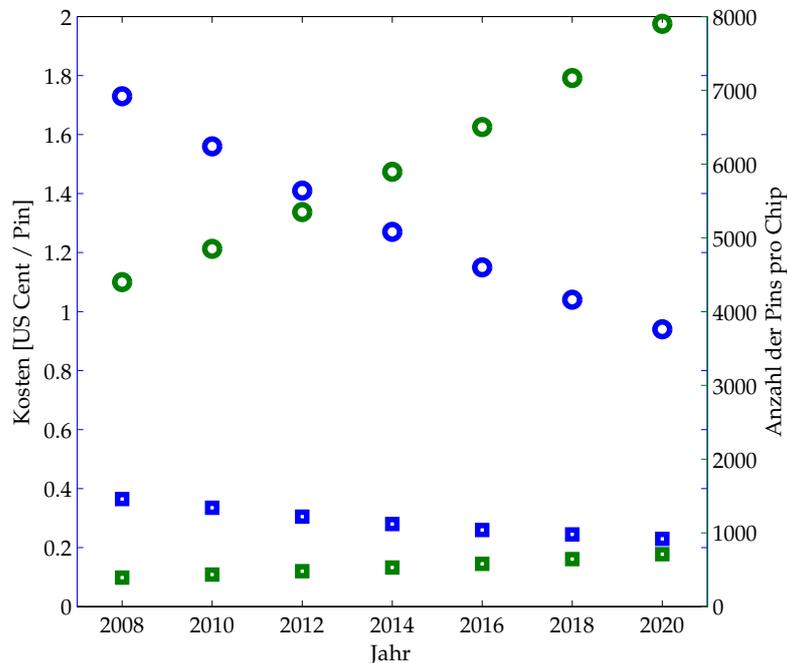


Abbildung 2.1: Daten aus der ITRS über die Kosten für das Packaging pro Pin (linke Skala - blau) bzw. die Anzahl der Pins pro Chip (rechte Skala - dunkelgrün). Exemplarisch sind die Daten für Low-cost Produkte (Quadrate) und für High-performance Produkte (Kreise) dargestellt [2].

Um die einzelnen Schaltungen zur Überprüfung ihrer Funktionsfähigkeit elektrisch kontaktieren zu können, werden sog. Testkarten¹¹ eingesetzt. Es gibt unterschiedliche Typen von Testkarten, die zum Einsatz kommen, beispielsweise Nadelkarten, Membran-Karten oder Microspring-Karten. Die einzelnen Kartentechnologien werden in Abschnitt 2.2 vorgestellt. Im Weiteren werden nur die Cantilever-Karten, die eine Variante der Nadelkarten darstellen, betrachtet. Dieser Kartentyp ist in der industriellen Produktion sehr weit verbreitet, da die Karten kostengünstig herzustellen und leicht zu warten sind [5,6]. Der elektrische Kontakt zwischen Testkarte und Chip wird erreicht, indem die Nadeln auf die Testpads¹² aufgesetzt werden. Dazu wird der Wafer von unten gegen die Probe Card gedrückt. Die Anordnung der Nadeln ist dem entsprechenden Layout der Probing Pads angepasst.

In Abbildung 2.2 ist die Oberfläche eines bereits getesteten Wafers zu sehen. Aktuell setzt die Industrie Nadelkarten ein, die mit bis zu 25.000 Nadeln bestückt sind. Mit diesen Karten ist es ausreichend, einen 300 mm-Wafer viermal zu kontaktieren, um alle Schaltungen überprüfen zu können [7]. Die Nadelkarten stellen zwar die am weitesten verbreitete Art dar, besitzen aber mit allen anderen Kartentypen eine Gemeinsamkeit. Der elektrische Kontakt wird durch eine mechanische Kontaktierung erreicht.

¹¹engl. probe card

¹²engl. probing pad

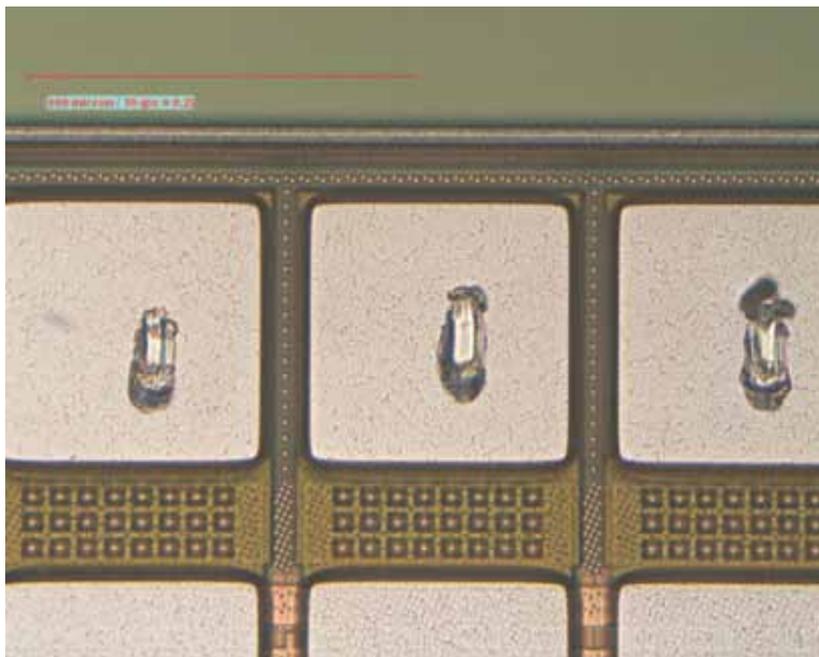


Abbildung 2.2: Lichtmikroskopische Aufnahme einzelner Probing Pads eines vollständig prozessierten Wafers.

