



(19)  
**Bundesrepublik Deutschland**  
**Deutsches Patent- und Markenamt**

(10) **DE 10 2005 030 661 B4 2007.01.25**

(12)

## Patentschrift

(21) Aktenzeichen: **10 2005 030 661.6**  
 (22) Anmeldetag: **30.06.2005**  
 (43) Offenlegungstag: **07.12.2006**  
 (45) Veröffentlichungstag  
 der Patenterteilung: **25.01.2007**

(51) Int Cl.<sup>8</sup>: **G11C 16/28 (2006.01)**

Innerhalb von drei Monaten nach Veröffentlichung der Patenterteilung kann nach § 59 Patentgesetz gegen das Patent Einspruch erhoben werden. Der Einspruch ist schriftlich zu erklären und zu begründen. Innerhalb der Einspruchsfrist ist eine Einspruchsgebühr in Höhe von 200 Euro zu entrichten (§ 6 Patentkostengesetz in Verbindung mit der Anlage zu § 2 Abs. 2 Patentkostengesetz).

(30) Unionspriorität:  
**11/145,520 03.06.2005 US**

(73) Patentinhaber:  
**Infineon Technologies AG, 81669 München, DE;**  
**Infineon Technologies Flash GmbH & Co. KG,**  
**01099 Dresden, DE**

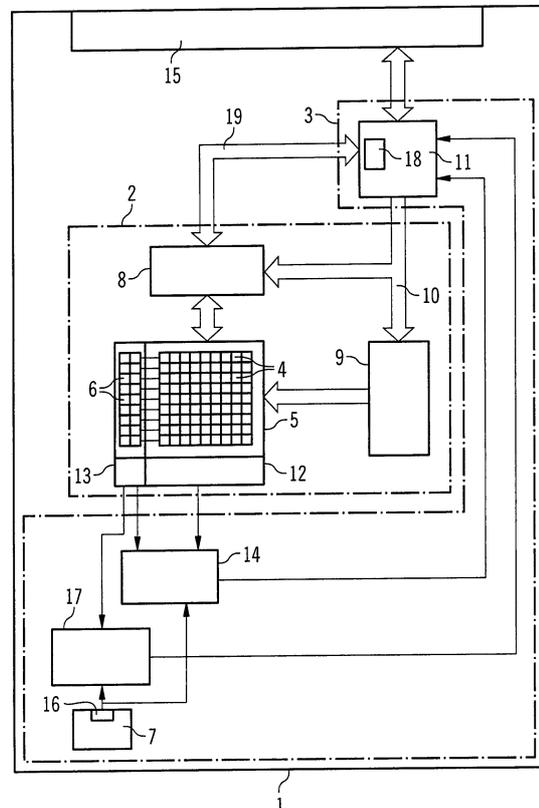
(74) Vertreter:  
**Epping Hermann Fischer,**  
**Patentanwaltsgesellschaft mbH, 80339 München**

(72) Erfinder:  
**Redaelli, Marco, Carate Brianza, Mailand/Milano,**  
**IT; Ambroggi, Luca de, 01099 Dresden, DE**

(56) Für die Beurteilung der Patentfähigkeit in Betracht  
 gezogene Druckschriften:  
**US2002/01 18 566 A1**  
**US 60 55 187 A**  
**US 60 11 725 A**  
**US 51 72 338 A**

(54) Bezeichnung: **Nichtflüchtiges Halbleiterspeicherbauelement und Verfahren zum Betreiben und Herstellen eines nichtflüchtigen Halbleiterspeicherbauelementes**

(57) Hauptanspruch: Nichtflüchtiges Halbleiterspeicherbauelement (1), das Folgendes umfasst:  
 einen Speicherbereich (2) und einen Schaltungsbereich (3),  
 wobei der Speicherbereich Folgendes umfasst:  
 – mehrere in einem Speicherarray (5) angeordnete Speicherzellen (4) und  
 – eine Menge von programmierbaren Arrayreferenzzellen (6), die als Referenzen zum Lesen der Speicherzellen (4) vorgesehen sind,  
 wobei der Schaltungsbereich (3) Folgendes umfasst:  
 – eine Menge von Hauptreferenzzellen (7), die als Referenzen zum Feststellen des Zustands von Arrayreferenzzellen (6) oder Speicherzellen (4) vorgesehen sind.



**Beschreibung**

## Stand der Technik

**[0001]** Die vorliegende Erfindung betrifft nichtflüchtige Halbleiterspeicherbauelemente. Sie betrifft insbesondere ein Speicherbauelement mit Referenzzellen, die so aufgebaut sind, dass sie eine Referenz liefern, die dafür ausgelegt ist, immer zu Eigenschaften von Speicherzellen zu passen, selbst nach einer Verschlechterung, die durch eine lange Arbeitszeit oder viele Lese- oder Schreibzyklen verursacht wird. Die Erfindung betrifft außerdem ein Verfahren zum Programmieren von Referenzzellen.

**[0002]** Flashspeicher sind in jüngster Zeit immer populärer geworden, insbesondere im Bereich von tragbaren Kommunikationseinrichtungen. Die grundlegende Struktur eines Flashspeichers ist der eines MOSFET ähnlich, einschließlich einer Gateelektrode, einer Drainelektrode und einer Sourceelektrode. Üblicherweise enthält der Flashspeicher als die Gateelektrode des MOSFET eine floatende Gateelektrode und eine Steuergateelektrode. Außerdem gibt es einige Arten von Flashspeichern ohne floatende Gateelektrode, wie etwa einen Nitridfestwertspeicher (NROM). Ein Nitridfestwertspeicher unterscheidet sich von anderen Arten von Flashspeichern, die eine leitende Polysilizium- oder metallische floatende Gateelektrode verwenden, durch die Verwendung einer Oxid-Nitrid-Oxid-Schicht (ONO) als ladungsfangendes Medium. Auf Grund einer stark kompaktierten Natur der Siliziumnitridschicht werden von dem MOS-Transistor in das Siliziumnitrid tunnelnde heiße Elektroden eingefangen, um eine ungleiche Konzentrationsverteilung zu bilden.

**[0003]** Im Allgemeinen weist der Flashspeicher die Funktionen Lesen, Programmieren und Löschen auf. Wenn Elektronen in die floatende Gateelektrode der Speicherzelle oder Elektronen in die ONO-Schicht der Speicherzelle injiziert werden, nimmt eine Schwellenspannung, anfänglich mit einer niedrigen Spannung, der Speicherzelle relativ zu und führt zu einem abnehmenden Strom von der Drainelektrode zu der Sourceelektrode. Dies ist der geschriebene oder programmierte Zustand der Speicherzelle. Wenn eine negative Spannung an die Steuergateelektrode angelegt wird, werden in der floatenden Gateelektrode (oder in der ONO-Schicht) gefangene Elektronen aus der floatenden Gateelektrode oder der Nitridschicht entfernt, um die Schwellenspannung der Speicherzelle zu senken. Dies ist der gelöschte Zustand. Ungeachtet des Zustands, in dem sich die Speicherzelle befindet, ist es erforderlich, einen Lesevorgang durchzuführen, währenddessen die in der Speicherzelle gespeicherten Bitinformationen gelesen werden.

**[0004]** Zum Lesen von in einer Speicherzelle ge-

speicherten Informationen wird die Spannung an dem Gateeingang erhöht, bis ein vordefinierter Strom von der Drainelektrode zur Sourceelektrode der Speicherzelle fließt. Die tatsächliche Gatespannung wird dann ausgewertet. Alternativ wird der Strom ausgewertet, während eine vordefinierte Spannung an die Gateelektrode angelegt wird. In beiden Fällen wird das Leitungsfenster der Speicherzelle durch einen Schwellwert zum Definieren des Zustands der Speicherzelle in zwei Gebiete unterteilt.

**[0005]** Genauer gesagt sind für das Lesen eines Zustands einer Speicherzelle unter Verwendung eines Referenzstroms zwei Mechanismen üblich. Bei einem ersten Mechanismus wird eine Zelle gelesen, indem an dem Gate- und dem Draineingang vorbestimmte, feste Spannungen angelegt werden. Ihr Drain-/Sourcestrom wird auf einen Speicherzustand abgebildet, indem er mit einem Referenzstrom verglichen wird. Wenn der gelesene Strom höher ist als die Referenz, dann wird bestimmt, daß sich die Zelle in einem Logikzustand befindet (beispielsweise in einem L-Zustand). Wenn andererseits der Strom kleiner ist als der Referenzstrom, wird bestimmt, daß sich die Zelle in dem anderen Logikzustand befindet (beispielsweise einem H-Zustand). Somit speichert eine derartige Zweizustandszelle ein Bit digitaler Informationen.

**[0006]** Ein zweiter Mechanismus zum Lesen des Zustands einer Speicherzelle besteht darin, die Gateelektrode der Speicherzelle anstatt mit einer konstanten Spannung mit einer variablen Spannung vorzuspannen. Hier wird der Drain-/Sourcestrom gelesen und mit einem konstanten Strom verglichen. Die Gatespannung, bei der der konstante Referenzstrom von dem gemessenen Strom erreicht wird, zeigt den Zustand der Speicherzelle an. Zum Programmieren und Löschen von Speicherzellen sind ähnliche Operationen erforderlich, so genannte Bestätigungsoperationen. Bestätigungsoperationen treten während des Programmierens oder Löschens von Speicherzellen auf, es handelt sich bei ihnen um Leseoperationen während Schreiboperationen, die eine Notwendigkeit eines Programmier- oder Löschimpulses bewerten, damit diejenigen Daten, die in die Zelle geschrieben werden sollen, ordnungsgemäß geschrieben werden.

**[0007]** Um die Speicherkapazität zu erhöhen, werden Flash-EEPROM-Bauelemente (elektrisch löschbare programmierbare Festwertspeicher) mit einer Dichte hergestellt, die mit dem Fortschritt der Halbleitertechnologie immer höher liegt. Ein weiteres Verfahren zum Erhöhen von Speicherkapazität besteht darin, dass jede Speicherzelle mehr als zwei Zustände speichert.

**[0008]** Für eine Mehrfachzustands- oder Mehrfachpegel-EEPROM-Speicherzelle wird das Leitungs-

fenster durch mehr als einen Schwellwert in mehr als zwei Gebiete unterteilt, so dass jede Zelle in der Lage ist, mehr als ein Datenbit zu speichern. Somit nehmen die Informationen, die ein gegebenes EEPROM-Array speichern kann, mit der Anzahl der Zustände zu, die jede Zelle speichern kann. Ein EEPROM oder ein Flash-EEPROM mit Mehrfachzustands- oder Mehrfachpegelspeicherzellen sind im US-Patent Nr. 5,172,338 beschrieben worden.

**[0009]** Eine weitere bekannte Möglichkeit zum Erhöhen der Speicherdichte besteht darin, mehr als ein Bit zu speichern, wobei unterschiedliche Mengen elektrischer Ladung auf einer leitenden Schicht nicht unterschieden werden, sondern in verschiedenen Bereichen einer Gateschicht gespeicherte elektrische Ladungen. Eine derartige Speicherzelle ist von den oben erwähnten Nitridfestwertspeichern (NROM) bekannt. Aus US 2002/0118566 ist bekannt, wie in Nitridfestwertspeicherzellen Zwei-Bit-Informationen gleichzeitig gelesen werden. Gemäß dem Drain-Source-Strom des NROM kann eine logische Zwei-Bit-Kombinationsnachricht identifiziert werden. Der beobachtete Strom wird in vier verschiedene Zonen unterteilt, und jede Zone stellt eine spezifische logische Zwei-Bit-Information dar, nämlich L und L, L und H, H und L oder H und H.

**[0010]** Das Speichern von zwei Bits an Informationen in verschiedenen Gebieten der Nitridschicht weist den Vorteil auf, dass der Unterschied zwischen den individuellen Zuständen leichter detektiert werden kann als in einer Mehrfachpegelspeicherzelle. Der Zustand des zweiten Bit beeinflusst jedoch den Strom, der durch die Zelle fließt, wenn der Zustand des ersten Bit detektiert wird, und umgekehrt. Dies wird auch als der Zweite-Bit-Effekt bezeichnet. Der beschriebene Effekt ist relevant beim Lesen von Daten, aber auch beim Bestätigen von Daten während Löscho- oder Programmieroperationen, und muss deshalb bei der Auswertung eines detektierten Zellenstroms berücksichtigt werden. Obwohl der Zweite-Bit-Effekt im Vergleich zu dem Effekt gering ist, der durch den Speicherzustand des zu lesenden Bits verursacht wird, kann er an Wichtigkeit zunehmen, wenn die Arbeitsspannung der Speicherzelle niedriger wird. Damit Leistung gespart wird und kleinere Einzelchipstrukturen für eine Halbleiterschaltung verwendet werden können, sinken die Arbeitsspannungen von Speichermodulen immer mehr. Wohingegen bisher 5 V und 3,3 V als Versorgungsspannungen verwendet wurden, verwenden neue Einrichtungen Spannungen von beispielsweise 1,6 V.

**[0011]** Zum Lesen dieser Flashspeicherzellen ist es wesentlich, den Leitungsstrom über die Source- und Drainelektrode der Zelle hinweg zu erfassen. Insbesondere zum Lesen von mehr als zwei Zuständen einer Speicherzelle ist es wichtig, einen Referenzstrom zu haben, der den Zustand der Speicherzelle genau

wiedergibt. Je mehr Zustände, für deren Speicherung eine Speicherzelle ausgelegt ist, unterschieden werden sollen, desto feiner muss ihr Gebiet durch Schwellwerte unterteilt werden. Dies erfordert höhere Präzision bei Programmier- und Leseoperationen, damit die erforderliche Auflösung erzielt werden kann.

**[0012]** Die verwendeten Referenzströme werden oftmals von Referenzzellen erzeugt, die sich in einem bestimmten Zustand befinden. In den meisten der früheren Einbit- oder Einpegel-Speicherarchitekturen wurde die Referenzstruktur zum Bereitstellen entsprechender Referenzströme durch ein Array aus vier oder fünf Flashzellen gebildet, die einem programmierten Zustand, einem gelöschten Zustand, einem überlöschten oder Verarmungszustand und einem Lesezustand entsprechen.

**[0013]** Diese Referenzzellen könnten, nachdem sie beispielsweise bei der Wafersortierung auf ein vordefiniertes Niveau programmiert wurden, überhaupt nicht mehr geändert werden.

**[0014]** Aus der US 6 055 187 A ist ein nicht-flüchtiger Halbleiterspeicher bekannt, der außer einer Referenzzelle für den normalen Lesebetrieb weitere Referenzzellen für den gelöschten, programmierten und verarmten Zustand aufweist.

**[0015]** Bei mehreren Anwendungen, beispielsweise zum Archivieren von Daten, ist es wichtig, dass Daten nach einer langen Zeit oder einer großen Anzahl von Schreibzyklen immer noch korrekt gelesen werden können. Außerdem müssen Temperaturschwankungen berücksichtigt werden. Diese Einflüsse beeinflussen die von einer Speicherzelle gelesenen Ströme. Temperaturschwankungen beeinflussen auch die Referenzströme. Bei typischen Einpegel- und Einbitbauelementen definieren die Toleranzen, die eingestellt sind, um die Zustände einer Speicherzelle zu trennen, große Fenster für die verschiedenen Zustände. Die Fenster sind groß genug, um sicherzustellen, dass Daten unter allen Bedingungen korrekt gelesen oder geschrieben werden. Bei Mehrfachpegel- und Mehrbitspeichern sind die Fenster signifikant reduziert worden, und es ist ein Problem geworden, unter allen Umgebungsbedingungen und über die ganze Lebensdauer hinweg eine korrekte Funktionalität sicherzustellen.

**[0016]** Die jeweilige Notwendigkeit für genauere Referenzen ist in **Fig. 17, 18 und 19** dargestellt. Das Diagramm von **Fig. 17** betrifft eine Zweizustandspeicherzelle zum Speichern eines Datenbit. Es zeigt die Spannungen und Ströme in einer Speicherzelle gemäß **Fig. 7**. Der Strom  $I_{CELL}$  durch die Speicherzelle hängt von der Gate-Source-Spannung ab. Ein niedrigerer Schwellwert **121** definiert den Beginn eines gelöschten Zustands. Unterhalb des gelöschten

Zustands befindet sich auch ein als Verarmungszustand bezeichneter überlöschter Zustand. In diesem Zustand, der beim normalen Betrieb der Speicherzelle nicht gestattet ist, fließt ein Strom selbst dann von der Drainelektrode zur Sourceelektrode, wenn keine Spannung an die Gateelektrode angelegt ist. Der gelöschte und der programmierte Zustand sind durch eine Referenzspannung **122** getrennt. Die Toleranzfenster für den gelöschten und den programmierten Zustand sind groß genug, um für alle Bedingungen zu passen.

