(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
30 March 2006 (30.03.2006)

PCT

(10) International Publication Number
WO 2006/033717 A2

(51) International Patent Classification:
H01L 29/788 (2006.01)

(21) International Application Number:
PCT/US2005/028212

(22) International Filing Date: 9 August 2005 (09.08.2005)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
10/946,727 21 September 2004 (21.09.2004) US

(71) Applicant *(for all designated States except US)*: **ATMEL CORPORATION** [US/US]; 2325 Orchard Parkway, San Jose, CA 95131 (US).

(72) Inventor: **LOJEK, Bohumil**; 4830 Langdale Way, Colorado Springs, CO 80906 (US).

(74) Agent: **SCHNECK, Thomas**; Schneck & Schneck, P.O. Box 2-E, San Jose, CA 95109-0005 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

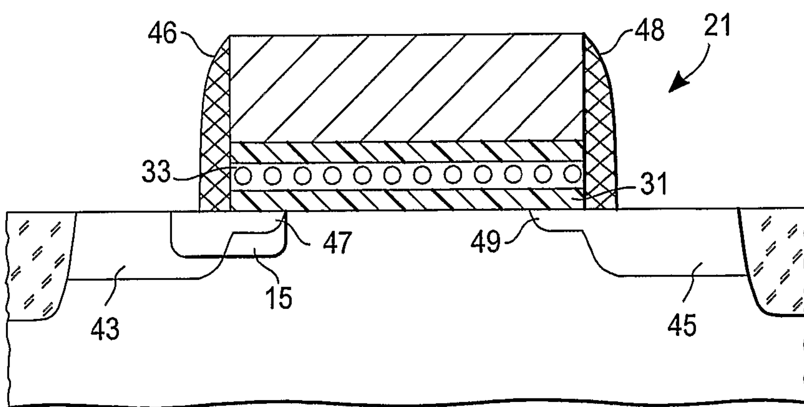
(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:
— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii)) for all designations*
— *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii)) for all designations*

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) **Title:** NON-VOLATILE NANOCRYSTAL MEMORY TRANSISTORS USING LOW VOLTAGE IMPACT IONIZATION



(57) Abstract: A low voltage non-volatile charge storage transistor (21) has a nanocrystal layer (33) for permanently storing charge until erased. A subsurface charge injector (15, 151, 152) generates secondary carriers by stimulating electron-hole current flowing toward the substrate (10), with some carriers impacting charge in nanocrystals. The charge injector is a p-n junction diode (15, 43, 47) where one polarity is source and drain electrodes (43 & 47, 45 & 49) and the other polarity is two split doped regions (151, 152) in the substrate partially overlapping the active area (11) on opposite sides of the active area. Any misalignment of masks for making the injected doped portions is inconsequential because a misalignment on one side of the active area offsets the corresponding misalignment on the other side. The injector implanted portions with overlap in the active area always have the same total area in the active area. This leads to programming reliability.

WO 2006/033717 A2

Description

NON-VOLATILE NANOCRYSTAL MEMORY TRANSISTORS USING LOW VOLTAGE IMPACT IONIZATION

5

TECHNICAL FIELD

The invention relates to non-volatile memory transistors and, in particular, to non-volatile nanocrystal memory transistors operating at low voltage
10 for programming and erasing.

BACKGROUND ART

Impact ionization has been known for several years. U.S. Patent No. 4,432,075 to B. Eitan and U.S.
15 Patent No. 4,821,236 to Hayashi et al. describe an EEPROM transistor adjacent to a charge generator, creating a substrate current near the EEPROM, creating excess charge or holes, resembling space charge, near subsurface electrodes of the EEPROM. Assume that the holes are
20 generated and accelerated toward one of the electrodes of the EEPROM. Resulting secondary electrons are sufficiently energetic to penetrate gate oxide over the substrate and become injected into a conductive floating gate. For very small EEPROMs, the floating gate becomes
25 charged by band-to-band tunneling, a situation which eliminates the need for a control gate over the floating gate.

U.S. Pat. No. 6,690,059 to B. Lojek and assigned to the assignee of the present invention
30 discloses a nanocrystal EEPROM transistor where charge is stored in a nanocrystal layer. Charge is supplied from a charge storage layer for facilitating electron communication between the two layers. Other EEPROM nanocrystal devices exist in the prior art without the charge storage
35 layer, such as U.S. Pat. No. 6,320,784 to Muralidhar. In

the '784 patent, charge is moved or removed from nanocrystal elements using traditional electrodes existing in non-nanocrystal EEPROMs.