Wie in Abbildung 2.2 zu erkennen ist, erzeugt das Probing eine Deformation der Padoberfläche, die im Englischen als “probe mark“ oder “scrub“ bezeichnet wird. Diese Verformung der Oberfläche wird durch den mechanischen Krafteintrag der Nadel erzeugt. Zusätzlich wird ein Teil dieser Kraft in die Materialschichten unter dem Pad eingeleitet und erzeugt dort mechanischen Stress. Dies kann zum Ausfall der getesteten Schaltung¹³ führen und reduziert somit die Ausbeute¹⁴ an funktionierenden Chips. Um Möglichkeiten zu finden dies zu vermeiden, müssen die Mechanismen, die hinter den Ausfällen stehen, verstanden werden. Aus diesem Grund wird in diesem Kapitel der Vorgang des Probing näher beschrieben und untersucht.

Mit der folgenden Abschätzung soll dargelegt werden, dass Probleme beim Probing nicht nur durch mechanische Belastungen verursacht werden. Eine Verschlechterung des Kontaktwiderstandes zwischen Kontaktnadel und Probing Pad kann dazu führen, dass die Schaltung nicht mehr mit einer ausreichenden Versorgungsspannung betrieben wird. Dadurch kann die getestete Schaltung den Funktionstest nicht erfolgreich absolvieren und wird als fehlerhaft aussortiert. Aktuelle ICs werden mit einer Versorgungsspannung von $V_{DD} = 1,2\text{ V}$ betrieben. Wird ein typischer Strom von $I = 100\text{ mA}$ angenommen, bedeutet ein Anstieg des Kontaktwiderstandes um $\Delta R = 1,0\ \Omega$, dass die Schaltung nur noch mit einer Spannung von $V = 1,1\text{ V}$ versorgt wird. Dies liegt außerhalb der Spezifikationen für die Variation der Versorgungsspannung, die mit $\Delta V_{DD} = +10\%/ -5\%$ angegeben ist. Wie die späteren Messungen zeigen werden, ist ein Anstieg des Kontaktwiderstandes um $\Delta R = 1\ \Omega$ aufgrund von Verunreinigungen durchaus möglich. Daher ist es erfor-

¹³engl. device under test (DUT)

¹⁴engl. yield

derlich einen möglichst geringen Kontaktwiderstand zwischen Nadel und Pad zu erzielen.

Zunächst wird ein Überblick über die verwendeten Geräte beim Testen gegeben. Ferner werden der Schichtaufbau der Probing Pads näher betrachtet und mögliche Probleme aufgezeigt, die sich beim Testen über aktivem Schaltungsgebiet¹⁵ ergeben.

2.1 Testequipment

Ein vollständiges Testsystem besteht aus drei einzelnen Systemen – dem Wafer Prober, dem Testkopf und der eigentlichen Testeinheit. Häufig wird der Begriff Wafer Prober als Synonym für ein komplettes Testsystem verwendet.

- Der **Wafer Prober** übernimmt das komplette mechanische Handling der Wafer während des Testens. Dies beinhaltet den Transport der Wafer innerhalb der Testanlage auf den Wafer Chuck¹⁶, sowie dessen genaue Ausrichtung gegenüber der Testkarte, damit die Nadeln und Probing Pads genau deckungsgleich zueinander positioniert sind. Zum Testen wird der Chuck nach oben angehoben und gegen die Testkarte gedrückt, um den elektrischen Kontakt herzustellen.
- Der **Testkopf** stellt das Interface zur Testeinheit dar, in das die Testkarte eingesetzt wird. Die Aufnahme für die Karten ist dabei standardisiert, um einen einfachen Austausch der Karten zu ermöglichen. Der Testkopf dient also zur Kontaktierung der Wafer und als elektrisches Verbindungsglied zwischen Wafer und Testeinheit.
- Die **Testeinheit**¹⁷ stellt den eigentlichen Tester des Systems dar. Diese übernimmt die Ausführung der Testprogramme und die anschließende Auswertung der Messdaten bzw. -werte. Auf Grundlage dieser Messungen und der vorgegebenen Spezifikationen wird entschieden, ob eine Schaltung funktionsfähig ist oder nach dem Zersägen des Wafers aussortiert wird.

2.2 Testkartentechnologien

Testkarten werden in unterschiedlichen Varianten und Bauformen in der Industrie verwendet. Beispiele hierfür sind Cantilever- [8–11], vertikale [12–14], Membran- [15–17], MEMS- [18–20] oder Microspring-Karten [21–23]. Die Kartentypen unterscheiden sich dabei in ihren elektrischen und mechanischen Eigenschaften, wodurch sich jeweils andere Einsatzgebiete ergeben. Zusammenfassungen unterschiedlicher Testkartentechnologien sind bei *Weeden* [24] und *Mann* [25] zu finden.

¹⁵engl. probing over active area (POAA)

¹⁶Als Wafer Chuck wird eine planare Haltevorrichtung bezeichnet, die die Wafer mittels Unterdruck oder elektrostatischer Anziehung fixiert.

¹⁷engl. automatic test equipment (ATE)