**[0017]** Fig. 18 zeigt, dass in einer Mehrfachpegelzelle die Toleranzfenster für alle Zustände reduziert sind. Für jeden der verschiedenen Zustände wird eine Schwellenspannung **121**, **123**, **124** und **125** definiert. Fig. 18, die die möglichen Zustände einer NROM-Zelle zeigt, macht klar, dass die Referenzen **121**, **123**, **124** und **125** präziser sein müssen als die Referenz **121** und **122** in dem Diagramm von Fig. 17.

**[0018]** Aus Fig. 19 kann man erkennen, dass es bei Mehrfachbitspeicherzellen zu einem zusätzlichen Problem kommt. Wenn das erste Bit der Speicherzelle gelesen wird, muss berücksichtigt werden, dass die Kennlinie von dem zweiten Bit abhängt. Ohne Berücksichtigung des zweiten-Bit-Effekts wird eine logische „1“ detektiert, wenn die Gate-Source-Spannung in dem als **126** bezeichneten Bereich liegt. Der Bereich einer logischen „0“ wird als **129** bezeichnet. Eine Schwellenspannung wird als **128** bezeichnet. Wenn der zweite-Bit-Effekt berücksichtigt wird, muss eine logische „1“ selbst dann detektiert werden, wenn die Gate-Source-Spannung höher ist als der Bereich **126**. Der Überschussbereich wird als **127** bezeichnet. Deshalb reduziert die Berücksichtigung des zweiten-Bit-Effekts weiterhin das Toleranzfenster für das Detektieren des Zustands der Speicherzelle.

**[0019]** Zusammenfassend gibt es mehrere Effekte, die erfordern, dass der Referenzstrom oder die Referenzspannung zum Detektieren des Logikzustands einer Speicherzelle präziser sind. Dies ist nicht nur beim ersten Betreiben des Speicherbauelements erforderlich, sondern auch über die ganze Lebenszeit hinweg, wenn verschlechternde Effekte das Verhalten der Speicherzellen verändern.

**[0020]** Außer diesen spezifischen Problemen gibt es eine allgemeine Notwendigkeit für hohe Leistung, hohe Kapazität und hohe Zuverlässigkeit von nichtflüchtigen Speicherbauelementen. Insbesondere gibt es einen Bedarf an kompakten nichtflüchtigen Speicherbauelementen, die ein Speichersystem aufweisen, das Störungseffekte minimiert.

#### Aufgabenstellung

**[0021]** Eine Aufgabe der vorliegenden Erfindung besteht deshalb in der Bereitstellung eines nicht-

flüchtigen Halbleiterspeicherbauelements und eines jeweiligen Verfahrens zum Herstellen und Betreiben eines Speicherbauelements mit einem besseren Referenzverfahren.

**[0022]** Die Aufgabe wird erfindungsgemäß durch ein nichtflüchtiges Halbleiterspeicherbauelement gelöst, das einen Speicherbereich und einen Schaltungsbereich umfasst, wobei der Speicherbereich mehrere in einem Speicherarray angeordnete Speicherzellen und eine Menge von programmierbaren Arrayreferenzzellen, die als Referenzen zum Lesen der Speicherzellen vorgesehen sind, umfasst und wobei der Schaltungsbereich eine Menge von Hauptreferenzzellen umfasst, die als Referenzen zum Feststellen des Zustands von Arrayreferenzzellen oder Speicherzellen vorgesehen sind.

**[0023]** Die vorliegende Erfindung schlägt die Bereitstellung einer Menge so genannter Arrayspeicherzellen vor, die in einem Speicherbereich angeordnet sind, und einer zweiten Menge von Referenzzellen, so genannten Hauptreferenzzellen, die sich in einem Schaltungsbereich befinden.

**[0024]** Genauer gesagt wird ein nichtflüchtiges Halbleiterspeicherbauelement vorgeschlagen, das einen Speicherbereich und einen Schaltungsbereich umfasst, wobei der Speicherbereich mehrere in einem Speicherarray angeordnete Speicherzellen und eine Menge programmierbarer Arrayreferenzzellen umfasst, die als Referenzen zum Lesen der Speicherzellen vorgesehen sind. Der Schaltungsbereich umfasst eine Menge von Hauptreferenzzellen, die als Referenzen zum Bestätigen des Zustands von Arrayreferenzzellen oder Speicherzellen vorgesehen sind.

**[0025]** Die Hauptreferenzzellen sind vorgesehen zum Setzen der Referenzzellen und zum Bestätigen eines Zustands der Speicherzellen. Die Arrayreferenzzellen sind vorgesehen als Referenz zum Lesen der Speicherzellen. Während die Hauptreferenzzellen vor der Auslieferung und deshalb vor Beginn des regelmäßigen Betriebs des Speicherbauelements auf einen spezifischen Zustand gesetzt werden, werden die Arrayspeicherzellen während ihrer Lebensdauer wiederholt programmiert und gelöscht. Deshalb sind die Arrayspeicherzellen der gleichen Beanspruchung ausgesetzt wie die Speicherzellen, und deshalb verschlechtern sie sich auf die gleiche Weise wie die Speicherzellen. Eine Bewegung der Kennlinien der Speicherzellen beeinflusst nicht Lese- oder Programmierergebnisse, da sich die Arrayreferenzzellenkennlinien auf entsprechende Weise bewegen.

**[0026]** Bei einer bevorzugten Ausführungsform der Erfindung umfasst die Menge von Hauptreferenzzellen eine programmierte Hauptreferenzzelle und eine gelöschte Hauptreferenzzelle. Die Menge von programmierbaren Arrayreferenzzellen umfasst Arrayre-

ferenzzellen, die für die Bereitstellung einer programmierten und einer gelöschten Referenz für das Lesen von Daten verwendet werden.

**[0027]** Bei einer zweiten bevorzugten Ausführungsform der Erfindung umfasst die Menge von Hauptreferenzzellen auch eine programmierte Hauptreferenzzelle und eine gelöschte Hauptreferenzzelle. Die Menge von programmierbaren Arrayreferenzzellen umfasst als die einzige Referenzzelle oder ergänzend zu den anderen Referenzzellen eine Lesereferenzzelle die vorgesehen ist, eine Schwellenspannung zu haben, die die Grenze zwischen einem gelöschten Zustand und einem programmierten Zustand einer Speicherzelle definiert und die zum Lesen von Daten von den Speicherzellen verwendet werden kann. Sie kann unter Verwendung der Hauptreferenzzellen als Referenz auf die spezifische Schwellenspannung gesetzt werden.

**[0028]** Bei einer dritten bevorzugten Ausführungsform der Erfindung umfasst die Menge von Hauptreferenzzellen zusätzlich eine Lesereferenzzelle, die vorgesehen ist, eine Schwellenspannung zu haben, die die Grenze zwischen einem gelöschten Zustand und einem programmierten Zustand einer Speicherzelle zu definieren, die vorgesehen ist als Referenz zum Bestätigen des Zustands der in der Menge von Arrayreferenzzellen enthaltenen Lesereferenzzelle.

**[0029]** Gemäß der beanspruchten Erfindung umfasst ein Verfahren zum Betreiben eines Speicherbauelements die folgenden Schritte: Setzen mindestens einer der Arrayreferenzzellen in einen vordefinierten Zustand unter Verwendung mindestens einer der vorprogrammierten Hauptreferenzzellen als Referenz, Lesen des Zustands mindestens einer Speicherzelle unter Verwendung der Arrayreferenzzellen als Referenz und Bestätigen des Zustands mindestens einer Speicherzelle beim Löschen oder Programmieren der Speicherzelle unter Verwendung mindestens einer der Hauptreferenzzellen als Referenz.

**[0030]** Bei bevorzugten Ausführungsformen des Verfahrens gemäß der vorliegenden Erfindung werden die Hauptreferenzzellen und die Arrayreferenzzellen wie oben in Verbindung mit den jeweiligen Speicherbauelementen angedeutet verwendet.

**[0031]** Ein zusätzlicher Vorteil der vorliegenden Erfindung besteht darin, dass nicht nur das Arrayverhalten von gespeicherten Informationen im Lauf der Zeit berücksichtigt wird, die so genannte Erhaltung nach wiederholten Lesevorgängen („retention after-cycling“), sondern auch die durch Schreibzyklen verursachte Beanspruchung der Zellen. Außerdem beeinflussen Umgebungsbedingungen wie etwa die Temperatur sowohl die Referenzzellen als auch die Speicherzellen. Andererseits stehen gute Referenzwerte

von den Hauptreferenzzellen zur Verfügung. Diese werden nicht von der typischen Beanspruchung beeinflusst, denen Speicher- und Arrayreferenzzellen ausgesetzt sind.

**[0032]** Bei einer weiteren verbesserten Ausführungsform der Erfindung wird für jede Wortleitung eine zusätzliche Statuszelle vorgesehen. Bei dieser als „Zeilenschutzreferenz“ bezeichneten Statuszelle wird eine Information von einem ersten Logikzustand gespeichert, der anzeigt, ob eine der jeweiligen Wortleitungen assoziierten Speicherzellen programmiert ist. Ansonsten sind alle der Wortleitung assoziierten Speicherzellen gelöscht, was durch den anderen Logikzustand angezeigt wird.

**[0033]** Ein weiterer Vorteil der vorliegenden Erfindung besteht darin, dass das vorgeschlagene Referenzverfahren auf alle Arten flüchtiger Speicherbauelemente angewendet werden kann.

#### Ausführungsbeispiel

**[0034]** Diese und Vorteile weitere der vorliegenden Erfindung ergeben sich dem Durchschnittsfachmann nach der Lektüre der folgenden ausführlichen Beschreibung der bevorzugten Ausführungsformen, die in den verschiedenen Figuren und Zeichnungen dargestellt ist.

**[0035]** [Fig. 1](#) ist ein Schemadiagramm einer Flashspeicherschaltung gemäß der vorliegenden Erfindung.

**[0036]** [Fig. 2](#) ist ein Flussdiagramm eines Verfahrens gemäß der vorliegenden Erfindung.

**[0037]** [Fig. 3](#) ist ein Flussdiagramm, das ein Verfahren zum Herstellen eines nichtflüchtigen Halbleiterspeichers durch einen Hersteller zeigt.

**[0038]** [Fig. 4](#) ist ein Flussdiagramm, das ein Verfahren zum Betreiben eines nichtflüchtigen Speicherbauelements mit Hauptreferenzzellen und Arrayreferenzzellen zeigt.

**[0039]** [Fig. 5A](#) zeigt ein Flussdiagramm, das sich auf das Auffrischen einer gelöschten Arrayreferenzzelle bezieht.

**[0040]** [Fig. 5B](#) ist ein Flussdiagramm, das sich auf das Auffrischen einer programmierten Arrayreferenzzelle bezieht.

**[0041]** [Fig. 6](#) zeigt eine schematische Zeichnung einer Arrayreferenzzelle und einer Speicherzelle.

**[0042]** [Fig. 7](#) ist eine schematische Zeichnung einer NROM-Speicherzelle.

[0043] [Fig. 8A](#) bis [Fig. 8D](#) zeigen verschiedene Zustände einer Zwei-Bit-NROM-Zelle.

[0044] [Fig. 9](#) zeigt ein Diagramm, das die Verbindung zwischen der Spannung  $V_G$  und dem Strom  $I_D$  für verschiedene Zustände der Speicherzelle veranschaulicht.

[0045] [Fig. 10A](#) bis [Fig. 10D](#) zeigen verschiedene Zustände einer Speicherzelle bei Berücksichtigung des Zweiten-Bit-Effekts.

[0046] [Fig. 11](#) ist eine graphische Darstellung, die eine zeitliche Schwankung einer Schwellenspannung  $V_{TH}$  jeder Speicherzelle darstellt, die sich aus einer Überlöschungsoperation und einer Soft-Programmierungs- und Bestätigungsoperation, für eine Zelle ausgeführt, ergibt.

[0047] [Fig. 12](#) ist ein Flussdiagramm, das die zyklische Beanspruchung von Referenzzellen betrifft.

[0048] [Fig. 13](#) ist ein Flussdiagramm, das einen Blocklöschbefehl darstellt.

[0049] [Fig. 14](#) ist eine graphische Darstellung, die eine zeitliche Schwankung der Schwellenspannung einer gelöschten Referenzzelle, einer programmierten Referenzzelle und einer Speicherzelle darstellt.

[0050] [Fig. 15](#) ist ein Flussdiagramm, das einen Wortprogrammbehehl betrifft.

[0051] [Fig. 16](#) ist ein Flussdiagramm, das eine alternative Implementierung eines Wortprogrammbehehls betrifft.

[0052] [Fig. 17](#), [Fig. 18](#) und [Fig. 19](#) sind Diagramme, die die Notwendigkeit von scharfen Referenzspannungen zeigen.

[0053] Die vorliegende Erfindung und Ausführungsformen werden spezifischer unter Bezugnahme auf die Zeichnungen beschrieben. In den Figuren sind identische oder ähnliche Komponenten mit identischen Zahlen bezeichnet, und ihre Beschreibung wird nicht wiederholt.

[0054] Nichtflüchtige Halbleiterspeicherbauelemente, die elektrisch wiederbeschrieben und gelöscht werden können, werden wie oben erwähnt als EEPROMs und NROMs bezeichnet.

[0055] [Fig. 1](#) ist ein Blockdiagramm, das eine Konfiguration eines Flashspeicherbauelements zeigt, das NROM-Speicherzellen zum Speichern von Daten verwendet. Das Speicherbauelement **1** umfasst einen Speicherbereich **2** und einen Schaltungsbereich **3**. Der Speicherbereich **2** umfasst ein Speicherarray **5** mit mehreren Speicherzellen **4**. Die Speicherzellen

sind in Spalten und Zeilen angeordnet und stellen deshalb eine Matrix dar. Jede der Speicherzellen **4** kann mit Hilfe eines Spaltendecodierers **8** und eines Zeilendecodierers **9** gewählt werden. Beide Decodierer sind mit einem Adressbus **10** verbunden, der auch mit einer Steuereinheit **11** des Schaltungsbereichs **3** verbunden ist. Die Steuereinheit **11** ist mit einer E/A-Schnittstelle **15** zum Empfangen und Senden von Daten von einer externen Einrichtung verbunden. Beispielsweise greift die Steuereinheit **11** zum Lesen von in dem Speicherbauelement gespeicherten Daten auf das Speicherarray **5** zu, indem es eine Adresse an den Spaltendecodierer **8** und den Zeilendecodierer **9** schickt. Der Spaltendecodierer **8** extrahiert Informationen dahingehend, welche der Bitleitungen des Speicherarrays **5** gewählt werden sollen. Der Zeilendecodierer **9** wählt eine Wortleitung entsprechend der von der Steuereinheit **11** übertragene Adresse. Daten werden zwischen der Steuereinheit **11** und dem Speicherarray **5** über einen Datenbus **19** übertragen. Zum Lesen und Schreiben von Daten umfasst die Steuereinheit **11** Schreib- und Lesehaltungen **18**, um Operationen an dem Speicherarray **5** auszuführen.

[0056] Zum Lesen einer Speicherzelle **4** ist es erforderlich, einen durch diese Zelle fließenden Strom mit einem durch eine Referenzzelle fließenden Strom zu vergleichen. Dies geschieht durch Leseverstärker **12** und **13**. Der Leseverstärker **12** ist ein Speicherzellenleseverstärker, der an die Bitleitungen der Speicherzelle **4** angeschlossen ist.

[0057] Gemäß der Erfindung werden Arrayreferenzzellen **6** vorgesehen, die an die gleichen Wortleitungen angeschlossen sind wie die Speicherzellen **4**. Bei einer bevorzugten Ausführungsform der Erfindung ist eine Menge von mindestens zwei Referenzzellen **6** für jede der Wortleitungen vorgesehen. Während des Betriebs des Speicherbauelements **1** wird eine der Referenzzellen **6** so konfiguriert, dass sie einen Schwellwert entsprechend einem gelöschten Zustand einer Speicherzelle **4** aufweist, und mindestens eine andere Referenzzelle **6** wird so konfiguriert, dass sie eine Schwellenspannung entsprechend einem programmierten Zustand einer Speicherzelle **4** aufweist. Dritte Haupt- und Arrayreferenzzellen können vorgesehen sein und so konfiguriert werden, dass sie eine Schwellenspannung aufweisen, die einem überlöschten Zustand einer Speicherzelle **4** entspricht. Ergänzende Arrayreferenzzellen können hinzugefügt werden, wie etwa eine Statuszelle zum Anzeigen, ob alle Speicherzellen **4** einer spezifizierten Gruppe von Speicherzellen **4** gelöscht sind.

