

"Electrically Programmable Read Only Memory".

The present invention relates to an electrically programmable read only memory (hereinafter, abbreviated to "EPROM").

A known EPROM with its write-down efficiency enhanced has been disclosed in the Specification of British Patent Application No. 81.19691 (published under Serial No. 2,).

In the above disclosed EPROM, a non-volatile memory element thereof has a structure which includes a floating gate electrode and a control gate electrode overlying the former gate electrode, and a high impurity concentration region having the same conductivity type as that of a substrate is formed in that peripheral portion of the semiconductor surface of the non-volatile memory element which underlies a gate insulating film and which extends in a channel direction. In the non-volatile memory element of such a structure, when a source - drain current flows during a write-down operation, the presence of the high impurity concentration region improves the generation efficiency of hot electrons, so that the write-down speed is enhanced.

The specification of the above-mentioned Application, however, does not teach a construction which solves other problems in EPROMs and a new problem occurring in the case of the structure formed with the high impurity concentration region as stated above.

We have found that in order to raise the write-down efficiency of charges into a floating gate, a write-down voltage to be applied to a non-volatile memory element should desirably be a greater magnitude in a range in which a pn-junction does not substantially break down.

However, in the case where the high impurity concentration region is arranged as stated above, the pn-junction thus formed by the drain of the non-volatile memory element and the high impurity concentration region causes the avalanche breakdown even at a comparatively low reverse voltage.

In the case where the pn-junction of the non-volatile memory element has undergone the avalanche breakdown due to a write-down voltage controlled unsatisfactorily, holes created in the pn-junction according to the avalanche breakdown are injected into the interface between a silicon substrate and silicon dioxide underlying the floating gate of the non-volatile memory element. Furthermore, we have found as a result of investigations, electrons which are injected by rewrite-down recombine with the holes in the floating gate, and the rewrite-down efficiency deteriorates.

Accordingly, the write-down voltage must have its level controlled satisfactorily.

In an integrated circuit, a plurality of non-volatile memory elements are arranged in a matrix form. The drains of a plurality of non-volatile memory elements arranged in the same column are coupled with a single bit line in common, while the control gates of a plurality of non-volatile memory elements arranged in the same row are coupled with a single word line in common.

In accordance with such arrangement, one non-volatile memory element is selected through a given word line and a given bit line.

In order to write down an information into one non-volatile memory element, a write-down voltage is supplied to one bit line (selected bit line) with which the non-volatile memory element is coupled, and a word line signal at a selection level is supplied to one word line (selected word line).

The high level of the write-down voltage is a comparatively high voltage, and raises the potential of the floating gate to an unnegligible level through the parasitic capacitance between the drain and floating gate of the non-volatile memory element.

As a result, channel leak currents flow to the non-selected non-volatile memory elements coupled with the selected bit line. The number of non-volatile elements to be coupled with one bit line increases in proportion to the increase of the memory capacity. Accordingly, the leakage currents increase with the increase of the memory capacity. In other words, in cases where the memory capacity has been increased, the idle current increases to render the power consumption unduly high.

When the leakage currents have increased excessively, a write-down current which flows to the selected non-volatile memory element decreases accordingly.

Accordingly, the current driving ability of a write-down circuit must be enhanced. This inevitably necessitates an increase in the size of elements which constitute the write-down circuit.

The leakage currents fluctuate comparatively greatly due to, not only the manufacturing dispersions of the non-volatile memory elements, but also the variation of the operating temperature. Accordingly, a write-down circuit appropriately designed is required. Unless the appropriate write-down circuit is designed, the yield of products will lower.

It has also been known that the leakage currents are reduced by connecting the sources of a plurality of non-volatile memory elements to earth through a single common resistor as disclosed in German Offenlegungsschrift 3,037,315.

We have found that one disadvantage of the known non-volatile memory elements is that since the magnitudes of the leakage currents depend upon the resistance of the common resistor, great dispersions are involved in the leakage currents in accordance with fluctuations in the manufacturing factors of the integrated circuit.