One of the problems that occurs is that either high voltages or relatively large currents are needed for programming many nanocrystal EEPROMs. Where programming is by Fowler-Nordheim tunneling, large voltages are needed. To obtain high charge, charge pumps are frequently employed. This means additional circuitry, including switches, capacitors and an oscillator. Where programming is by hot electrons, a good current is required. This, too, requires additional circuitry. An object of the invention was to devise a nanocrystal EEPROM that could be programmed at low voltage and with low current.

SUMMARY OF THE INVENTION

The above object has been achieved in an MOS transistor by providing a tiny low voltage, low current, split charged particle injector in the substrate of a chip having a nanocrystal layer EEPROM. The injector allows for low voltage generation of charged species that cause impact ionization to supply electrons for storage on the nanocrystals. The tiny current injector is a diode split by design in two halves or portions that are spaced apart and overlay the edges of the active area of the chip. Misalignments between the two portions are mutually offset because a loss to one portion is a gain to the other portion. The overall area of the injector remains the same.

In a P-substrate, the two injector halves will be P-type material since the subsurface electrodes are typically N-type. This allows formation of P-N junctions. Alternatively, in an N-well, the halves will

be N-type since subsurface electrodes would be P-type to form the P-N junctions.

BRIEF DESCRIPTION OF THE DRAWINGS

5 Fig. 1 is a side sectional plan view of a portion of a wafer substrate illustrating a first manufacturing step in making a nanocrystal memory transistor using impact ionization in accordance with the present invention.

10 Fig. 2 is a top plan view thereof.

Fig. 3 is a side sectional plan view showing a manufacturing step subsequent to the step of Fig. 1 involving placement of split injector portions in the substrate.

15 Fig. 4 is a top view thereof.

Fig. 5 is a side sectional plan view taken along lines 5-5 in Fig. 4, showing manufacturing steps subsequent to the step of Fig. 3 involving manufacture of a gate stack structure upon the substrate.

20 Fig. 6 is a side sectional plan view showing a manufacturing step further to Fig. 5.

Fig. 7 is a side sectional plan view of Fig. 6 with a programming voltage plan.

Fig. 8 is a voltage versus distance plot for 25 the device of Fig. 7.

Fig. 9 is a side sectional plan view of Fig. 6 with an erase voltage plan.

Fig. 10 is a magnified view of circle 20 in Fig. 9 with a stylized representation of distances X, Y 30 and Z shown in Figs. 8 and 11.

Fig. 11 is an energy versus distance plot for the device of Fig. 9.

BEST MODE FOR CARRYING OUT THE INVENTION

With reference to Fig. 1, a P-type substrate 10 is defined in a silicon wafer with an active area 11 bounded by an isolation region periphery, such as STI region 13. Alternatively, the active area 11 may be in a P-well in an N-type substrate. In the top view of Fig. 2, the active area 11 is seen to be rectangular and surrounded by the isolation region 13.

In Fig. 3, the active area 11 of the substrate 10 has been implanted with a doped P-type injector region 15. The injector region is made in two portions that may be halves and so there are two implants 151 and 152 seen in Fig. 4.

In Fig. 4, the injector halves 151 and 152 are each P-type doped regions that overlap the boundary of the active area 11 by approximately fifty per cent. The regions of overlap are indicated by hatching. By making the injector regions in two halves, any misalignment in overlap with the boundary of the active area on one side of the active area will offset on the other side. Thus, mask misalignments are inconsequential insofar as the total amount of injector region being within the active area. For example, if one injector half has sixty per cent of its area in the active region, the other injector half will have forty per cent of its area in the active region. The injector portions or halves are positioned where they will overlap with one or the other of source and drain of a transistor device to be built in subsequent stages.

In Fig. 5, a gate stack 21 has been fabricated on the surface 23 of a semiconductor substrate in the active area 11. The stack 21 partly overlaps the P-type injector region 15. The overlap is provided such that the injector 15 will be in contact with a later fabricated source or drain in order to form a P-N

junction. The stack 21 is formed by etching a plurality of overlying layers. The first or lowermost layer is an insulative tunnel oxide layer 31 having a thickness in the range of 10-50 Angstroms, followed by an overlying nanocrystal charge storage layer 33, an overlying insulative control oxide layer 35, and an overlying conductive control poly layer 37.