[0058] Der Speicherzellenleseverstärker **12** ist vorgesehen zum Messen eines Stroms durch die Speicherzellen **4** des Speicherarrays **5**. Ein Arrayreferenzleseverstärker **13** ist vorgesehen zum Messen eines Stroms durch die Arrayreferenzzellen **6**. Ein Ar-

rayauswertungsmittel **14** ist an den Speicherzellenleseverstärker **12** und an den Arrayreferenzleseverstärker **13** angeschlossen. Durch Vergleichen der von den Leseverstärkern **12** und **13** empfangenen Signale wird ein Ergebnissignal erzeugt, das einen Logikzustand einer gewählten Speicherzelle **4** angibt.

**[0059]** Dieser Vorgang zum Lesen des Logikzustands einer Speicherzelle **4** wird immer dann ausgeführt, wenn die gespeicherten Daten beispielsweise von einem Hostsystem angefordert werden, das an die E/A-Schnittstelle **15** angeschlossen ist, oder wenn der Zustand einer Speicherzelle **4** nach einer Löscho- oder einer Programmieroperation bestätigt und deshalb gelesen werden soll.

**[0060]** Im Kontext dieser Anmeldung bedeutet "Schreiben" das Setzen des Zustands einer Speicherzelle **4** in einen gelöschten oder einen programmierten Zustand. "Programmieren" bedeutet das Setzen des Zustands auf einen ersten Logikwert und "Löschen" das Setzen des Zustands auf einen zweiten Logikwert.

**[0061]** Wie oben erwähnt ist es insbesondere bei stark integrierten Speicherbauelementen erforderlich, einen Referenzwert zu liefern, der sehr präzise ist, um eine scharfe und enge Programmier-/Löschdifferenzierung zu ermöglichen. Dies ist insbesondere dann wichtig, wenn mehr als zwei Zustände detektiert werden sollen, wie dies bei Verwendung von Zwei-Bit-NROM-Zellen der Fall ist.

**[0062]** Ein Vorteil der vorliegenden Erfindung besteht darin, dass die Referenzzellen **6** neben den Speicherzellen **4** angeordnet sind, die gelesen werden sollen. Sie werden auch als normale Speicherzellen behandelt und weisen deshalb zeitlich und während Schreibzyklen das gleiche Verhalten wie Speicherzellen **4** auf und sind den gleichen Temperaturschwankungen ausgesetzt. Außerdem beeinflussen alle anderen Arten von Störungen die Referenzzellen **6** auf die gleiche Weise wie die Speicherzellen **4**. Die jeweiligen Referenzzellen **6** sind mit der gleichen Wortleitung wie die zugewiesenen Speicherzellen **4** verbunden. Deshalb aktiviert auch die Wahl einer Speicherzelle **4** durch Aktivieren der Wortleitung die entsprechende Arrayreferenzzelle **6**. Dieser Mechanismus stellt sicher, dass der durch eine Speicherzelle **4** fließende und der durch eine Referenzzelle **6** fließende Strom zur gleichen Zeit gemessen werden.

**[0063]** Zur Bereitstellung exakter Referenzwerte und gemäß der Erfindung wird vorgeschlagen, eine Menge von Hauptreferenzzellen **7** bereitzustellen, die als Referenz für das Schreiben von Referenzwerten in die Arrayreferenzzellen **6** verwendet werden. Ein Hauptreferenzleseverstärker **16** ist ebenfalls mit dem Arrayauswertungsmittel **14** verbunden. Deshalb

können Zellen **4** und **6** des Speicherarrays **5** unter Verwendung der Hauptreferenzen **7** auch gelesen oder bestätigt werden. Im Allgemeinen werden zum Lesen gespeicherter Daten die Arrayreferenzzellen **6** verwendet, da diese Referenzen liefern, die sich auf die gleiche Weise verschlechtern wie die in den Speicherzellen **4** gespeicherten Daten. Dies schützt zusätzlich die Hauptreferenzzellen **7** vor weiterer Beanspruchung, was die Schwellwerte der Referenzzellen **7** verschlechtern könnte. Die Arrayreferenzzellen **6** und die Speicherzellen **4**, die dieser Zeile zugewiesen sind, können später über die von den Hauptreferenzzellen **7** bereitgestellten Referenzen aufgefrischt werden. Dadurch werden die betroffenen Zellen **4** und **6** auf optimale Werte gesetzt.

**[0064]** Für das Setzen der Arrayreferenzwerte **6** wird der gleiche Mechanismus implementiert wie bereits für das Lesen von Speicherzellen **4** erläutert. Wie oben beschrieben kann ein Strom durch eine Arrayreferenzzelle **6** durch den Arrayreferenzleseverstärker **13** gemessen werden. Ein Strom durch die Hauptreferenzzellen **7** wird von den Hauptreferenzleseverstärkern **16** gemessen. Beide Leseverstärker **13** und **16** sind an ein Hauptauswertungsmittel **17** angeschlossen, das die von den Leseverstärkern **13** und **16** gemessenen Werte vergleicht und ein dem Logikzustand einer gemessenen Arrayreferenzzelle **6** entsprechendes Signal ausgibt. Im Gegensatz zu dem Setzen von Speicherzellen **4** unter Verwendung der Arrayreferenzzellen **6** sind die Hauptreferenzzellen **7** und die Arrayreferenzzellen **6** nicht mit der gleichen Wortleitung verbunden. Die Hauptreferenzzellen **7** werden von einem zusätzlichen Steuersignal gesteuert.

**[0065]** Sowohl das Arrayauswertungsmittel **14** als auch das Hauptauswertungsmittel **17** sind mit der Steuereinheit **11** verbunden. Zum Setzen einer Arrayreferenzzelle **6** in einen vorgegebenen Zustand wird der tatsächliche Zustand von den Leseverstärkern **13** und **16** und dem Hauptauswertungsmittel **17** detektiert. Der Zustand wird an die Schreibschaltung **18** der Steuereinheit **11** übermittelt, und gegebenenfalls wird eine Schreiboperation ausgeführt. Nach dem Schreiben von Daten in die Arrayreferenzzelle **6** wird eine Leseoperation ausgeführt, um zu bestätigen, ob die Schreiboperation ausreichte, um den erforderlichen Zustand zu erreichen.

**[0066]** Bei der Ausführungsform nach [Fig. 1](#) werden zwei Leseverstärker **12** und **13** bereitgestellt, um den Zustand einer Speicherzelle **4** oder einer Arrayreferenzzelle **6** auszuwerten. Beide Leseverstärker **12** und **13** sind mit einem Auswertungsmittel **14** zum Erzeugen eines dem detektierten Logikzustand entsprechenden Signals verbunden. Es stehen andere Arten von Leseverstärkern zu Verfügung, die die gemessenen Ströme durch die Speicherzelle **4** bzw. die Arrayreferenzzelle **6** direkt vergleichen und ein ent-

sprechendes Ausgangssignal erzeugen. Die Erfindung kann unabhängig von den verwendeten Leseverstärkern angewendet werden, so dass alle Arten von Leseverstärkern verwendet werden können.

**[0067]** [Fig. 2](#) ist ein Flussdiagramm, das den Prozess der Herstellung eines nichtflüchtigen Speicherbauelements **1** gemäß der vorliegenden Erfindung beschreibt. Im Schritt **31** wird ein nichtflüchtiges Speicherbauelement **1** bereitgestellt, das mehrere, in einem Speicherarray **5** angeordnete Speicherzellen **4** umfasst, mindestens zwei Arrayreferenzzellen **6**, die programmiert werden können, so dass sie eine Schwellenspannung aufweisen, die einem gelöschten Zustand bzw. einem programmierten Zustand entsprechen, mindestens zwei Hauptreferenzzellen **7**, die programmiert werden können, um eine Schwellenspannung aufzuweisen, die einem gelöschten Zustand bzw. einem programmierten Zustand entspricht. Ein derartiges Speicherbauelement entspricht dem einen, das in [Fig. 1](#) als ein mögliches Beispiel beschrieben wird.

**[0068]** Im Schritt **32** wird eine gelöschte Hauptreferenzzelle **7** auf eine Schwellenspannung gesetzt, die einem gelöschten Zustand einer Speicherzelle **4** entspricht. Im Fall einer NROM-Zelle erfolgt dies durch die oben beschriebenen Mechanismen. Genauer gesagt erfolgt dies durch Löschen der jeweiligen Hauptreferenzzelle **7** und danach Bestätigen, ob die Schwellenspannung die von dem Hersteller definierten Anforderungen erfüllt.

**[0069]** Im Schritt **33** geschieht das Gleiche für eine andere Hauptreferenzzelle **7**, die später als Referenz für das Programmieren oder Lesen von Arrayreferenzzellen **6** oder Speicherzellen **4** verwendet wird.

**[0070]** In einem weiteren Schritt **34** wird das Speicherbauelement beispielsweise an einen Kunden verschickt. Das Schreiben von Referenzwerten in die Hauptreferenzzellen **7** kann am besten während der Wafersortierung erfolgen. Es ist aber auch möglich, die Hauptreferenzzellen **7** auf die vordefinierten Schwellenspannungen zu setzen, wenn das nichtflüchtige Halbleiterspeicherbauelement getestet wird, wie in [Fig. 3](#) dargestellt.

**[0071]** Gemäß dem Diagramm von [Fig. 3](#) wird ein Testprogramm nach dem Setzen der Hauptreferenzzellen **7** ausgeführt. Das Testprogramm enthält Löscho- und Programmieroperationen, die für die Hauptreferenzzellen **7** durchgeführt werden.

**[0072]** Nach dem Abschließen des Testprogramms werden die Hauptreferenzzellen nicht länger zyklisch beansprucht. Dies bedeutet, dass an den Hauptreferenzzellen **7** keine Schreiboperationen ausgeführt werden. Sie werden nur in Leseoperationen verwendet, dies beschleunigt aber den Verschlechterungs-

prozess der Hauptreferenzzellen **7** nicht wesentlich. Deshalb liefern die Hauptreferenzzellen **7** während der ganzen Lebenszeit eines nichtflüchtigen Speicherbauelements **1** eine präzise Referenz.

**[0073]** Wenn ein nichtflüchtiges Speicherbauelement **1** gemäß der Erfindung betrieben wird, müssen die Arrayreferenzzellen **6** auf entsprechende Schwellwerte gesetzt werden, damit sie als Referenzen für das Programmieren und Lesen von Speicherzellen **4** des Speicherarrays **5** dienen, wie in [Fig. 4](#) gezeigt. Im Schritt **31** wird ein nichtflüchtiges Speicherbauelement **1** mit Hauptreferenzzellen **7** versehen, die vor dem Verschicken auf die Referenzwerte gesetzt werden. In den Schritten **41** und **42** werden Arrayreferenzzellen **6** auf einen programmierten bzw. einen gelöschten Zustand für das spätere Lesen oder Programmieren von Speicherzellen **4** des Speicherarrays **5** gesetzt (Schritt **43**).

**[0074]** Die Arrayreferenzzellen **6** zur Bereitstellung einer Löschrferenz werden auf einen Gelöscht-Schwellwert gesetzt, indem zuerst die Zellen gelöscht werden und danach bestätigt wird, ob die Schwellwerte die für eine gelöschte Zelle definierten Anforderungen erfüllen. Für diesen Schritt wird die gelöschte Hauptreferenzzelle **7** als Referenz verwendet (Schritt **41**).

**[0075]** Im Schritt **42** erfolgt das Gleiche für andere Arrayreferenzzellen **6**, die unter Verwendung der programmierten Hauptreferenzzelle **7** als Referenz auf einen programmierten Zustand gesetzt werden.

**[0076]** Nach Durchführung dieser Schritte können im Schritt **43** die Speicherzellen **4** gelesen oder geschrieben werden. Für das Lesen von Speicherzellen **4** kann der tatsächliche Zellenstrom mit mehreren Referenzströmen verglichen werden, um herauszufinden, in welchem Zustand sich eine Speicherzelle **4** befindet. Es ist außerdem möglich, aus dem Strom der gelöschten Arrayreferenzzelle **6** und dem Strom der programmierten Arrayreferenzzelle **6** einen Vergleichsstrom zu erzeugen und den tatsächlichen Zellenstrom mit dem Vergleichsstrom zu vergleichen.

**[0077]** Das Verhalten von Speicherzellen **4** und **6** ändert sich im Lauf der Zeit und nach vielen Lesezyklen. Das erste Kriterium bedeutet, dass die Speicherzellen **4** und **6** nach einer langen Zeit Ladung verlieren. Bezeichnet wird dies als „Erhaltung nach dem periodischen Durchlaufen von Lesevorgängen (RAC – Retention after Cycling)“. Die Änderung des Verhaltens nach vielen Schreibzyklen wird als die Beanspruchungsdauer bezeichnet. Wie oben beschrieben, ist es ein Vorteil der Erfindung, dass die Arrayreferenzzellen **6**, die zum Lesen der Speicherzellen **4** verwendet werden, der gleichen Beanspruchung wie die Speicherzellen **4** selbst unterworfen sind. Um gute Lese- oder Schreibeergebnisse zu erzielen, ist es

jedoch erforderlich, die Referenzzellen **6** zyklisch zu belasten. Dies geschieht durch Schreiben von Referenzzellen **6** und Lesen von Referenzzellen **6**, wie dies auch mit normalen Speicherzellen **4** geschieht. Es ist ein Aspekt der Erfindung, ein Verfahren bereitzustellen, das sicherstellt, dass eine Auffrischung der Referenzzellen **6** auf einer regelmäßigen Basis erfolgt.

**[0078]** Gemäß dem in dem in [Fig. 5A](#) gezeigten Flussdiagramm kann das Auffrischen von gelöschten Arrayreferenzzellen **6** an den Aufruf von vordefinierten Benutzerbefehlen gekoppelt werden. In einem ersten Schritt **51** wird ein Benutzerbefehl erhalten. In einem Speicherbauelement **1** geschieht dies durch eine Steuereinheit, wie in [Fig. 1](#) gezeigt. Bei einem weiteren Schritt **52** wird geprüft, ob der erhaltene Befehl einer von vordefinierten speziellen Befehlen ist. Wenn dies der Fall ist, werden die gelöschten Arrayreferenzzellen **6** aufgefrischt, und danach wird der Benutzerbefehl in einem Schritt **54** ausgeführt. Wenn in Schritt **52** detektiert wurde, dass der erhaltene Befehl kein spezieller Benutzerbefehl ist, wird der Befehl in Schritt **54** direkt ausgeführt. Nach den jeweiligen Operationen fährt das Verfahren mit Schritt **51** fort, der darin besteht, auf einen neuen Benutzerbefehl zu warten.

**[0079]** [Fig. 5B](#) ist ein Flussdiagramm, das sich auf eine modifizierte Vorgehensweise zum Auffrischen von programmierten Arrayreferenzzellen bezieht. In Schritt **51** wird ein Benutzerbefehl erhalten. Im Schritt **55** wird getestet, ob die tatsächliche Schwellenspannung der programmierten Referenzzelle **6** bezüglich der Hauptprogrammreferenzspannung niedrig ist. Wenn dies der Fall ist, werden die programmierten Referenzzellen **6** des Arrays zuerst aufgefrischt. Falls nicht, wird der Benutzerbefehl direkt ausgeführt.

**[0080]** [Fig. 6](#) ist ein Schemadiagramm eines beispielhaften Schaltplans, der eine Hauptarrayreferenzzelle **6** mit zugewiesenem Leseverstärker **13** und auch eine Speicherzelle **4** des Speicherarrays **5** mit einem zugewiesenen Leseverstärker **12** zeigt. Die Arrayreferenzzelle **6** und die Speicherzelle **4** sind der gleichen Wortleitung **61** zugewiesen. Deshalb wird durch die Wahl der Speicherzelle **4** auch die Arrayreferenzzelle **6** gewählt. Unter Verwendung der Wortleitung **61** werden die Gateelektroden der Transistoren **6** und **4** vorgespannt. Für die Wahl der Speicherzellen **6** und **4**, die von den gezeigten Transistoren gebildet werden, müssen auch die jeweiligen Bitleitungen **75** und **77** gewählt werden. Dies geschieht durch Decodierer **63**, die die Spaltendecodierer sind. Wenn alle Schalttransistoren des Decodierers **63** geschlossen sind, entspricht die Spannung an einem Knoten **67** der Drainelektrode des Transistors der Arrayreferenzzelle **6**, und die Spannung am Knoten **69** der an der Drainelektrode des Transistors **4** angelegten Spannung.