It is an object of the present invention to provide an EPROM which can reduce its total leakage current to a very small value and in which the total leakage current is not substantially affected by fluctuations in the manufacturing factors of an integrated circuit.

According to the present invention there is provided an electrically programmable read only memory including:

- (a) a first word line;
- (b) a plurality of non-volatile memory elements, each having source, drain and control electrodes, the control electrodes being connected in common to the first word line;
- (c) a first switch MISFET having source, drain and gate electrodes; and
- (d) a source of reference potential;
at least one of the source electrodes of the plurality of non-volatile memory elements being connected to the source of reference potential through the drain-source path of the first switch MISFET, the gate electrode of the first switch MISFET being connected to the first word line.

In one preferred form, the control gate electrodes of a plurality of non-volatile memory elements are connected to a first word line, the sources of the plurality of non-volatile memory elements are connected to earth through the drain-source path of a single switching MISFET in common, and the gate electrode of the switching MISFET is connected to the first word line.

In another preferred form, the control gate electrodes of a plurality of non-volatile memory elements are connected to a first word line, the control gate electrodes of a further plurality of non-volatile memory elements are connected to a second word line, the sources of the former non-volatile memory elements and those of the latter non-volatile memory elements are connected to earth through the drain-source path of a first switching MISFET and that of a second switching MISFET in common, the gate electrode of the first switching MISFET is connected to the first word line, and the gate electrode of the second switching MISFET is connected to the second word line.

In the preferred forms described above, in the case where a word line driving signal at a non-selection level is applied to one word line, the switching MISFET whose gate is connected to this word line becomes non-conductive, so that leakage current is inhibited from flowing to the switching MISFET.

In an EPROM of very large memory capacity (for example, 64 kilobits), the last mentioned preferred form is the most favourable from both the viewpoints of the integration density and the leakage current reduction.

The present invention will now be described in greater detail by way of example with reference to the accompanying drawings, wherein:-

Figure 1 shows a block diagram of a preferred form of an EPROM;

Figures 2(A) and 2(B) show a circuit diagram of essential portions of the EPROM shown in Figure 1;

Figure 3 shows a circuit diagram of a sense amplifier which is included in the EPROM;

Figures 4 and 5 are circuit diagrams showing modified embodiments of a write-down circuit;

Figure 6(A) shows a plan view of a non-volatile memory element which is included in the EPROM;

Figure 6(B) shows a view of one section in Figure 6(A) taken along the line B-B', and

Figure 6(C) shows a view of another section in Figure 6(A) taken along the line A-A';

Figures 7(A) to 7(F) are process sectional views for explaining the method of manufacturing the EPROM shown in Figure 1;

Figure 8(A) shows a structural sectional view of a zener diode of the write-down circuit in the EPROM; while

Figure 8(B) shows a plan view corresponding to Figure 8(A); and

Figure 9(A) shows a plan view of a non-volatile memory array which includes switching MISFETs, while

Figure 9(B) shows a view of one section in Figure 9(A).

Referring first to Figure 1, the part enclosed with a dotted line IC is formed within a single silicon chip by the method of manufacturing a semiconductor integrated circuit to be described later which method includes well-known semiconductor manufacturing techniques.

The EPROM has input terminals which are supplied with address signals A_0 to A_{12} , input terminals which are supplied with control signals $\overline{PGM}$, $\overline{CE}$ and $\overline{OE}$, input/output terminals I/O_0 to I/O_7 , a first power source terminal which is supplied with a power source voltage V_{CC} of, e.g., 5 volts, a second power source terminal which is supplied with a reference voltage GND such as earth potential, and a write-down power source terminal which is supplied with a write-down voltage V_{pp} that is made a value of, e.g. 25 volts during a programing operation (that is, during the write-down of information). The EPROM is operated by supplying the respective terminals with the appropriate voltages or signals.

The EPROM includes memory array circuits MAR_0 to MAR_7 (hereinafter written as ' MAR_0 to MAR_7 ') divided into eight groups. Each of the MAR_0 to MAR_7 circuits is furnished with a memory capacity of 8192 bits, whereby the EPROM has a total memory capacity of about 64 kilobits.