The nanocrystal layer 33 may be formed by any of the various methods known in the art. For example, silicon atoms may be implanted into a dielectric material. A subsequent annealing step causes the implanted silicon atoms to group together through phase separation to form the nanocrystals. Alternatively, amorphous silicon may be deposited on top of the tunnel oxide layer 31, as disclosed, followed by a subsequent annealing step to recrystallize the amorphous silicon into nanocrystals. Other techniques utilize LPCVD nucleation and growth process to form crystalline nanocrystals directly on a tunnel dielectric layer. Nanocrystals are typically from 3 nm to 6 nm (30 Å - 60 Å) in diameter with a surface density of about $5(10^{12})/cm^2$ to $1.0(10^{13})/cm^2$, but other sizes and surface densities are known.

The control poly layer 37 is the first layer of polysilicon and has a thickness in the range of 1.5 - 2.5 kÅ. After layers 31, 33, 35, and 37 have been deposited, they are masked and etched to form stack 21.

The gate stack 21 has near vertical sidewalls, such as sidewall 41, that allows self-aligned implantation of subsurface n+ regions 43 and 45, seen in Fig. 6. In Fig. 6, after implantation of regions 43 and 45, n+ LDD regions 47 and 49 are formed as extensions of the subsurface source and drain regions 43 and 45. After formation of the n+ LDD regions 47 and 49, the sidewall spacers 46 and 48 may be established. The spacers are

typically made of nitride and are formed by disposing a nitride layer over the wafer to the height of the top of the stack 21. The nitride is partially removed, leaving spacers 46 and 48 but removing the remainder of the nitride layer. The spacers prevent entry or loss of mobile ions or electrons from the charge storage layer of the stack 21.

The source and drain regions 43 and 45, together with extensions 47 and 49 are subsurface electrodes of the same n-conductivity type that interact with the charge storage layer of stack 21 by forming a channel between source and drain regions 43 and 45. One of the source and drain regions 43 and 45 forms a P-N junction with injector 15 so that a diode is formed relative to the P-substrate 10. This P-N junction, when reverse biased, generates hole and electron (e-h) current mostly flowing to the substrate. The e-h current results in space charge within the active area 11, with whole impaction on the source and drain electrodes 43 and 45. The secondary electrons generated by impact ionization are sufficiently energetic as to be available for tunneling through the tunnel oxide layer 31 to the nanocrystal layer 33. The tunneling mechanism for electrons from the substrate through the tunnel oxide to the nanocrystal layer is thought to be hot electron transfer. In other words, by activation of the P-N junction involving injector 15 by a relatively low voltage, i.e. less than 5 volts, programming of the nanocrystal layer occurs.

With reference to Fig. 7, a programming voltage, V_g, of approximately 5 volts is applied to the control poly member 37 by contact 51. The source 43 has a contact wire 53 where a positive voltage, less than the programming voltage, is applied. The voltage V_s, applied to wire 53, is somewhat less than +5 volts to establish a

gradient between control poly member 37 and source region 43. At the same time, the voltage applied on line 53 is sufficient to reverse bias the P-N junction formed by the injector region 15 and the source 43. This generates e-hole current in substrate 10 which is grounded by a grounding contact 55. Some of the carriers in the substrate impact upon the source and drain electrodes 43 and 45, creating electron space charge, some of which is accelerated toward control poly electrode 37 and the contact wire 51 carrying the positive voltage V_g. These electrons, known as hot electrons, penetrate the tunnel oxide and some electrons become trapped in nanocrystal layer 33. The drain electrode 45 is floating. A contact wire 57 is not connected.

The electrical potential diagram of Fig. 8 is similar to the diagram of Fig. 2 in U.S. Patent No. 4,821,236. Substrate 10 is shown to have a width dimension, S. In Fig. 8, the lower curve 61 represents positive voltage applied at the top of the mesa structure as seen by an electron in the substrate. The positive voltage V_g is attenuated across the distance X representing the thickness of the control oxide. At the nanocrystal layer, the electrical potential becomes higher represented by line 63 and increases across the nanocrystal layer having a thickness Y with the increasing potential represented by line 65. Upon encountering tunnel oxide, the electrical potential again falls indicated by line 67 and is attenuated toward the substrate, the tunnel oxide thickness being indicated by the letter Z until the substrate 10 is reached whereupon the electrical potential again diminishes, indicated by line 69. An inverse voltage arrangement is experienced by holes where the highest negative voltage, indicated by line 71, is closest to the substrate. Traversing the tunnel oxide, voltage increases across the tunnel oxide

or thickness dimension Z to line 73. The electrical potential increases to line 75 and then decreases to line 77 forming a potential well 80. Electrons can be trapped in nanocrystals which form due to quantum confinement in
5 a potential well, such as electron 81. Here the electron is shown in well 80 (polarity reversal) so that the electron appears to be trapped. Electrical potential increases across the thickness of the control oxide apex to line 79.