**[0081]** Nach dem Schließen der Schalttransistoren des Decodierers **63** beginnt ein Strom  $I_{\text{CELL-REF}}$  vom Knoten **67** über die Drainelektrode und Sourceelektrode des Transistors **6** zum Masseknoten GND zu fließen. Der Strom  $I_{\text{CELL-REF}}$  hängt von der Leitfähigkeit des Kanals von der Drainelektrode zur Sourceelektrode des Transistors **6** ab. Die Leitfähigkeit hängt von der Ladung ab, die sich auf der Nitridschicht der Speicherzelle **6** befindet, die bei dieser Ausführungsform der Erfindung eine NROM-Zelle ist. Der Strom hängt mit anderen Worten vom Zustand der Speicherzelle **6** ab, wenn die Spannung am Knoten **67** konstant ist. Dazu ist ein Steuerverstärker **62** vorgesehen, der an einem Eingang die Spannung am Knoten **67** erhält. Der Steuerverstärker **62** steuert einen weiteren Transistor **73**, der zwischen den Knoten **67** und einen Knoten **68** geschaltet ist. Der Knoten **68** ist auch mit einem Eingang eines Vergleichers **65** und einer Stromquelle **71** verbunden, die auf einen Referenzstrom  $I_{\text{REF}}$  eingestellt wird. Über den Steuerverstärker **62** wird sichergestellt, dass die Spannung am Knoten **67** unabhängig von dem Zellenstrom  $I_{\text{CELL-REF}}$  konstant gehalten wird. Andere Ausführungsformen eines Leseverstärkers sind möglich und unterliegen dem durchschnittlichen Geschick einer Person, die mit jeder Art von Speicherbauelementen vertraut ist.

**[0082]** Die Spannung am Knoten **68** hängt von dem Strom  $I_{\text{CELL-REF}}$  und dem Referenzstrom  $I_{\text{REF}}$  ab. Wenn der Strom  $I_{\text{CELL-REF}}$  höher ist als der Strom  $I_{\text{REF}}$  gibt der Vergleichler **65** ein H-Signal aus, ansonsten gibt er ein L-Signal aus. Jeder der Vergleichler **65** und **66** weist einen zweiten Eingang auf, an dem eine Referenzspannung  $V_{\text{REF}}$  angelegt wird. Die Spannung  $V_{\text{REF}}$  wird als Vergleichsspannung für die Spannungen an Knoten **68** bzw. **70** verwendet. Sie ist gemäß der Dynamik des Knotens **68** und **70** entsprechend bemessen.

**[0083]** Wenn die Speicherzelle **4**, die als ein Beispiel für mehrere Speicherzellen **4** gezeigt ist, durch Anlegen einer entsprechenden Gatespannung an die Speicherzelle **4** gelesen wird, steigt der Zellenstrom  $I_{\text{CELL}}$  auf Grund einer erhöhten Leitfähigkeit der Speicherzelle **4** an. Dies geschieht unter der Vorbedingung, dass die Drain-Source-Spannung, die die Spannung am Knoten **69** ist, konstant gehalten wird. Dies geschieht durch Erfassen der Spannung am Knoten **69** und Steuern der Gatespannung des Transistors **74**, um seine Leitfähigkeit zu erhöhen. Da die Stromquelle **72** unverändert bleibt, nimmt die Spannung am Knoten **70** ab. Wenn die an die Speicherzelle **4** angelegte Gatespannung hoch genug ist, wird der Zellenstrom  $I_{\text{CELL}}$  größer als ein Schwellwert, die Spannung am Knoten **70** fällt unter die Spannung  $V_{\text{REF}}$  ab und der Vergleichler **66** kippt.

**[0084]** Die gleiche Vorgehensweise läuft für die Referenzzelle **6** ab, die in [Fig. 6](#) mit der gleichen Wortleitung **61** verbunden ist. Unter der Annahme, dass

die Gatespannungen der Speicherzelle **4** und der Referenzzelle **6** variiert werden, d.h. linear erhöht werden, kippen sowohl der Referenzzellenvergleich **65** als auch der Speicherzellenvergleich **66**, aber zu unterschiedlichen Zeiten. Die Information, welcher Vergleich zuerst kippt, zeigt den Zustand der Speicherzelle **4** an.

**[0085]** Die Schaltung der Speicherzelle **4** und des zugewiesenen Leseverstärkers **12** arbeitet auf die gleiche Weise. Aus diesem Teil der Schaltungsanordnung werden Informationen über den Zustand der Speicherzelle **4** erhalten.

**[0086]** Die Speicherzellen sind wie oben erwähnt in einer bevorzugten Ausführungsform der Erfindung NROM-Zellen (Nitridfestwertspeicher). Ein prinzipieller Aufbau einer NROM-Zelle ist in [Fig. 7](#) gezeigt. Ein prinzipielles Verständnis eines solchen Bauelements ist nützlich für das Verständnis einiger Aspekte der Erfindung, die unten beschrieben sind. In [Fig. 7](#) ist ein Querschnitt durch eine herkömmliche Oxid-Nitrid-Oxid-EEPROM-Speicherzelle dargestellt. Die allgemein als **4** bezeichnete Speicherzelle umfasst ein Silizium **81** vom p-Typ, zwei pn-Übergänge zwischen n+-Source- und -Drainebereichen **82** und **83** und ein Substrat **81** vom p-Typ, eine zwischen zwei Oxidschichten **87** und **88** geschichtete nichtleitende Nitridschicht **86** und eine polykristalline leitende Schicht **89**.

**[0087]** Zum Programmieren der Zelle werden Spannungen an die Drainelektrode **83** und an die Gateelektrode **85** angelegt, während die Sourceelektrode **82** mit Masse verbunden ist. Beispielsweise werden 10 V an die Gateelektrode und 9 V an die Drainelektrode angelegt. Diese Spannungen erzeugen ein vertikales und seitliches elektrisches Feld **84** entlang einer Länge des Kanals von der Sourceelektrode **82** zur Drainelektrode **83**. Dieses elektrische Feld bewirkt, dass Elektronen von der Sourceelektrode **82** abgezogen werden und sich in Richtung auf die Drainelektrode **83** zu beschleunigen beginnen. Während sie sich entlang der Länge des Kanals bewegen, gewinnen sie Energie. Wenn die ausreichend Energie gewinnen, sind sie in der Lage, über die Potentialbarriere der Oxidschicht **88** in die Siliziumnitridschicht **86** zu springen und gefangen zu werden. Die Wahrscheinlichkeit, dass dies eintritt, ist am höchsten in dem Gebiet der Gateelektrode neben der Drainelektrode **83**, definiert durch einen gestrichelten Kreis **90a**, weil nämlich die Elektronen die meiste Energie in der Nähe der Drainelektrode **83** gewinnen. Diese beschleunigten Elektronen werden als heiße Elektronen bezeichnet, und nach ihrer Injektion in die Nitridschicht werden sie gefangen und bleiben dort gespeichert. Die gefangenen Elektronen können sich wegen der geringen Leitfähigkeit der Nitridschicht **86** und des schwachen seitlichen elektrischen Felds in dem Nitrid nicht durch die Nitridschicht ausbreiten.

Somit bleibt die gefangene Ladung in einem lokalisierten Fanggebiet **90a** in dem Nitrid, in der Regel in der Nähe der Drainelektrode **83** gelegen.

**[0088]** Analog wird das linke Bit programmiert, indem Programmierspannungen an die Gateelektrode **85** und die Sourceelektrode **82** angelegt werden, die nun als die Drainelektrode für das linke Bit fungiert. Heiße Elektronen werden in die Ladungen fangende Schicht **86** in dem durch den gestrichelten Kreis **90b** definierten Gebiet injiziert und dort gefangen. Die Schwellenspannung des Abschnitts des Kanals unter der gefangenen Ladung umfassend das linke Bit nimmt zu, wenn mehr und mehr Elektronen in die Nitridschicht injiziert werden.

**[0089]** Eine ausführlichere Beschreibung dessen, wie eine NROM-Zelle arbeitet, findet man in US 6,011,725. Das Prinzip dessen, wie mehr als ein Bit in eine Ein-Transistor-Zelle programmiert werden kann, ist in US 2002/0118566 beschrieben.

**[0090]** Zusammenfassend sind in den [Fig. 8A](#) bis [Fig. 8D](#) die verschiedenen möglichen Zustände dargestellt. Es ist möglich, zwischen einem linken Bit und einem rechten Bit zu unterscheiden, wie auch in [Fig. 7](#) veranschaulicht, wo der Aufbau der Speicherzelle gezeigt ist. In [Fig. 8A](#) befinden sich beide Bits der Speicherzelle in einem L-Zustand. In [Fig. 8B](#) befindet sich das rechte Bit in einem H-Zustand und das linke Bit in einem L-Zustand. In [Fig. 8C](#) ist eine Situation gezeigt, bei der sich das linke Bit in einem H-Zustand und die rechte Seite sich in einem L-Zustand befindet. Gemäß [Fig. 8D](#) befinden sich beide Bits in einem H-Zustand.

**[0091]** Beim Messen des Stroms durch die Zelle bei einer bestimmten Gatespannung  $V_G$  und in Abhängigkeit von der tatsächlichen Schwellenspannung können verschiedene Ströme  $I_D$  gemessen werden, wie in [Fig. 9](#) gezeigt. Beispielsweise können bei einer Gatespannung von 2 V vier verschiedene Ströme gemessen werden. Es kann deshalb detektiert werden, zu welchem Zustand der gemessene Strom gehört.

**[0092]** Das Programmieren der Nachbarbits beeinflusst auch die Schwellenspannung des anderen Bit, die beispielsweise auf einem niedrigen Pegel gehalten werden sollte. Dieser Effekt erscheint, da das zweite Bit nicht nur die Leseoperationen beeinflusst, sondern auch die Bestätigungsoperationen, die während des Schreibens auftreten, zum Beispiel Löschen oder Programmieren des Arrays. Eine Bestätigungsoperation ist eine Leseoperation, die während Schreiboperationen ausgeführt wird, und eine Notwendigkeit für einen Programmier- oder Löschimpuls bewertet, um die beabsichtigten Daten ordnungsgemäß zu schreiben. Das zweite Bit beeinflusst das Bestätigungsergebnis und deshalb das Programmierverhalten der Zelle. Angesichts dieses Effekts wer-

den die in den [Fig. 8A](#) bis [Fig. 8D](#) gezeigten Zustände erweitert, wie in [Fig. 10A](#) bis [Fig. 10D](#) gezeigt.

**[0093]** Bei einem zukünftigen Szenarium, bei dem mehr als zwei Bits in einer Speicherzelle gespeichert werden, wird die Situation schwieriger, da weitere Zustände gemäß unterschiedlicheren Schwellenspannungen detektiert werden müssen.

**[0094]** Um beim Verständnis zusätzlicher Aspekte der Erfindung behilflich zu sein, wird auch erläutert, was ein überlöscher Zustand einer Speicherzelle ist. Unter Bezugnahme auf [Fig. 11](#), die Schwellenspannungen von mehreren Referenz- oder Speicherzellen **6** oder **4** zeigt, kann man sehen, dass eine Schwellenspannung einer Speicherzelle einen programmierten Pegel  $V_{TH(PROGRAM)}$  und einen gelöschten Pegel  $V_{TH(ERASE)}$  umfasst. Für einen programmierten Zustand einer Speicherzelle wird die Schwellenspannung  $V_{TH}$  auf einen über dem programmierten Pegel liegenden Pegel gebracht. Für einen gelöschten Zustand wird die Schwellenspannung auf einen geringfügig über dem gelöschten Pegel liegenden Pegel reduziert. Beim Übergang vom programmierten Zustand in den gelöschten Zustand könnte es passieren, dass die Schwellenspannung unter den gelöschten Pegel  $V_{TH(ERASE)}$  abfällt. Dieser Zustand wird als ein „überlöscher Zustand“ bezeichnet. Ein Grund für das Auftreten des überlöscher Zustands besteht darin, dass mehrere Speicherzellen gemeinsam gelöscht werden, aber das Verhalten dieser Zellen verschieden ist. Während eine Zelle den gelöschten Zustand erreicht, ist eine andere Zelle bereits übergelöscht. Ein weiterer beispielhafter Grund ist eine zu einer Nachbarzelle geleitete Löschoption, was zu einem Ladungsverlust führt. Eine Softprogrammoperation, auch als Auffrischoperation bezeichnet, für jede Zelle ausgeführt, die sich in dem übergelöschten Zustand befindet, stellt sicher, dass die Schwellenspannung aller Zellen auf den gelöschten Zustand eingestellt wird. Dies ändert nicht den Logikzustand der Zelle **4**, doch wird die tatsächliche Schwellenspannung auf einen ordnungsgemäßen Wert gemäß der Schwellenspannung der gelöschten Referenzzelle **6** eingestellt.

**[0095]** Mit den [Fig. 12](#) bis [Fig. 15](#) wird beschrieben, wie sichergestellt werden kann, dass die Arrayreferenzzellen **6** der gleichen Beanspruchung unterworfen sind wie die Speicherzellen **4**. Dies kann erreicht werden, wenn nach jeder Löschoption oder Programmdatenoperation auch die Referenzzellen **6** zyklisch beansprucht werden, was bedeutet, dass sie gelöscht oder programmiert werden. In Schritt **51** von [Fig. 12](#) wird ein Benutzerbefehl erhalten. Wenn es ein Lesebefehl ist, werden in Schritt **91** Lesedatenoperationen durchgeführt. Wenn es ein Löschoption ist, werden die Operationen gemäß Schritt **92** ausgeführt. Nach dem Ausführen der Operationen von Schritt **92** werden mindestens einige der Arrayreferenzzellen im

Schritt **94** zyklisch beansprucht. Das Gleiche geschieht, wenn ein Programmbefehl erhalten wird. In diesem Fall werden Operationen gemäß Schritt **93** ausgeführt und danach werden zumindest einige der Referenzzellen im Schritt **94** zyklisch beansprucht. Alternativ kann das zyklische Beanspruchen von Referenzzellen gemäß Schritt **94** gleichzeitig mit dem Löschen oder Programmieren von Speicherzellen **4** durchgeführt werden, wie durch den gestrichelten Pfeil angezeigt. Die Arrayreferenzzellen **6** werden nicht zum Löschen oder Programmieren verwendet. Bestätigungsoperationen, die erforderlich sind, wenn Speicherzellen **4** gelöscht oder programmiert werden, werden unter Verwendung der Hauptreferenzzellen **7** als Referenzen durchgeführt.

**[0096]** In [Fig. 13](#) wird dargestellt, wie ein Blocklöschbefehl arbeitet. Zunächst werden alle gelöschten Arrayreferenzzellen **6** im Schritt **101** programmiert, eine Schwellenspannung gemäß der gelöschten Hauptreferenzzelle **7** aufzuweisen. In einem folgenden Schritt **102** werden die Speicherzellen **4** unter Verwendung der gelöschten Arrayreferenzzellen **6** als Referenz gelöscht.

**[0097]** In einem Schritt **103** werden die gelöschten Arrayreferenzzellen **6** für spätere Leseoperationen vorbereitet. Dann werden auch die programmierten Arrayreferenzzellen **6** im Schritt **104** für spätere Leseoperationen vorbereitet.

**[0098]** Bei einer alternativen Implementierung eines Blocklöschbefehls werden die Schritte **103** und **104** nicht anschließend nach dem Löschen der Speicherzellen **4** im Schritt **102** ausgeführt. Die alternative Ausführungsform wird realisiert, wenn eine Statuszelle oder ein Statusbit dazu verwendet wird anzuzeigen, dass alle Speicherzellen einer vordefinierten Gruppe gelöscht sind. Eine derartige Gruppe kann von allen Speicherzellen **4** und Referenzzellen **6** gebildet werden, die mit der gleichen Wortleitung **61** verbunden sind. Wenn ein Blocklöschbefehl gestartet wird, werden alle diese Zellen gelöscht, auch die Referenzzellen **6** und die Statuszelle. Vor dem Lesen von Daten müssen die Referenzzellen **6** auf entsprechende Werte gesetzt sein, d.h. auf einem gelöschten und einem programmierten Zustand, wenn die jeweiligen Zellen vorgesehen sind. Die „gelöschte Arrayreferenzzelle **6**“ muss gemäß Schritt **103** auf einen „schlechten“ Referenzwert gesetzt sein, und die „programmierte Arrayreferenzzelle“ muss gemäß Schritt **104** programmiert werden. Dies kann auch vor dem Programmieren des ersten Bit nach einem Blocklöschbefehl geschehen, wie unten unter Bezugnahme auf [Fig. 15](#) und [Fig. 16](#) beschrieben.

**[0099]** [Fig. 14](#) hilft zu erläutern, wie dies im Einzelnen funktioniert. [Fig. 14](#) ist ein Zeitdiagramm, das die Schwellenspannungen der gelöschten Arrayreferenzzelle **6**, der programmierten Arrayreferenzzelle **6**

und einer Speicherzelle **4** zeigt.