One memory array circuit is constructed of 256 word lines, 32 bit lines, 256 x 32 memory cells, and switching elements for reducing leakage currents as will be described later.

An X-address buffer circuit ADB_1 (hereinafter written as ' ADB_1 ') responds to the externally supplied address signals $A_5 - A_{12}$ of 8 bits at, e.g. TTL (Transistor Transistor Logic) levels, and provides address signals of a true level and a false level $a_5, \overline{a_5}$ to $a_{12}, \overline{a_{12}}$ respectively made appropriate levels.

By receiving the address signals $a_5, \overline{a_5}$ to $a_{12}, \overline{a_{12}}$, an X-decoder circuit X-DEC (hereinafter written as 'X-DEC') forms word line driving signals W_1 to W_{256} to be fed to the word lines of the MAR_0 to MAR_7 circuits.

Only one of the word line driving signals W_1 to W_{256} as corresponds to one combination of the address signals A_5 to A_{12} is brought to a selection level, whereas all the others are brought to a non-selection level such as earth potential.

By supplying the X-DEC circuit with the power source voltage V_{CC} , the write-down voltage V_{PP} and a control signal We from a control circuit CONT to be described later, the selection level of the word line driving signal is at a level which is substantially equal to the power source voltage V_{CC} during a read-out operation of information and a level which is substantially equal to the voltage V_{PP} during a write-down operation of information.

A Y-address buffer circuit ADB_2 (hereinafter written as ' ADB_2 ') responds to the externally fed address signals $A_0 - A_4$ of 5 bits at, e.g. the TTL levels, and provides address signals of the true level and the false level $a_0, \overline{a_0}$ to $a_4, \overline{a_4}$.

The address signals $a_0, \overline{a_0}$ to $a_4, \overline{a_4}$ are

used for selecting the 32 bit lines of each of the memory array circuits MAR_0 to MAR_7 .

The EPROM is arranged so that the bit lines of the respective memory array circuits MAR_0 to MAR_7 are selected by column switch circuits CS_0 to CS_7 and read-out/write-down circuits R/W_0 to R/W_7 .

Accordingly, among the address signals a_0 , $\overline{a_0}$ to a_4 , $\overline{a_4}$, those a_0 , $\overline{a_0}$ to a_2 , $\overline{a_2}$ are supplied to a first Y-decoder circuit Y-DEC₁ (hereinafter written as 'Y-DEC₁'), and the remaining signals a_3 , $\overline{a_3}$ and a_4 , $\overline{a_4}$ are supplied to a second Y-decoder circuit Y-DEC₂ (hereinafter written as 'Y-DEC₂').

By receiving the address signals a_0 , $\overline{a_0}$ to a_2 , $\overline{a_2}$, the Y-DEC₁ circuit forms bit line selection signals BS_1 to BS_8 to be supplied to the column switch circuits CS_0 to CS_7 .

Only one of the bit line selection signals BS_1 to BS_8 as corresponds to one combination of the address signals A_0 to A_2 is brought to a selection level, whereas all the others are brought to a non-selection level such as earth potential.

By supplying the Y-DEC₁ circuit with the voltages V_{CC} and V_{PP} and the control signal We similar to those of the X-DEC circuit, the selection level of the bit line selection signal is at a level which is substantially equal to the power source voltage V_{CC} during the read-out operation of information and a level which is substantially equal to the voltage V_{PP} during the write-down operation of information.

By receiving the address signals a_3 , $\overline{a_3}$ and a_4 , $\overline{a_4}$, the Y-DEC₂ circuit forms selection signals

K_0 to K_3 .

Only one of the selection signals K_0 to K_3 which corresponds to one combination of the address signals A_3 and A_4 is brought to a selection level substantially equal to the power source voltage V_{CC} , whereas all the others are brought to a non-selection level such as earth potential.