10 When programming voltages are applied, as shown in Fig. 7 electrons, such as electron 83, seen in Fig. 8, overcome the potential barrier of the tunnel oxide and tunnel through the oxide as indicated by the horizontal arrows A and wind up in a potential well, similar to
15 potential well 80 (polarity reversal), where they will remain indicating a specific memory state such as a digital 1 or a digital 0.

In the erase mode, voltages are applied as shown in Fig. 9, but the transistor structure is the same
20 as in Fig. 7. Note that circle 20 in Fig. 9 has been magnified in Fig. 10 to illustrate distances X, Y and Z. The Y dimension is a very thin nanocrystal layer. The X dimension is above the nanocrystal layer but below the control gate and the Z dimension is below the nanocrystal
25 layer and above the substrate. Returning to Fig. 9, wire 53 is floating and therefore carries no electrical potential to source 43. The wire 51, in contact with the control poly layer 37, is biased at a negative voltage, -V_g. Wire 57 is floating and so no electrical potential is
30 transferred to drain electrode 45. The negative voltage on control poly has an influence on the transistor device, as shown in Fig. 11. Here the electron 81 is influenced by the negative gate voltage V_g which may be up to -10 volts. The electron 81 may recombine with holes
35 that tunnel through the thickness, Z, of the tunnel

oxide, for example hole 82 that becomes trapped in potential well 84, also within the Y dimension. The hole 82 can recombine with electron 81, indicated by wavy line B. The charge from the nanocrystal region is reduced and channel conductance will change accordingly.

Reading of the state of the device depends upon channel conductivity. When charge is stored in the nanocrystal layer, channel conductivity is different from the situation where charge is not stored. The conductivity of the channel is ascertained by sense transistors, not described in this application.

Claims

1. An EEPROM transistor comprising:
a substrate of semiconductor conductivity having a conductive contact,
a gate stack structure disposed on the substrate, the gate stack structure having a nanocrystal layer disposed in an insulative manner over the substrate and an overlying control layer insulated from the nanocrystal layer,
source and drain electrodes in the substrate flanking the gate stack structure, both of a first conductivity type,
an injector electrode in the substrate of a second conductivity type in contact with one of the source and drain electrodes thereby forming a p-n junction in the substrate, and
bias means for creating a space charge in the substrate from the p-n junction with current flow into the conductive contact of the substrate and into the nanocrystal layer thereby programming the nanocrystal layer.

2. The EEPROM transistor of claim 1 wherein the injector electrode has two portions, both portions in contact with one of the source and drain electrodes.

3. The EEPROM transistor of claim 2 wherein said substrate has an active area bounded by isolation regions, the two portions of the injector overlapping the active area at spaced apart locations.

4. The EEPROM transistor of claim 1 wherein said bias means comprises a first low voltage source in contact with one of the source or drain electrode in P-N communication with the injector electrode and a second
5 low voltage source in contact with the control layer.

5. The EEPROM transistor of claim 4 wherein the other of the source or drain not in contact with the first low
10 voltage source is electrically floating.

6. The EEPROM transistor of claim 1 wherein the conductive contact of the substrate is electrical ground.
15

7. The EEPROM transistor of claim 1 wherein the gate stack structure at least partially overlies the injector electrode.
20

8. The EEPROM transistor of claim 1 wherein the source and drain electrodes are self-aligned with the gate stack structure.
25

9. The EEPROM transistor of claim 1 wherein the substrate is P-type semiconductor material, the injector electrode is P-type semiconductor material and the source
30 and drain electrodes are N-type semiconductor material.

10. The EEPROM transistor of claim 1 wherein the gate stack structure has vertical sidewalls with insulative
35 spacers capping the sidewalls.