**[0100]** Zum Zeitpunkt T1 startet Schritt **101** und gelöschte Referenzzellen **6** werden bis zu der Hauptprogrammreferenzschwellsenspannung MPV vorprogrammiert. Zum Zeitpunkt T2 werden die gelöschten Arrayreferenzzellen **6** gelöscht, bis sie die Hauptlöschreferenzschwellsenspannung MEV erreichen.

**[0101]** Zum Zeitpunkt T3 startet Schritt **102**. Zuerst werden die Speicherzellen **4** auf eine höhere Schwellsenspannung vorprogrammiert, und dann, im Moment T4, werden die Speicherzellen **4** auf die durch die gelöschten Hauptreferenzzellen **7** definierte gelöschte Hauptreferenzschwellsenspannung MEV gelöscht.

**[0102]** Zum Zeitpunkt T5 startet Schritt **103** mit dem Vorbereiten der gelöschten Arrayreferenzzellen **6** für das Lesen von Daten. Dies bedeutet, dass eine gelöschte Worst-Case-Referenz einschließlich dem zweiten-Bit-Effekt erzeugt wird. Durch diese Operation steigt die Schwellsenspannung der gelöschten Arrayreferenzzellen **6** an.

**[0103]** Zum Zeitpunkt T6 startet Schritt **104** mit dem Vorbereiten der programmierten Arrayreferenzzellen **6** für das Lesen von Daten. Dies geschieht durch Erhöhen der Schwellsenspannung auf eine Hauptprogrammreferenzschwellsenspannung MPV. Nach einiger Zeit nimmt die Schwellsenspannung ab, und deshalb wird eine Auffrischung benötigt. Die Auffrischung ist erforderlich, weil beispielsweise für andere Zellen durchgeführte Blocklöschoperationen oder Programmoperationen sich auf benachbarte Zellen auswirken, was zu einem Ladungsverlust führt. Außerdem werden die Referenzzellen **6** beeinflusst. Das Auffrischen der programmierten Arrayreferenzzellen **6** stellt sicher, dass der Ladungsverlust kompensiert wird.

**[0104]** Wie oben angedeutet, ist das Verfahren, wie es unter Bezugnahme auf die [Fig. 13](#) und [Fig. 14](#) beschrieben wird, ein Beispiel. Aus Zuverlässigkeitsgründen ist es besser, die Arrayreferenzzellen **6** vorzubereiten, wenn Daten in die Speicherzellen **4** programmiert werden.

**[0105]** [Fig. 15](#) zeigt ein weiteres Flussdiagramm für einen Wortprogrammbehl. Vor dem Schreiben von Daten wird das linke Bit der programmierten Arrayreferenzzelle bei Schritt **110** eingelesen. Wie dies gemäß der Konvention bei der vorliegenden Ausführungsform der Erfindung geschieht, wird üblicherweise das rechte Bit der Referenzzellen zum Speichern des Referenzwerts verwendet. Der Zustand des linken Bits gibt die Information an, ob zuvor Bits auf dieser Wortleitung programmiert wurden. Wenn das linke Bit eine logische „1“ ist, wie in Schritt **111** geprüft, bedeutet dies, dass auf dieser Wortleitung **61** zuvor

kein Bit programmiert wurde. Deshalb sollte in Schritt **112** eine Auffrischoperation des rechten Bit unter Verwendung der programmierten Hauptreferenz **7** ausgeführt werden. Eine Auffrischung, auch als Soft-Programmierung bezeichnet, bedeutet, dass die Schwellsenspannung bestätigt wird und gegebenenfalls auf einen neuen Pegel gesetzt wird, während der Logikzustand der Zelle der gleiche bleibt. Dies unterscheidet sich von einer Programmieroperation, die den Logikzustand von gelöscht zu programmiert ändert, oder der Löschoption, die den Logikzustand einer Zelle von programmiert zu gelöscht ändert.

**[0106]** Nach dem Auffrischen des rechten Bit erfolgt in Schritt **113** eine Programmierung des linken Bit. Da das linke Bit nun programmiert ist, ist zu erkennen, dass ein Bit auf dieser Wortleitung **61** geschrieben wurde, wenn das nächste Mal ein Wortprogrammbehl ausgeführt wird.

**[0107]** Nach dem Programmieren des linken Bit in Schritt **114** können Daten in die Speicherzellen **4** gemäß der Benutzerdaten geschrieben werden. Bei einem folgenden Schritt **115** wird eine Auffrischung der zuvor programmierten Bits, die möglicherweise Ladung verloren haben könnten, ausgeführt.

**[0108]** [Fig. 16](#) zeigt ein weiteres Beispiel des Detektierens, ob ein Bit zuvor auf einer Wortleitung **61** programmiert wurde. Bei dem dargestellten Beispiel wird die oben erwähnte Zeilenschutzreferenz zum Speichern der erforderlichen Information verwendet. Wenn alle mit der Wortleitung **61** verbundenen Speicherzellen **4** gelöscht werden, wird auch die Zeilenschutzreferenzzelle gelöscht. Wenn eine logische „1“ von der Zeilenschutzzelle gelesen wird, was in Schritt **120** unter Verwendung der Hauptreferenzzellen **7** als Referenz erfolgt, weiß das Arraybewertungsmittel **14** (Schritt **121**), dass vor der Programmierung von Daten die programmierte Arrayreferenzzelle **6** programmiert und bestätigt werden muss, wobei die programmierte Hauptreferenzzelle **7** verwendet wird (Schritt **122**). Dadurch wird eine programmierte Arrayreferenz **6** bereitgestellt. Um eine „schlechte“ gelöschte Arrayreferenz **6** bereitzustellen, d.h., auch verschlechterte Speicherzellen **4** sollten ordnungsgemäß als zu löschend identifiziert werden, wird in Schritt **123** das Nachbarbit, das das linke Bit der gelöschten Arrayreferenzzelle **6** ist, programmiert.

**[0109]** Danach können Daten in Schritt **124** programmiert und bestätigt werden, wobei die gelöschten und programmierten Arrayreferenzzellen **6** verwendet werden. In Schritt **125** wird die Zeilenschutzreferenzzelle programmiert, so dass sie anzeigt, dass mindestens ein Bit auf dieser Wortleitung **61** programmiert worden ist. Wahlweise wird in Schritt **126** ein Trimmen der gelöschten Arrayreferenzzelle **6** zum Ändern der rechten Bitladung der gelöschten Arrayreferenzzelle **6** ausgeführt.

**[0110]** Weitere Ausführungsformen der Erfindung unterliegen dem Wissen eines Durchschnittsfachmanns auf dem Gebiet. Sie werden deshalb von der Erfindung wie beansprucht abgedeckt.

### Patentansprüche

1. Nichtflüchtiges Halbleiterspeicherbauelement (1), das Folgendes umfasst:  
einen Speicherbereich (2) und einen Schaltungsbereich (3),  
wobei der Speicherbereich Folgendes umfasst:  
– mehrere in einem Speicherarray (5) angeordnete Speicherzellen (4) und  
– eine Menge von programmierbaren Arrayreferenzzellen (6), die als Referenzen zum Lesen der Speicherzellen (4) vorgesehen sind,  
wobei der Schaltungsbereich (3) Folgendes umfasst:  
– eine Menge von Hauptreferenzzellen (7), die als Referenzen zum Feststellen des Zustands von Arrayreferenzzellen (6) oder Speicherzellen (4) vorgesehen sind.

2. Nichtflüchtiges Halbleiterspeicherbauelement (1) nach Anspruch 1,  
wobei die Menge von Hauptreferenzzellen (7) Folgendes umfasst  
– eine erste Hauptreferenzzelle (7), die vorgesehen ist, um eine Schwellenspannung aufzuweisen, die einem programmierten Zustand einer Speicherzelle entspricht,  
– eine zweite Hauptreferenzzelle (7), die vorgesehen ist, um eine Schwellenspannung aufzuweisen, die einem gelöschten Zustand einer Speicherzelle entspricht, und  
wobei die Menge von programmierbaren Arrayreferenzzellen (6) Folgendes umfasst:  
– eine erste Arrayreferenzzelle (6), die vorgesehen ist, um eine aufzuweisen, die einem programmierten Zustand einer Speicherzelle entspricht,  
– eine zweite Arrayreferenzzelle (6), die vorgesehen ist, um eine Schwellenspannung aufzuweisen, die einem gelöschten Zustand einer Speicherzelle entspricht.

3. Nichtflüchtiges Halbleiterspeicherbauelement (1) nach Anspruch 1,  
wobei die Menge von Hauptreferenzzellen (7) Folgendes umfasst  
– eine erste Hauptreferenzzelle (7), die vorgesehen ist, um eine Schwellenspannung aufzuweisen, die einem programmierten Zustand einer Speicherzelle entspricht,  
– eine zweite Hauptreferenzzelle (7), die vorgesehen ist, um eine Schwellenspannung aufzuweisen, die einem gelöschten Zustand einer Speicherzelle entspricht, und  
wobei die Menge von programmierbaren Arrayreferenzzellen (6) Folgendes umfasst  
– eine Lesereferenzzelle (6), die vorgesehen ist, um

eine Schwellenspannung aufzuweisen, die die Grenze zwischen einem gelöschten Zustand und einem programmierten Zustand einer Speicherzelle (4) definiert und die zum Lesen von Daten aus den Speicherzellen (4) verwendet werden kann.

4. Nichtflüchtiges Halbleiterspeicherbauelement (1) nach Anspruch 3,  
wobei die Menge von Hauptreferenzzellen (7) zusätzlich Folgendes umfasst:  
– eine Lesereferenzzelle (7), die vorgesehen ist, um eine Schwellenspannung aufzuweisen, die die Grenze zwischen einem gelöschten Zustand und einem programmierten Zustand einer Speicherzelle (4) definiert und die als eine Referenz zum Feststellen des Zustands der Lesereferenzzelle (6) vorgesehen ist, die in der Menge von Arrayreferenzzellen (6) enthalten ist.

5. Nichtflüchtiges Halbleiterspeicherbauelement (1) nach einem der Ansprüche 2, 3 oder 4, wobei die Menge von Hauptreferenzzellen (7) zusätzlich eine Referenzzelle (7) umfasst, die vorgesehen ist, um eine Schwellenspannung aufzuweisen, die einem übergelöschten Zustand einer Speicherzelle (4) entspricht.

6. Nichtflüchtiges Halbleiterspeicherbauelement (1) nach einem der Ansprüche 2, 3 oder 4, wobei aus den Speicherzellen (4) mehrere Gruppen ausgebildet werden und für jede Gruppe eine Menge von Arrayreferenzzellen (6) vorgesehen ist.

7. Nichtflüchtiges Halbleiterspeicherbauelement (1) nach Anspruch 6, wobei in jeder der Menge von programmierbaren Arrayreferenzzellen (6) eine zusätzliche Statuszelle vorgesehen ist, um über einen vordefinierten Logikzustand anzuzeigen, ob alle Zellen (4) einer vordefinierten Gruppe von Speicherzellen (4) einen gelöschten Zustand aufweisen.

8. Nichtflüchtiges Halbleiterspeicherbauelement (1) nach Anspruch 1, wobei die Speicherzellen (4) Mehrbit-Speicherzellen sind.

9. Nichtflüchtiges Halbleiterspeicherbauelement (1) nach Anspruch 8, wobei jede Speicherzelle (4) aus einem Transistor mit einer elektrischen Ladungen fangenden Schicht (86) konfiguriert ist, die in der Lage ist, zwei binäre Informationsbits (90a, 90b) zu speichern.

10. Nichtflüchtiges Halbleiterspeicherbauelement (1) nach Anspruch 9, wobei die elektrischen Ladungen fangende Schicht (86) der Transistoren (4) aus einem Nitridmaterial besteht.

11. Nichtflüchtiges Halbleiterspeicherbauelement (1) nach Anspruch 8, wobei die Arrayreferenzzellen (6) aus einem Transistor mit einer elektrischen Ladun-

gen fangenden Schicht (86) konfiguriert ist, die in der Lage ist, zwei binäre Informationsbits (90a, 90b) zu speichern, wobei eines dieser Bits (90a) den Logikzustand der Referenzzelle (6) definiert.

12. Nichtflüchtiges Halbleiterspeicherbauelement (1) nach Anspruch 11, wobei das zweite Bit (90b) des Transistors einer gelöschten Arrayreferenzzelle (6) in einen entgegengesetzten Logikzustand gesetzt wird.

13. Nichtflüchtiges Halbleiterspeicherbauelement (1) nach Anspruch 1, wobei das Speicherbauelement (1) so eingerichtet ist,

– dass es in einem Testmodus betrieben wird, bei dem ein Setzen der Hauptreferenzzellen (7) in einem programmierten oder gelöschten Zustand freigegeben ist,

– und ein Setzen der Hauptreferenzzellen (7) in einen programmierten oder gelöschten Zustand nach dem Verlassen des Testmodus blockiert wird.

14. Verfahren zum Betreiben eines nichtflüchtigen Halbleiterspeicherbauelements (1), wobei das Speicherbauelement (1) Folgendes aufweist:

einen Speicherbereich (2) und einen Schaltungsbereich (3),

wobei der Speicherbereich (2) Folgendes umfasst:

– mehrere in einem Speicherarray (5) angeordnete Speicherzellen (4) und

– eine Menge von programmierbaren Arrayreferenzzellen (6), die als Referenzen zum Lesen der Speicherzellen (4) vorgesehen sind,

wobei der Schaltungsbereich (3) Folgendes umfasst:

– eine Menge von vorprogrammierten Hauptreferenzzellen (7), die als Referenzen zum Feststellen des Zustands von Arrayreferenzzellen (6) oder Speicherzellen (4) vorgesehen sind,

wobei das Verfahren die folgenden Schritte umfasst:

a) Setzen mindestens einer der Arrayreferenzzellen (6) in einen vordefinierten Zustand unter Verwendung mindestens einer der vorprogrammierten Hauptreferenzzellen (7) als Referenz,

b) Lesen des Zustands mindestens einer Speicherzelle (4) unter Verwendung der mindestens einen der Arrayreferenzzellen (6) als Referenz,

c) Feststellen des Zustands mindestens einer Speicherzelle (4, 6) beim Löschen oder Programmieren der Speicherzelle (4, 6) unter Verwendung mindestens einer der Hauptreferenzzellen (7) als Referenz.

15. Verfahren nach Anspruch 14, wobei

– eine erste Referenzzelle (7) der Menge von Hauptreferenzzellen (7) vorprogrammiert ist, um eine Schwellenspannung aufzuweisen, die einem programmierten Zustand einer Speicherzelle (4, 6) entspricht,

– eine zweite Referenzzelle (7) der Menge von Hauptreferenzzellen (7) vorprogrammiert ist, eine

Schwellenspannung aufzuweisen, die einem gelöschten Zustand einer Speicherzelle (4, 6) entspricht, und

das Verfahren die folgenden Schritte umfasst:

a) Setzen einer ersten der Arrayreferenzzellen (6) in einen gelöschten Zustand unter Verwendung der gelöschten Hauptreferenzzelle (7) als Referenz,

b) Setzen einer zweiten der Arrayreferenzzellen (6) in einen programmierten Zustand unter Verwendung der programmierten Hauptreferenzzelle (7) als Referenz,

c) Lesen mindestens einer der Speicherzellen (4) unter Verwendung der gelöschten und programmierten Referenzzellen (6, 7) zum Erzeugen eines Vergleichswerts für die Leseoperation.

16. Verfahren nach Anspruch 15, wobei die Schritte a) und b) von Anspruch 15 ausgeführt werden, wenn Daten in die Speicherzellen (4) geschrieben werden.

17. Verfahren nach Anspruch 14, wobei

– eine erste Referenzzelle (7) der Menge von Hauptreferenzzellen (7) vorprogrammiert ist, um eine Schwellenspannung aufzuweisen, die einem programmierten Zustand einer Speicherzelle (4, 6) entspricht,

– eine zweite Referenzzelle (7) der Menge von Hauptreferenzzellen (7) vorprogrammiert ist, eine Schwellenspannung aufzuweisen, die einem gelöschten Zustand einer Speicherzelle (4, 6) entspricht, und

– die Menge von Arrayreferenzzellen (6) eine Lesereferenzzelle (6) umfasst,

wobei das Verfahren die folgenden Schritte umfasst:

– Erzeugen eines Vergleichswerts unter Verwendung der gelöschten und programmierten Hauptreferenzzelle (7),

– Setzen der Lesereferenzzelle (6) auf eine Schwellenspannung, die die Grenze zwischen einem gelöschten Zustand und einem programmierten Zustand einer Speicherzelle (4, 6) unter Verwendung des Vergleichswerts definiert,

– Lesen mindestens einer der Speicherzellen (4) unter Verwendung der Lesereferenzzelle (6) als Referenz.