The Y-DEC₂ circuit is supplied with a control signal Ce from the control circuit CONT. By way of example, the control signal Ce is at a high level during the selection of the chip and a low level during the non-selection of the chip. The selection signals K_0 to K_3 which are provided from the Y-DEC₂ circuit are held at the non-selection level during the chip non-selection irrespective of the levels of the address signals A_3 and A_4 .

By receiving the above bit selection signals, the column switch circuits CS₀ to CS₇ (hereinafter written as 'CS₀ - CS₇') select the bit lines of the corresponding MAR₀ - MAR₇ circuits in numbers of four at the same time. That is, four of the 32 bit lines B₁ to B₃₂ of, for example, the MAR₀ circuit are coupled with common bit lines CB₀ to CB₃ through the column switch circuit CS₀.

By way of example, the signals to be switched by the column switch circuits CS₀ to CS₇ are at a high level which is substantially equal to the power source voltage V_{CC} or at a low level which is substantially equal to earth potential during the read-out operation of information, and they are at a high level which is

substantially equal to the voltage V_{pp} or at a low level which is substantially equal to earth potential during the write-down operation of information.

5 To the end of coupling the bit lines to be selected and the common bit lines with a low voltage loss, the selection level of the bit line selection signals which are supplied to the column switch circuits CS_0 to CS_7 is at the level substantially equal to V_{pp} during the write-down operation of information as described above.

10 The sets of common bit lines, each set including the four lines, are respectively coupled to the read-out/write-down circuits R/W_0 to R/W_7 (hereinafter written as ' R/W_0 to R/W_7 ') as illustrated in the figure.

Each of the R/W_0 to R/W_7 circuits is constructed of four unit read-out/write-down circuits which correspond to the four common bit lines $CB_0 - CB_3$ one to one.

20 Input/output terminals of the four unit read-out/write-down circuits in the respective R/W_0 to R/W_7 circuits are connected in common to the common input/output terminals I/O_0 to I/O_7 .

25 Each unit read-out/write-down circuit is selected by one of the selection signals K_0 to K_3 delivered from the Y-DEC₂ circuit. When it is selected, the corresponding common bit line and the common input/output terminal are substantially coupled.

30 Each unit read-out/write-down circuit is operated by the power source voltage V_{CC} and the write-down voltage V_{pp} supplied during the write-down

operation of information.

Each unit read-out/write-down circuit has its operation mode specified by control signals We , $\overline{We}$, CS and $\overline{CS}$ provided from the control circuit CONT.

5 A read-out mode is specified by the high levels of the control signals $\overline{CS}$ and $\overline{We}$. In the read-out mode, a signal on one of the four common bit lines is supplied to the common input/output terminal through the unit read-out circuit selected by the
10 Y-DEC₂ circuit. That is, the common input/output terminal is supplied with a data stored in a memory cell selected by the X-DEC, Y-DEC₁ and Y-DEC₂ circuits.

15 A write-down mode is specified by the high level of the control signal We and the low level of the control signal $\overline{CS}$. In the write-down mode, among the unit read-out/write-down circuits supplied with data signals from a circuit (not shown) through the common input/output terminals, only one selected by the Y-DEC₂ circuit is brought into its operative status.
20 Accordingly, one of the four common bit lines is at substantially earth potential or an appropriate write-down voltage level in response to the data signal to be written down. The remaining three common bit lines are at substantially earth potential level
25 because the corresponding unit read-out/write-down circuits are not selected.

In a chip non-selection mode, the selection signals K_0 to K_3 which are provided from the Y-DEC₂ circuit are made the non-selection level as described
30 above. Therefore, all the unit read-out/write-down

circuits R/W_0 to R/W_7 are brought into their inoperative status.

Each read-out/write-down circuit is constructed so that its input/output terminals are in the floating state in the chip non-selection mode.

The control circuit (hereinafter written as 'CONT') forms the control signal $\overline{We}$ as indicated in the following Table 1, by receiving an external program control signal $\overline{PGM}$, chip enable signal $\overline{CE}$ and output enable signal $\overline{OE}$ and the write-down voltage V_{pp} at the predetermined level:

Table 1

	<u>Read</u>	<u>Program</u>
V_{PP}	5 V	25 V
$\overline{PGM}$	H	L
$\overline{CE}$	L	L
$\overline{OE}$	L	-
$\overline{we}$	H	L

Using these control signals, the selection of the write-down/read-out operations (in other words, the control of various operation modes such as read-out under a high-voltage application state) is executed.