11. A method of making an EEPROM transistor comprising,
forming an active area in a wafer substrate with spaced apart opposed sides,
fabricating a source, drain, channel in the active area below a nanocrystal gate insulatively over and between the source and drain and a control gate insulatively over the nanocrystal gate,
fabricating an injector as a P-N junction in relation to at least one of the source and drain, the injector fabricated in two portions, each portion overlapping an opposed side of the active area.

12. The method of claim 11 wherein said two injector portions are halves.

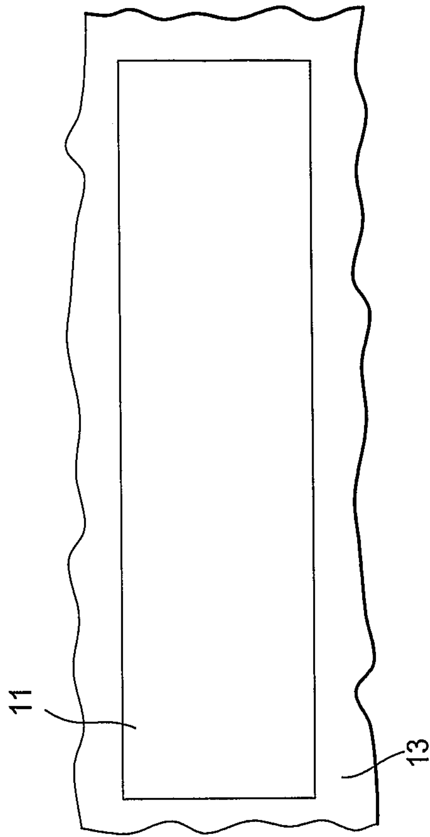


Fig._2

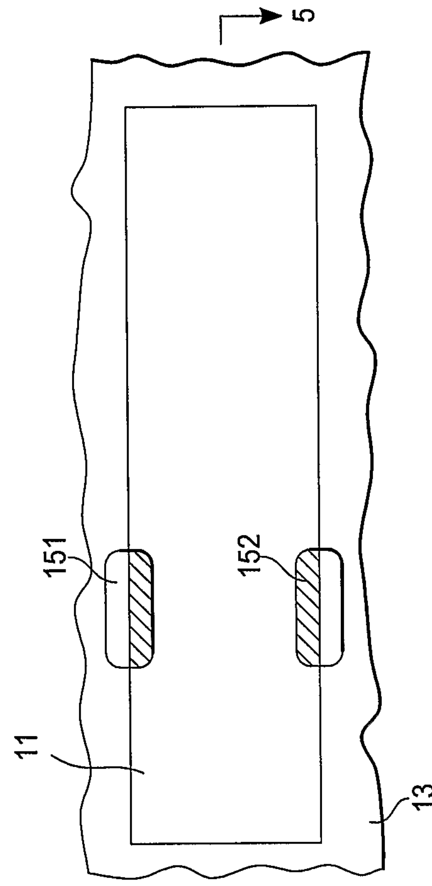


Fig._4

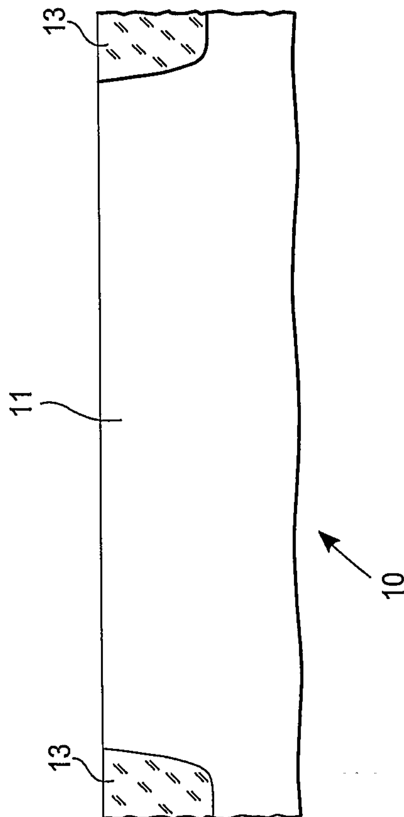


Fig._1

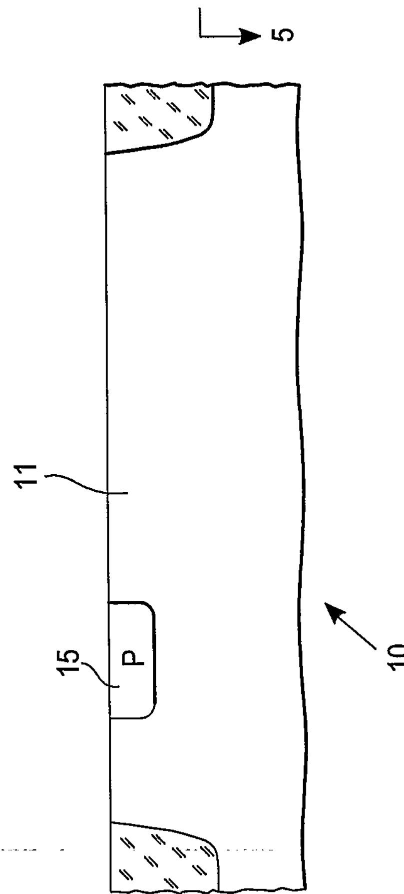


Fig._3

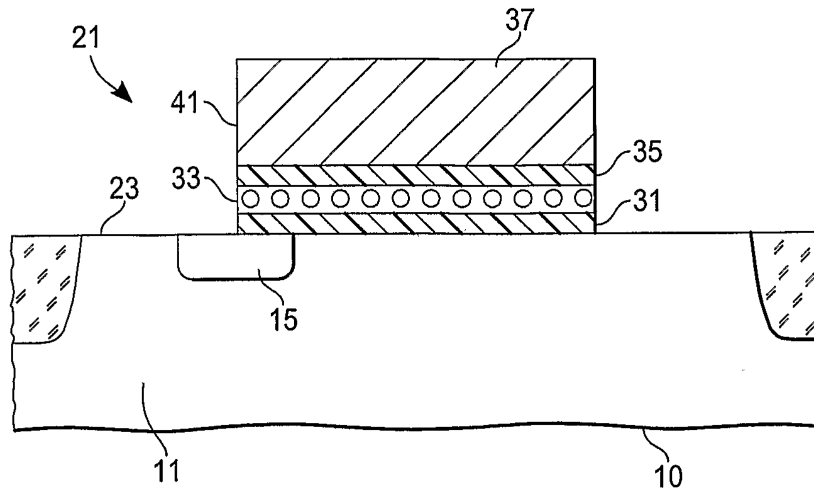


Fig._5