18. Verfahren nach Anspruch 14, wobei

– eine erste Referenzzelle (7) der Menge von Hauptreferenzzellen (7) vorprogrammiert ist, um eine Schwellenspannung aufzuweisen, die einem programmierten Zustand einer Speicherzelle (4, 6) entspricht,

– eine zweite Referenzzelle (7) der Menge von Hauptreferenzzellen (7) vorprogrammiert ist, eine Schwellenspannung aufzuweisen, die einem gelöschten Zustand einer Speicherzelle (4, 6) entspricht, und

– eine dritte Referenzzelle (7) der Menge von Hauptreferenzzellen (7) vorprogrammiert ist, um eine

Schwellenspannung aufzuweisen, die die Grenze zwischen einem gelöschten Zustand und einem programmierten Zustand einer Speicherzelle (4, 6) definiert, und

– die Menge von Arrayreferenzzellen (6) eine Lesereferenzzelle (6) umfasst,

wobei das Verfahren die folgenden Schritte umfasst:

– Setzen der Lesereferenzzelle (6) auf einen Lesezustand unter Verwendung der dritten Hauptreferenzzelle (7) als Referenz,

– Lesen mindestens einer der Speicherzellen (4) unter Verwendung der Lesereferenzzelle (6) als Referenz.

19. Verfahren nach Anspruch 14, mit den folgenden Schritten:

– Zuweisen vordefinierter Speicherzellen (4) zu einer Gruppe,

– Bereitstellen einer Statuszelle für die Gruppe von Speicherzellen (4),

– Setzen der Statuszelle in einen vordefinierten Logikzustand, wenn alle der Gruppe zugewiesene Speicherzellen (4) einen gelöschten Zustand aufweisen.

20. Verfahren nach Anspruch 19, wobei alle Zellen (4) einer Gruppe mit einer gemeinsamen Wortleitung (61) verbunden sind, wobei jede Wortleitung (61) mit einer Statuszelle verbunden ist.

21. Verfahren nach Anspruch 20, mit den folgenden Schritten:

– Lesen der Statuszelle vor dem Schreiben von Daten in die der Wortleitung (61) zugewiesenen Zellen (4),

– falls der Lesezustand der Statuszelle der vordefinierte Logikzustand ist, Ausführen einer Neuprogrammierung der programmierten Arrayreferenz (6) unter Verwendung der programmierten Hauptreferenzzelle (7).

22. Verfahren nach Anspruch 14, wobei Schritt a) während der Lebensdauer des Speicherbauelements (1) wiederholt wird.

23. Verfahren nach Anspruch 14, wobei Schritt a) ausgeführt wird, wenn ein vordefinierter Benutzerbefehl, vorgesehen für Aktionen unabhängig von dem Setzen von Arrayreferenzzellen (6), gestartet wird.

24. Verfahren nach einem der Ansprüche 15 bis 23, wobei Schritt a) von Anspruch 15 die folgenden Schritte umfasst:

– Setzen der ersten Arrayreferenzzelle (6) in einen programmierten Zustand und dann,

– Setzen der ersten Arrayreferenzzelle (6) in einen gelöschten Zustand.

25. Verfahren zum Herstellen eines nichtflüchtigen Halbleiterspeicherbauelements (1), umfassend:

Bereitstellen

– mehrerer in einem Speicherarray (5) angeordneter Speicherzellen (4),

– einer Menge von mindestens zwei programmierbaren Arrayreferenzzellen (6),

– einer Menge von mindestens zwei programmierbaren Hauptreferenzzellen (7),

Setzen mindestens einer der Hauptreferenzzellen (7), so dass sie eine Schwellenspannung aufweist, die einem gelöschten Zustand entspricht, und

Setzen mindestens einer anderen Hauptreferenzzelle (7), so dass sie eine Schwellenspannung aufweist, die einem programmierten Zustand entspricht.

26. Verfahren nach Anspruch 25, wobei die Schritte hinsichtlich des Setzens der Hauptreferenzzellen (7), so dass sie vordefinierte Schwellenspannungen aufweisen, während der Wafersortierung ausgeführt werden.

27. Verfahren nach Anspruch 25 oder 26, wobei die Schritte hinsichtlich des Setzens von Hauptreferenzzellen (7), so dass sie vordefinierte Schwellenspannungen aufweisen, beim Testen des nichtflüchtigen Halbleiterspeicherbauelements (1) ausgeführt werden.

Es folgen 13 Blatt Zeichnungen

FIG 1

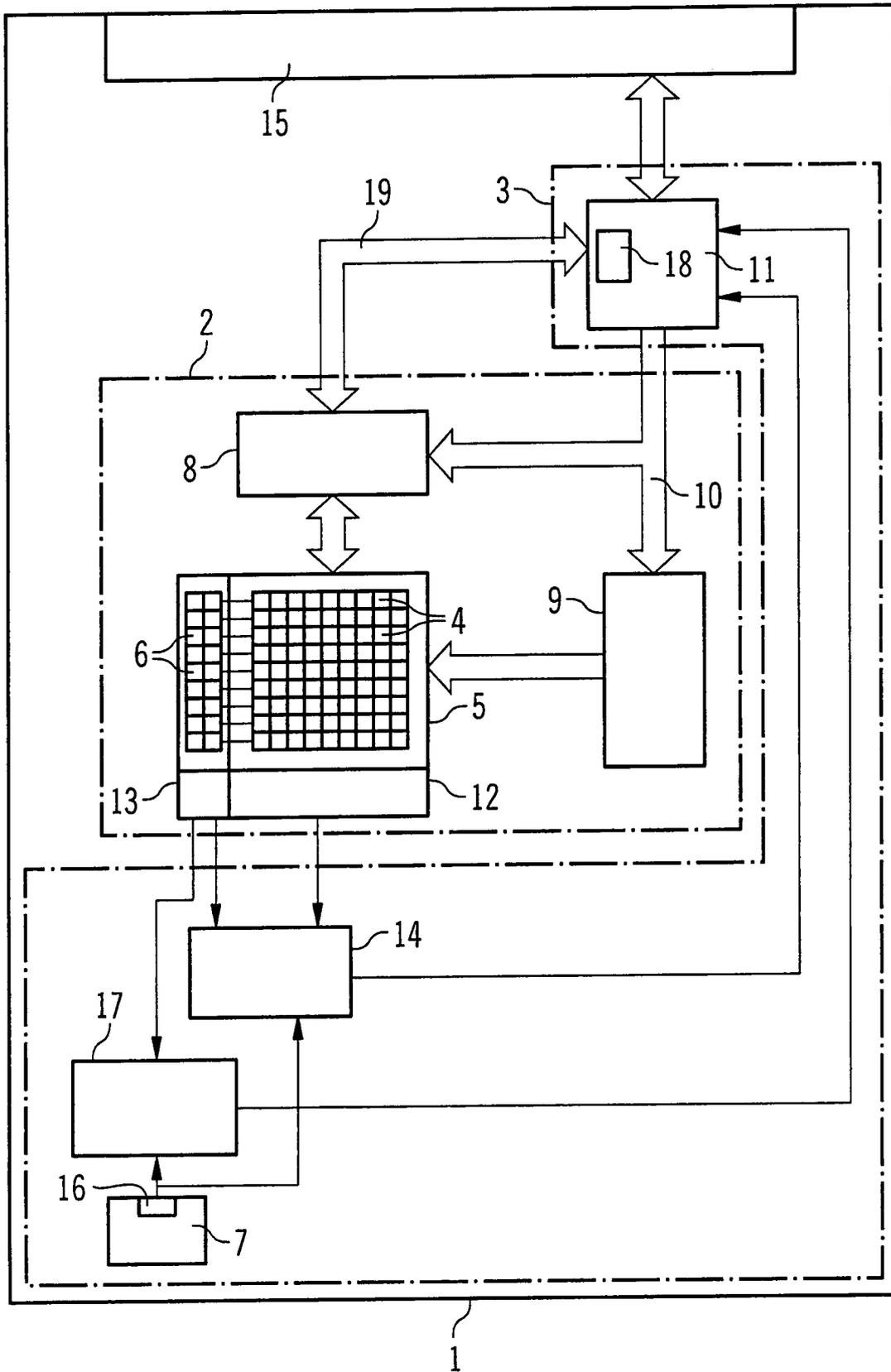


FIG 2

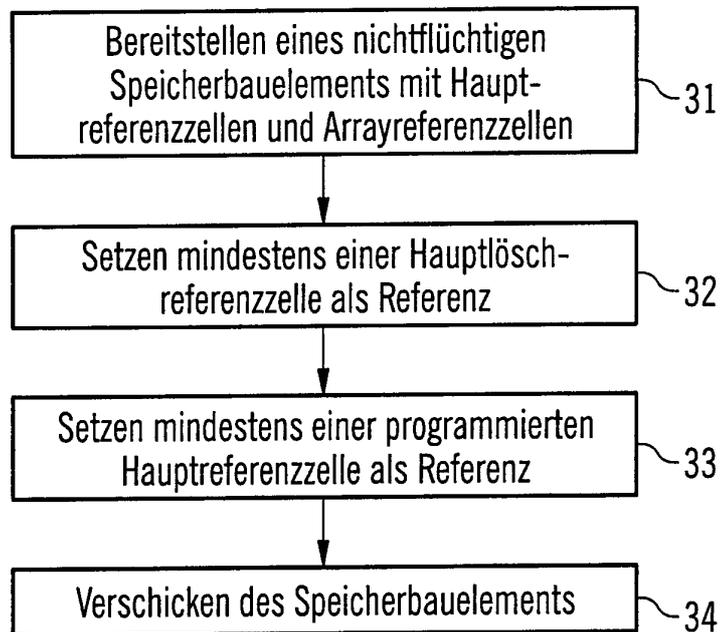


FIG 3

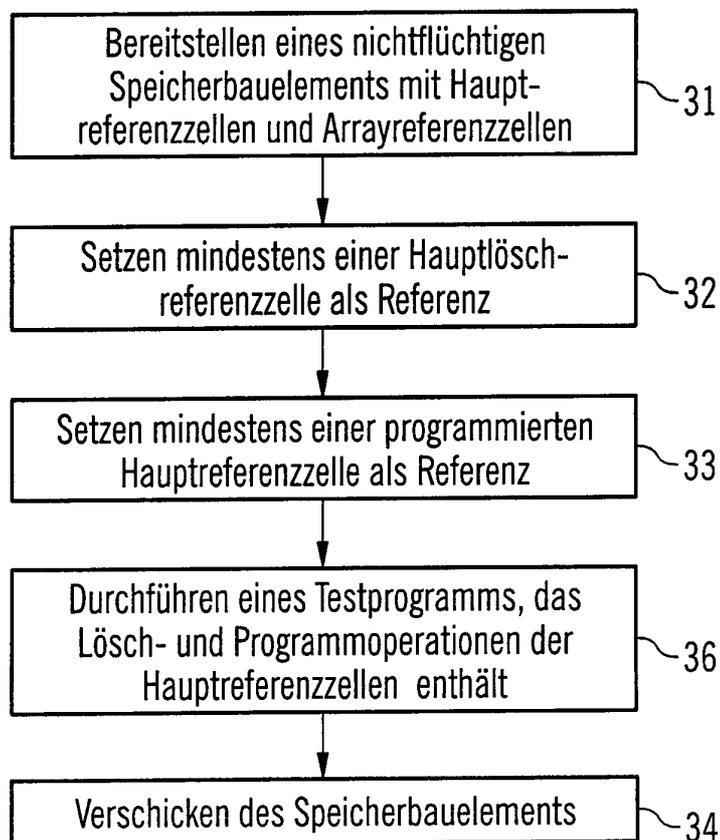


FIG 4

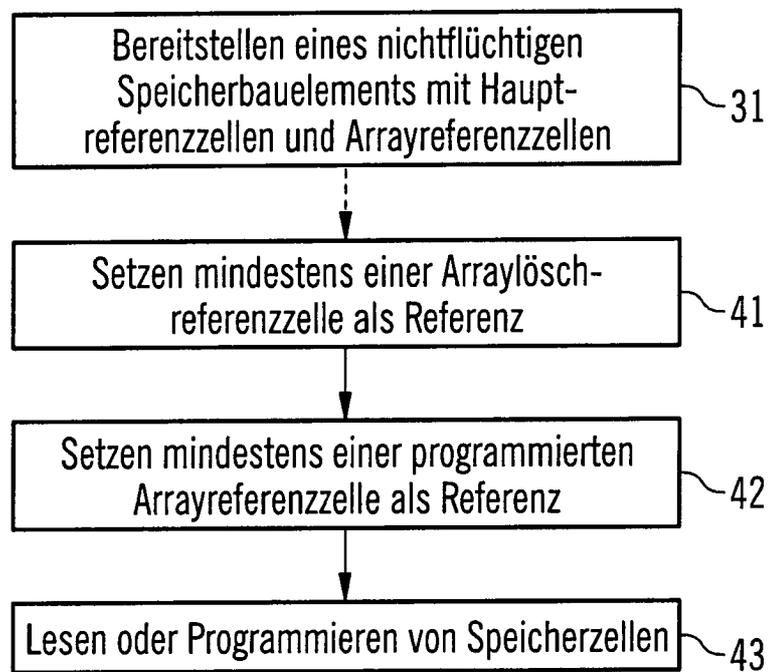


FIG 5A

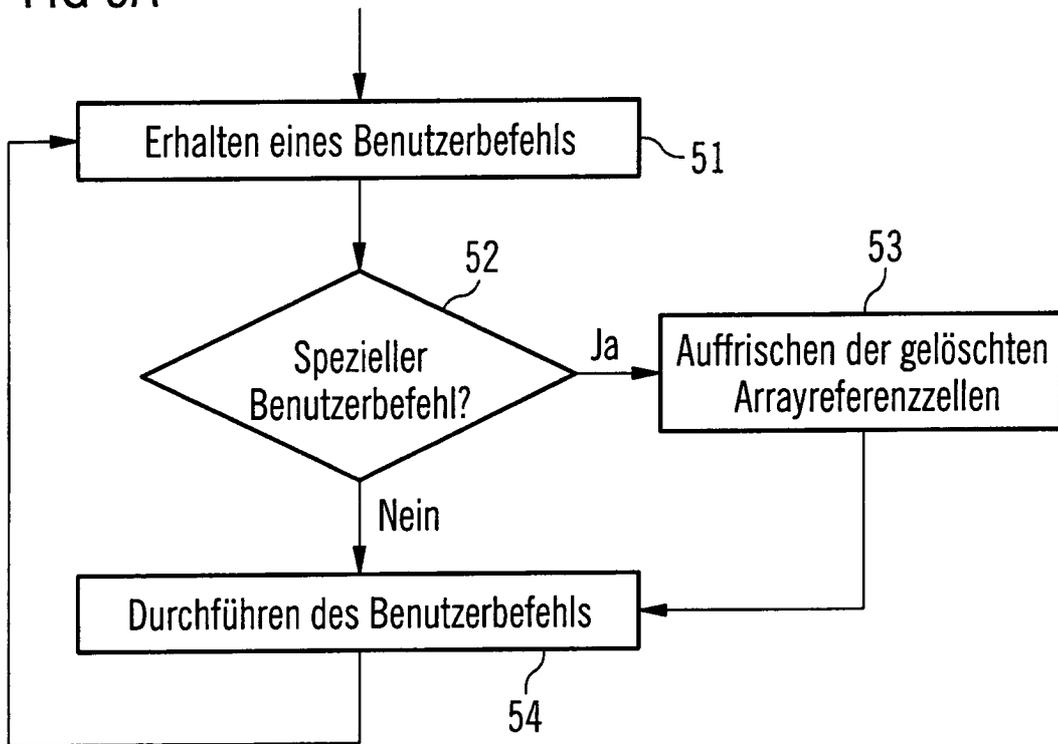


FIG 5B

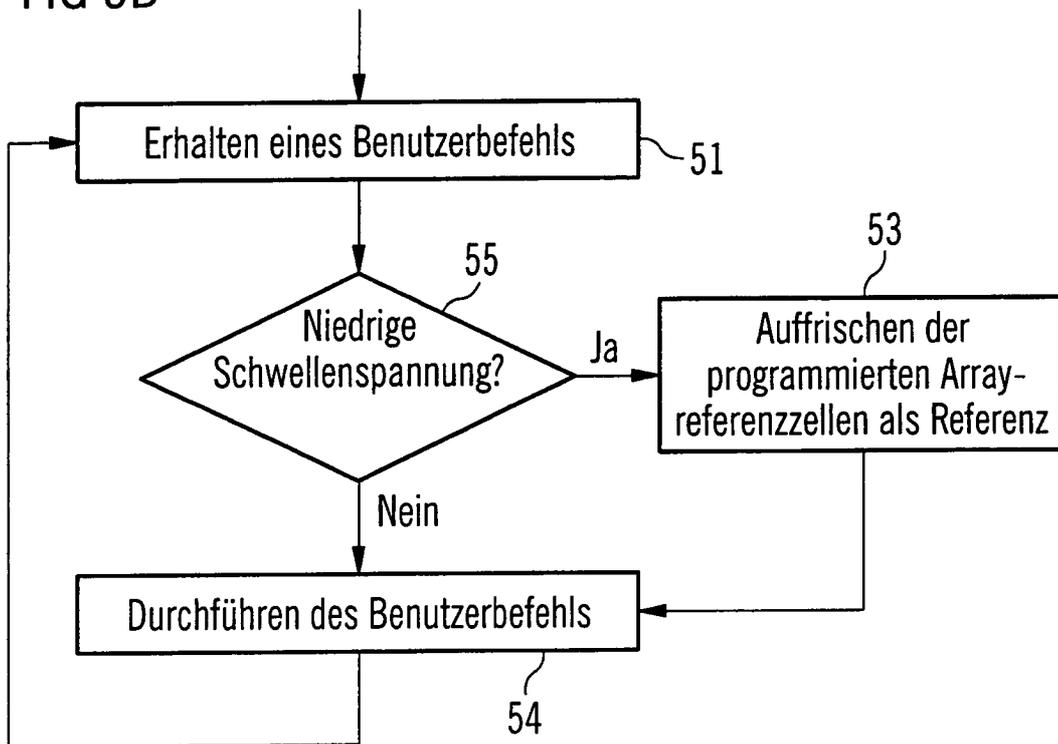
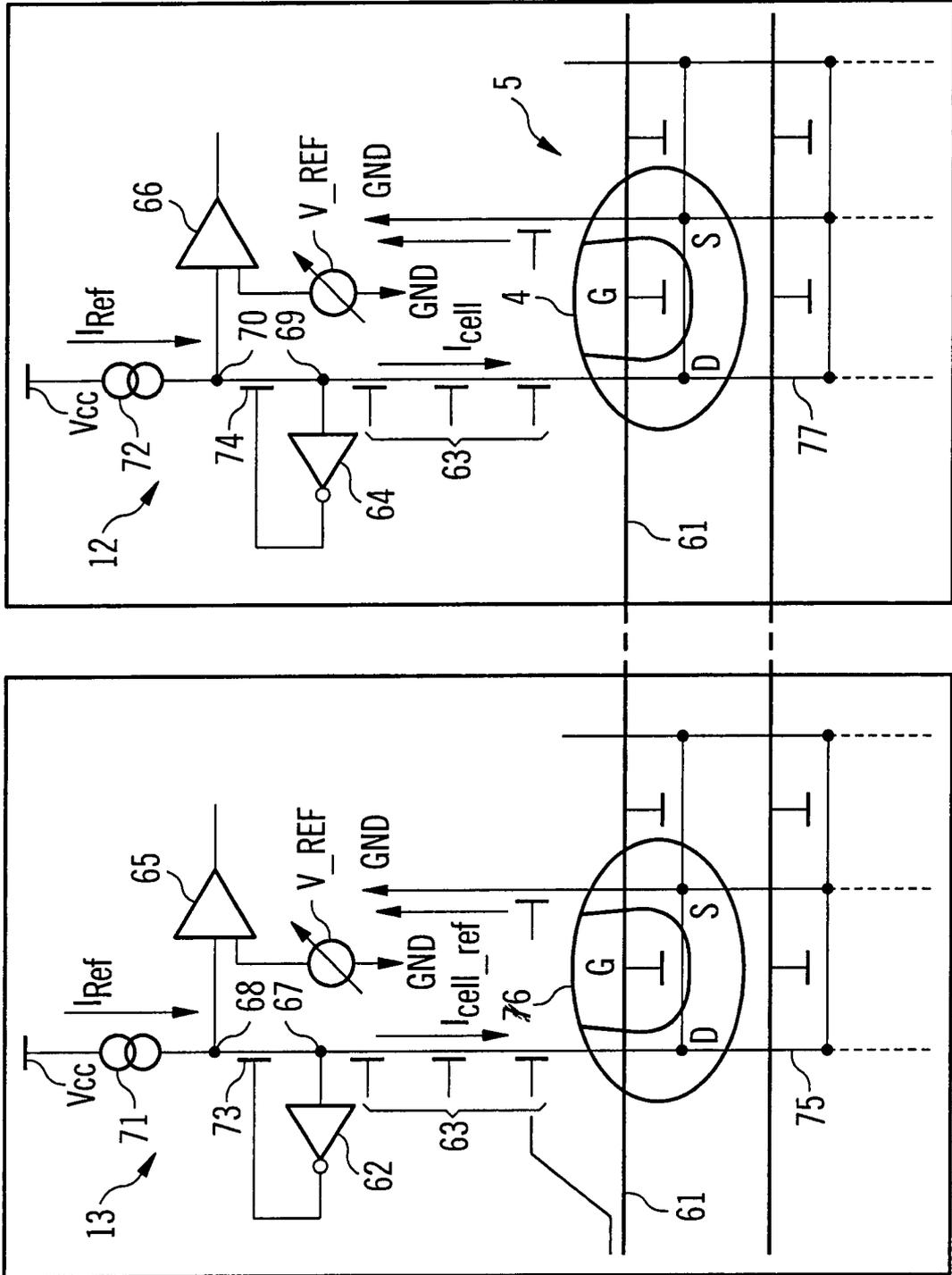
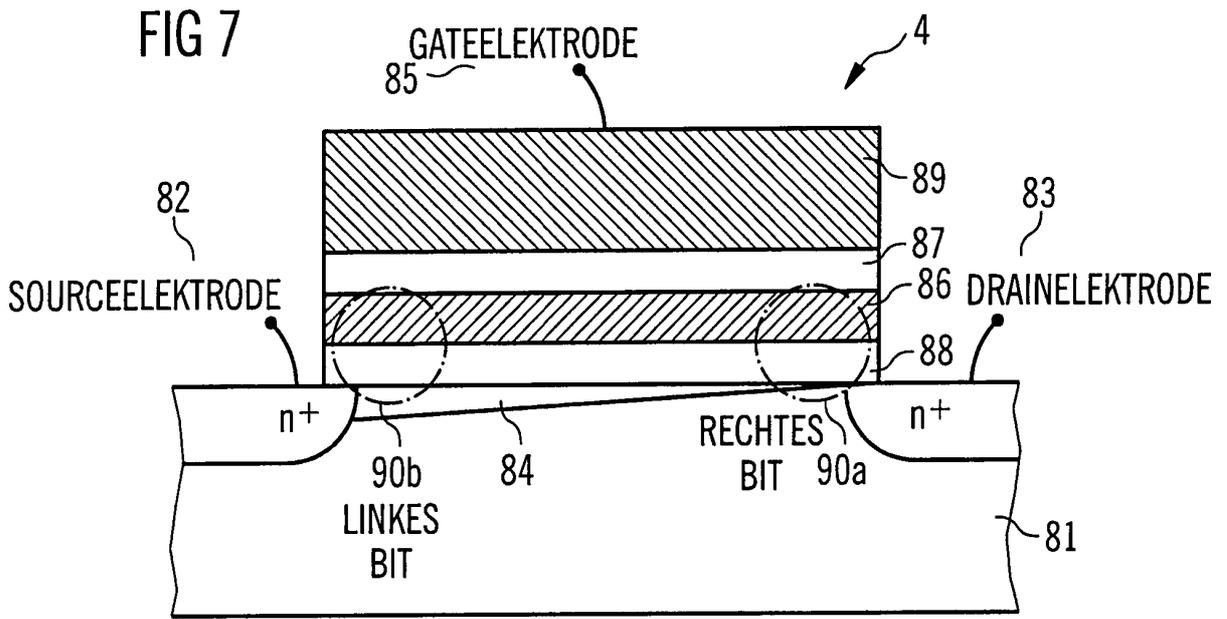
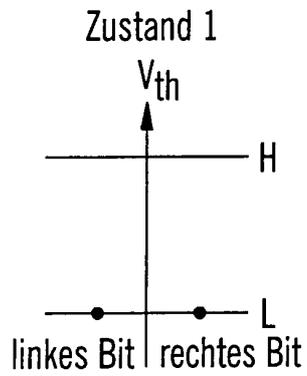


FIG 6

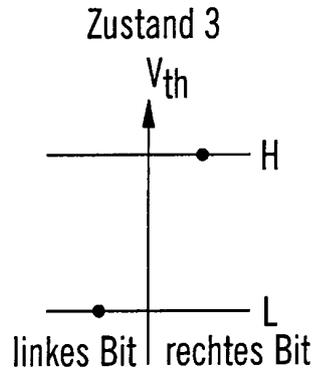




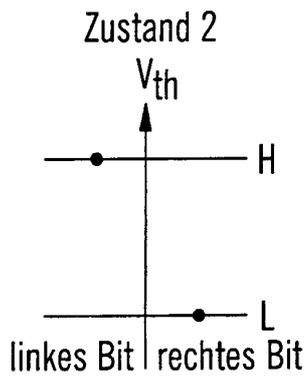
**FIG 8A**



**FIG 8C**



**FIG 8B**



**FIG 8D**

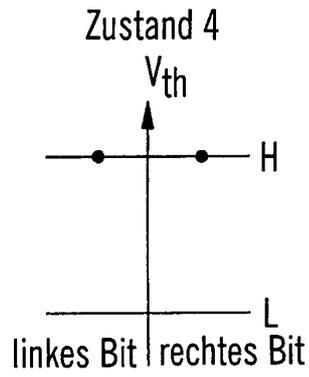


FIG 9

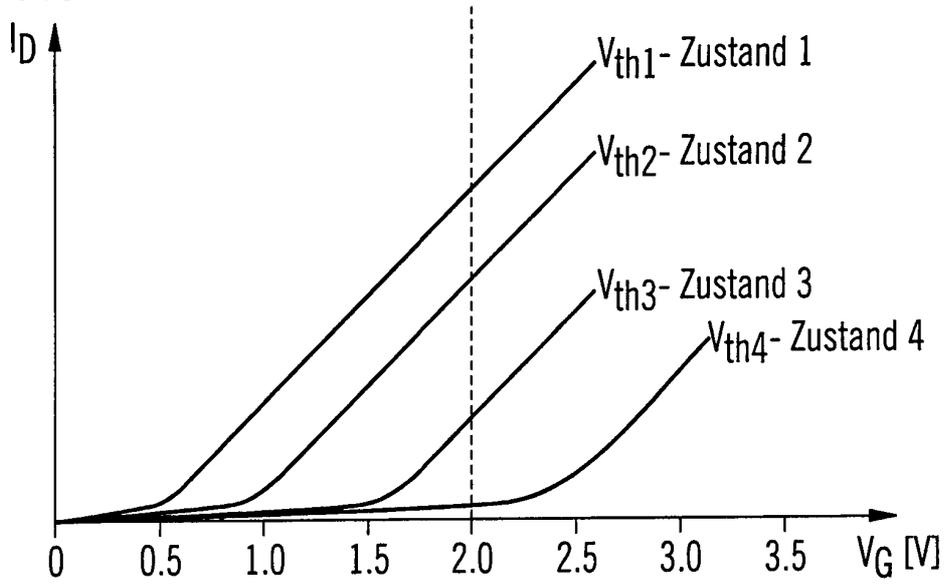


FIG 10A

Zustand 1

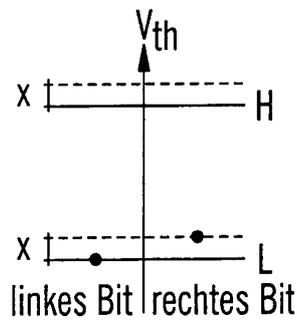


FIG 10C

Zustand 3

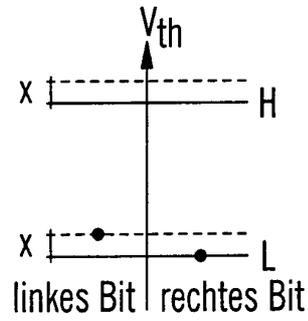


FIG 10B

zustand 2

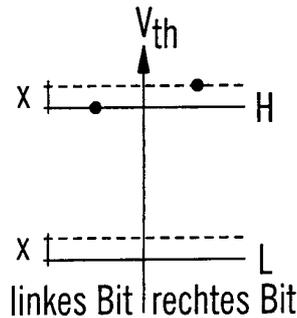


FIG 10D

Zustand 4

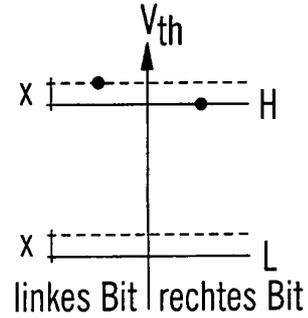


FIG 11

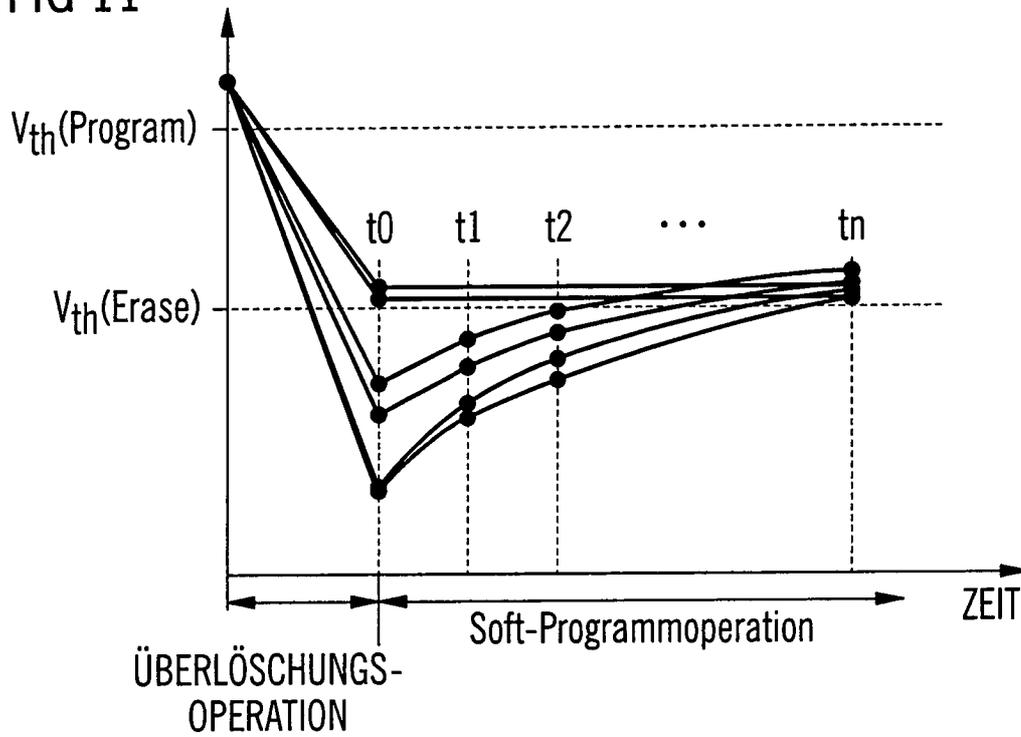
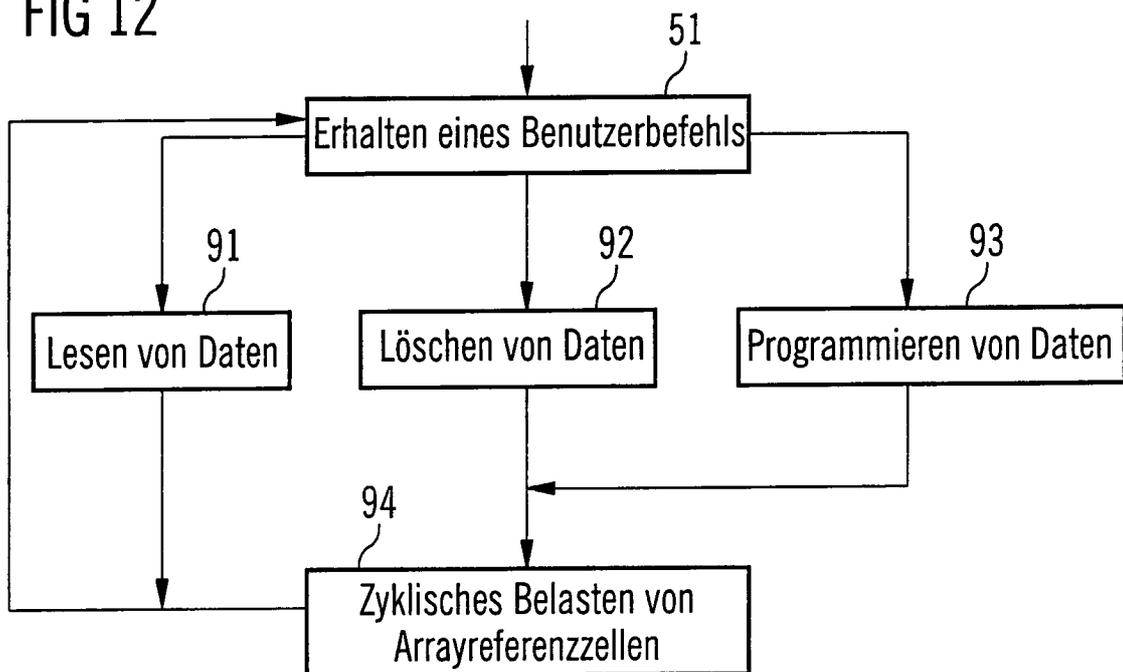