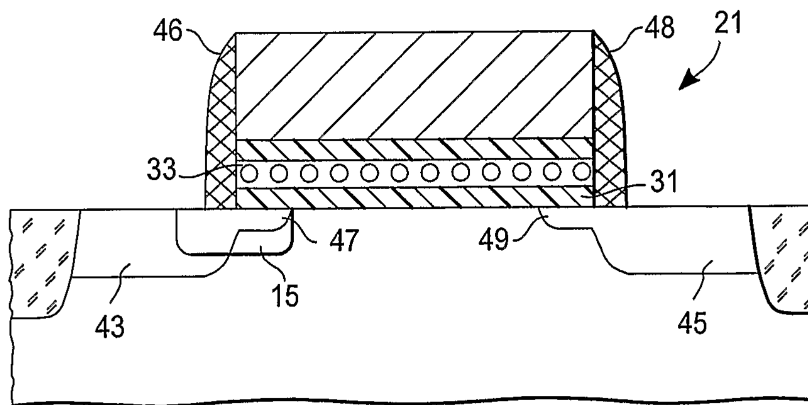


Fig._6

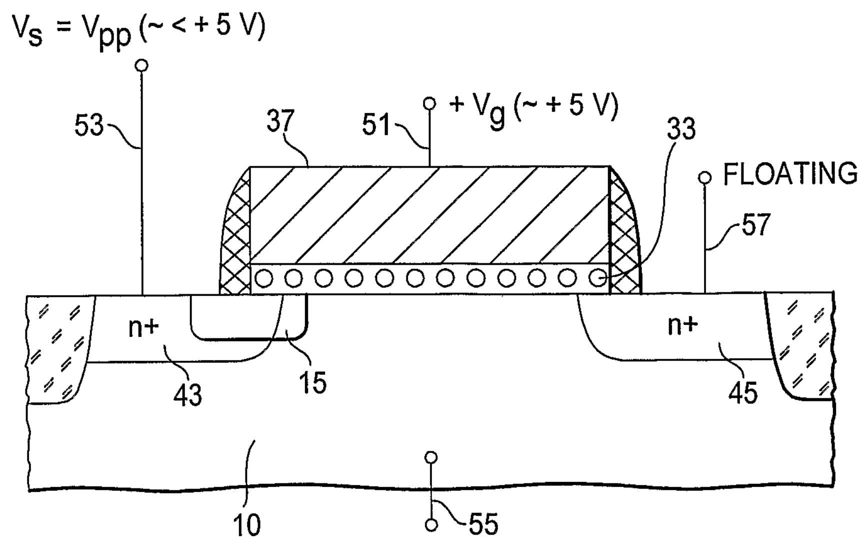


Fig._7

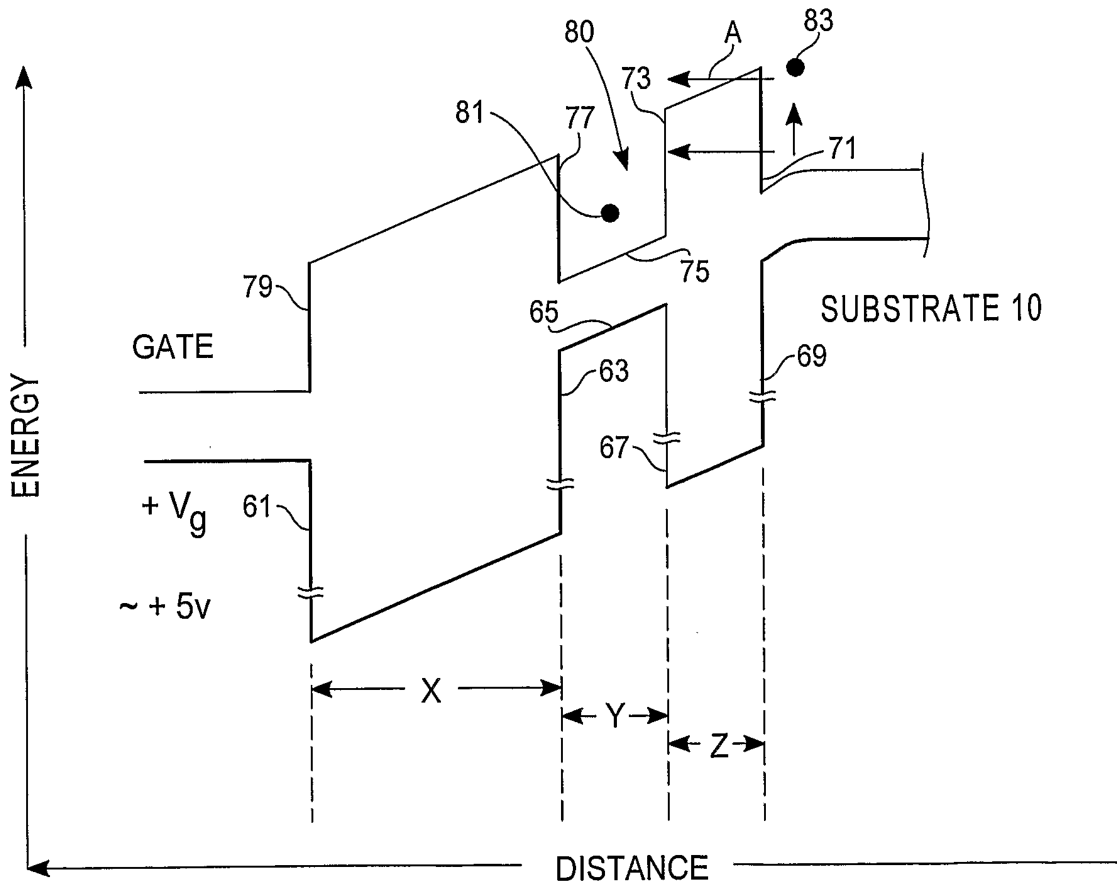


Fig._8

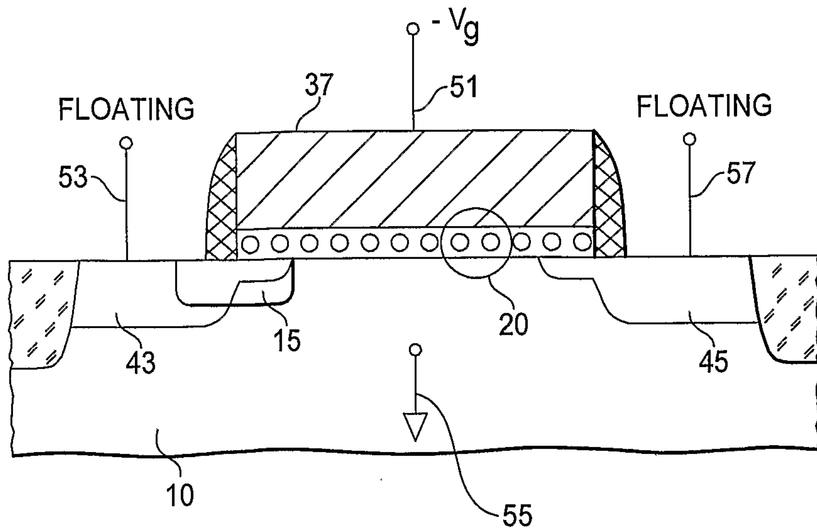


Fig._9

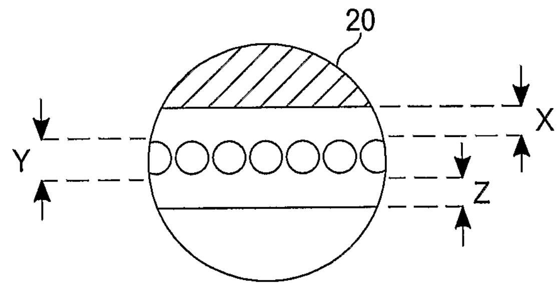


Fig._10

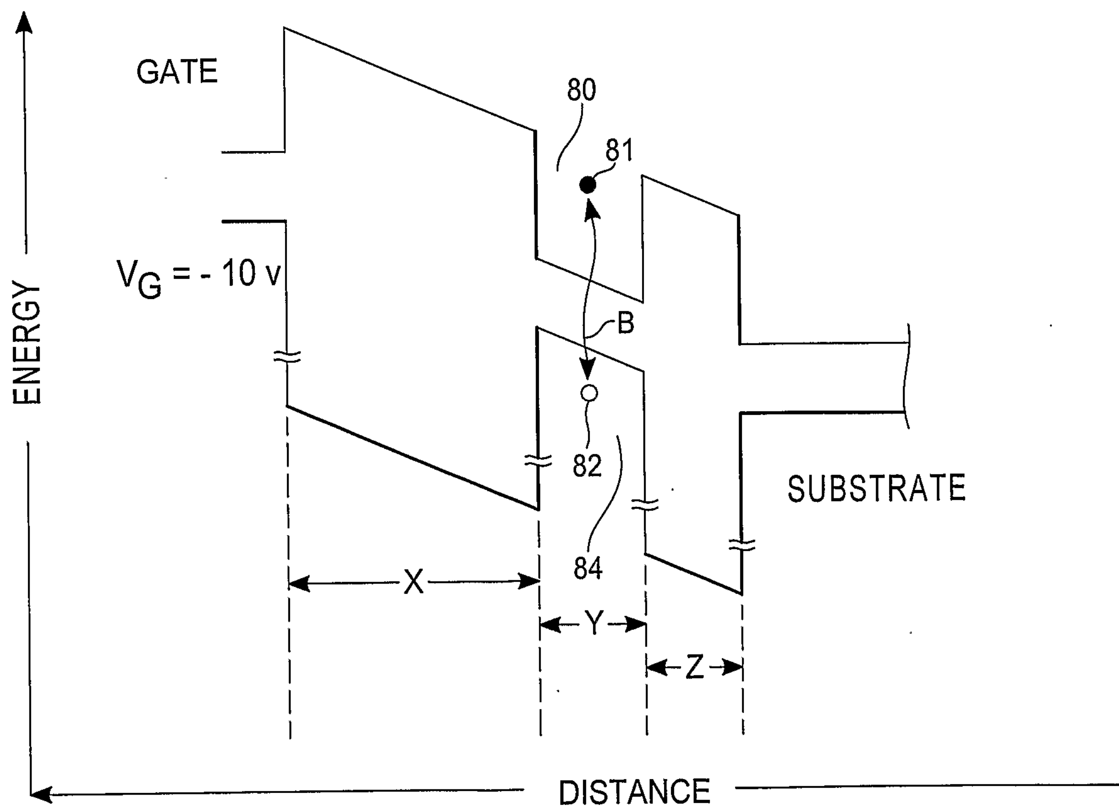


Fig._11