FIG 12



Benutzerblocklöschbefehl

FIG 13

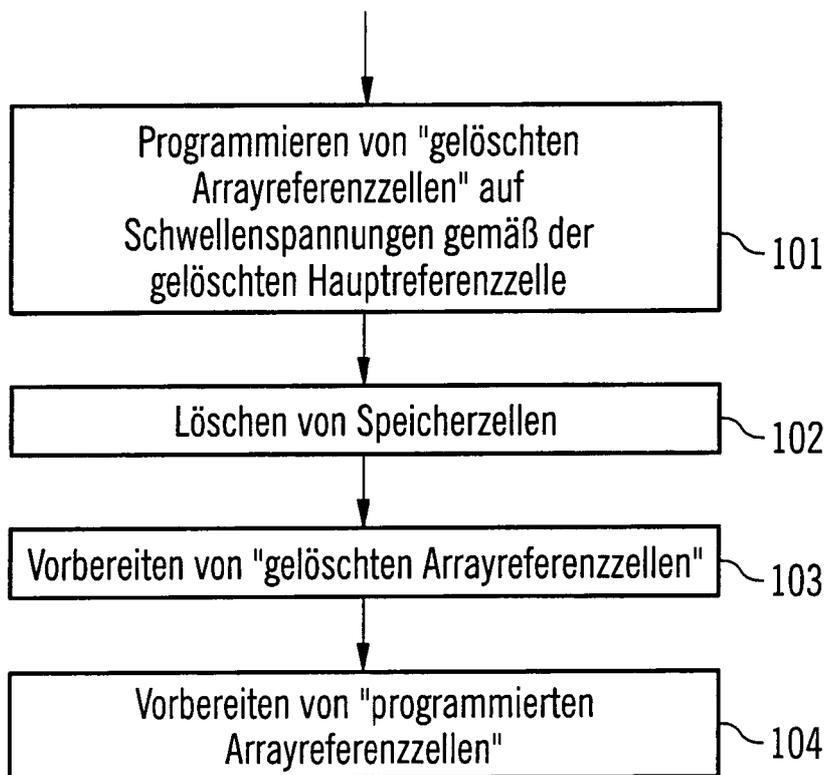
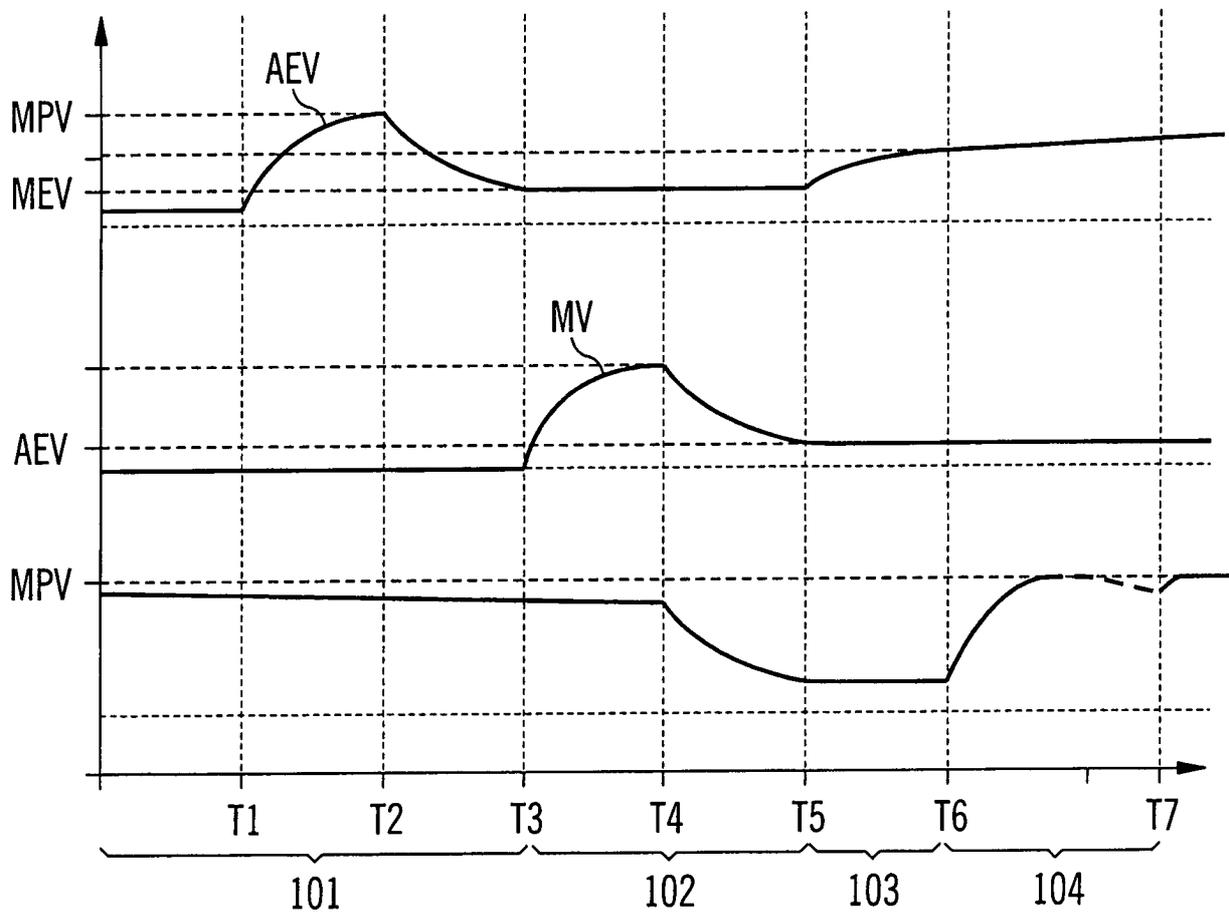


FIG 14 Benutzerblocklöschbefehl



Wortprogrammmbefehl

FIG 15

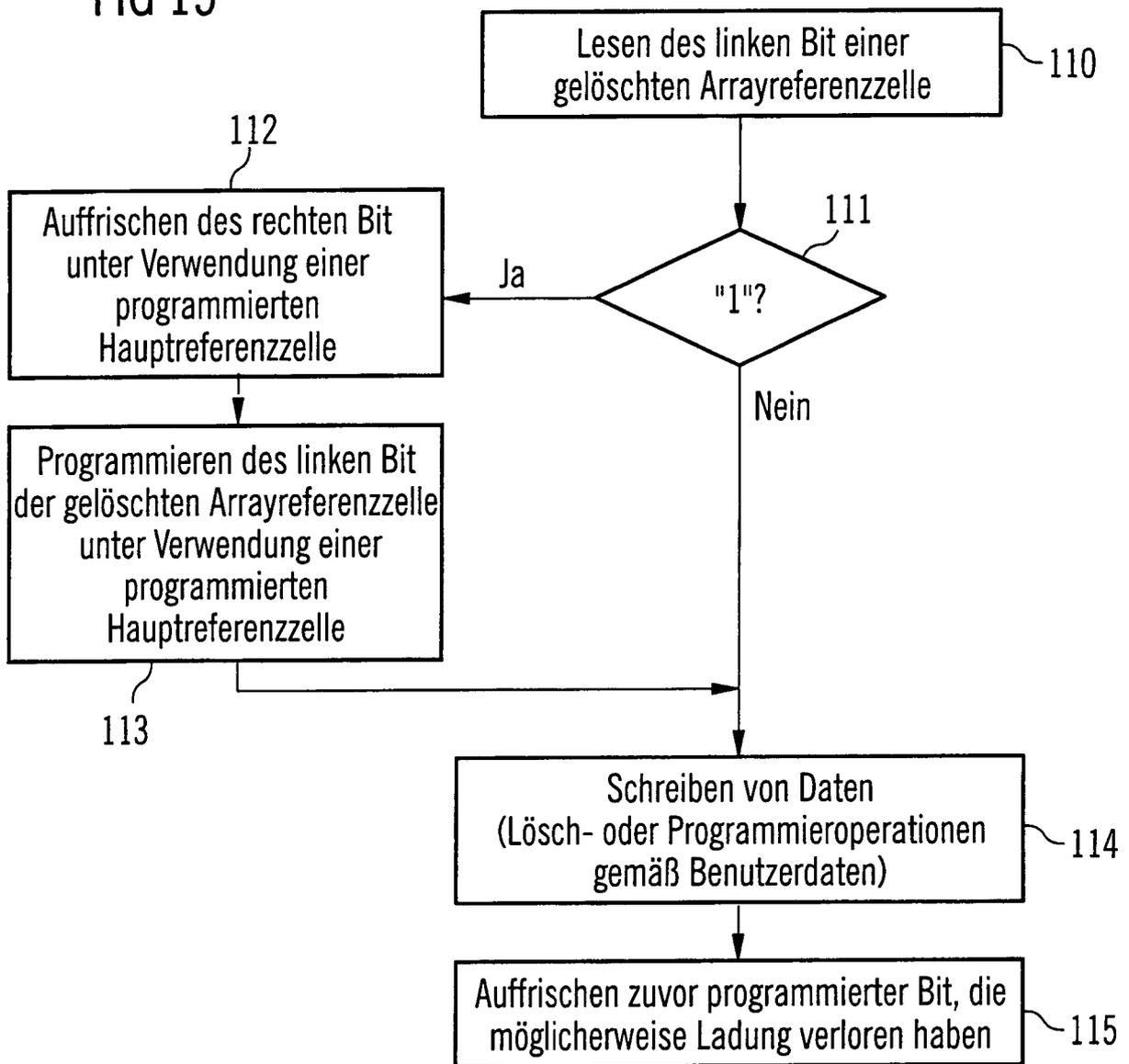


FIG 16

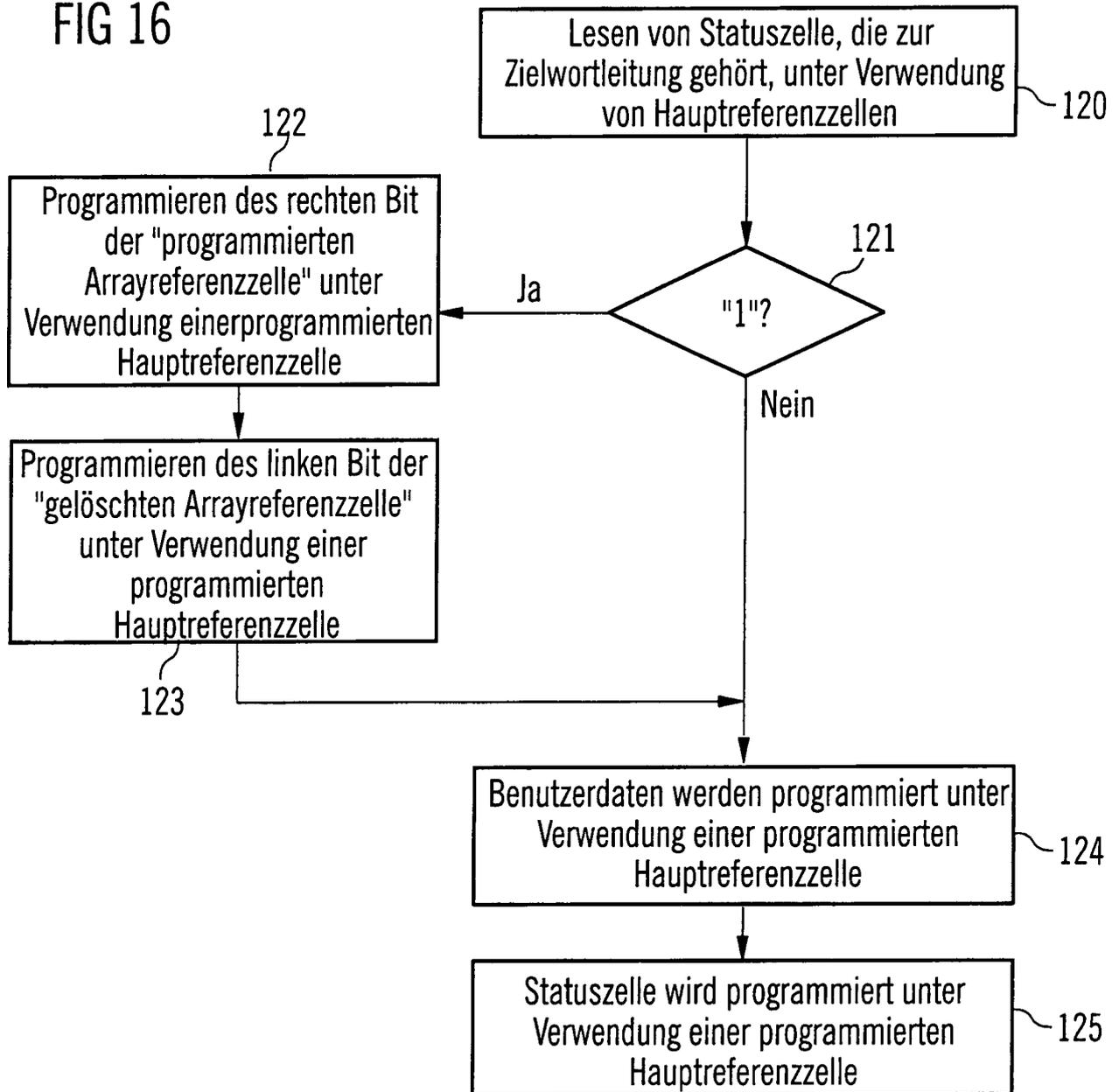


FIG 17

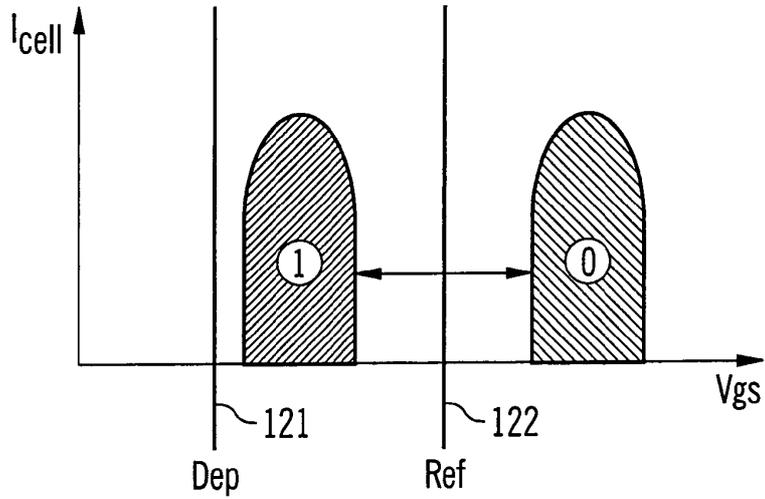


FIG 18

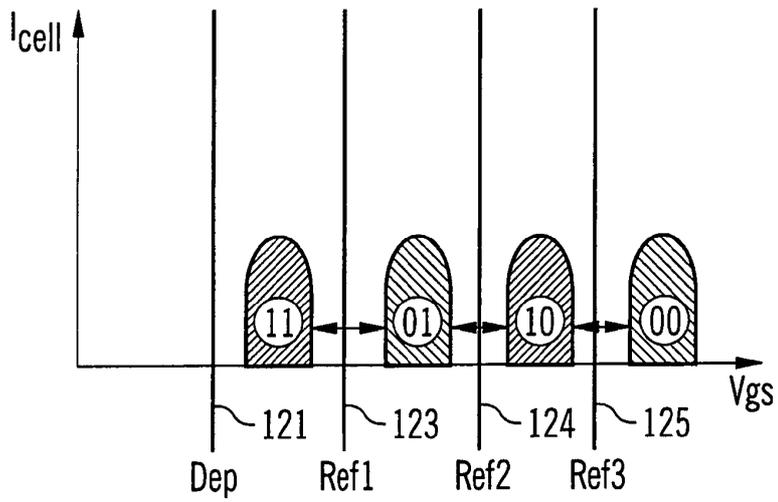


FIG 19