Figures 2(A) and 2(B) show detailed circuit diagrams of the EPROM shown in Figure 1.

The MAR_0 circuit includes 256×32 non-volatile memory elements $F_{11} - F_{25632}$ which are arranged

in 32 rows and 256 columns.

Whilst its structure will be further clarified later, each individual memory element is constructed of a FAMOS (Floating gate Avalanche injection MOS transistor) which has a floating gate electrode and a control gate electrode.

Amongst the memory elements, 32 memory elements arranged in the same column have their control gate electrodes connected in common to the corresponding one of the word lines $W_1 - W_{256}$.

The memory array circuit MAR_0 is provided with reference potential lines R_1 to R_{128} each of which corresponds to two of the word lines such as the word lines W_1 and W_2 and those W_3 and W_4 . Thus, the source electrodes of the 32 memory elements belonging to one column and those of the 32 memory elements belonging to the next column adjoining the first-mentioned column are connected in common to one reference potential line.

Switching MISFETs S_1 to S_{256} which are controlled "on" and "off" by the signals fed to the word lines are respectively arranged between the reference potential lines R_1 to R_{128} and earth potential.

The drain electrodes of 256 memory elements arranged in the same row are connected in common to the corresponding one of the bit lines $B_1 - B_{32}$.

The bit lines $B_1 - B_{32}$ are divided into four groups by the column switch circuits CS_0 .

Switching MISFETs are arranged between the respective bit lines and the common bit line. More specifically, switching MISFETs $Q_{11} - Q_{18}$ which are controlled by the

selection signals formed by the Y-DEC₁ circuit are respectively arranged between the bit lines B₁ - B₈ and the common bit line CB₀.

Likewise, switching MISFETs Q₂₁ - Q₂₈, Q₃₁ - Q₃₈ and Q₄₁ - Q₄₈ similar to the aforecited ones are respectively arranged between the bit lines B₉ - B₁₆, B₁₇ - B₂₄ and B₂₅ - B₃₂ and the common bit lines CB₁, CB₂ and CB₃. Accordingly, when one selection signal, for example, BS₁ formed by the Y-DEC₁ circuit is brought to the selection level, the four bit lines B₁, B₉, B₁₇ and B₂₅ are simultaneously coupled with the common bit lines CB₀, CB₁, CB₂ and CB₃ respectively.

The common bit lines CB₀ to CB₃ are coupled with the read-out/write-down circuit R/W₀.

As shown in Figure 2(B), the read-out/write-down circuit R/W₀ is constructed of a write-down circuit W₀ and a read-out circuit R₀.

The write-down circuit W₀ comprises four unit write-down circuits W₀₁ to W₀₄ corresponding to the common bit lines CB₀ to CB₃ one to one, and a single input circuit W₀₅. The input circuit W₀₅ comprises MISFETs Q₆₂ to Q₇₀. The MISFETs Q₆₅ and Q₆₉ which have straight lines drawn between their sources and drains are of the depletion mode and the other MISFETs are of the enhancement mode.

The output terminal of the input circuit W₀₅ (i.e., the drain of the MISFET Q₇₀) delivers a signal which is equal to the logical product between the inverted signal of the write-down data signal supplied to a common input/output terminal I/O and the write-down

control signal We .

More specifically, when the write-down control signal We is at the high level, the input circuit W_{05} provides a signal at a low level which is substantially equal to earth potential or at a high level which is substantially equal to the power source voltage V_{CC} in accordance with the high level or low level of the data signal supplied to the common input/output terminal I/O. When the write-down control signal We is at the low level, it provides the signal at the low level which is substantially equal to earth potential irrespective of the data signal supplied to the common input/output terminal I/O.

The unit write-down circuit W_{01} includes an output MISFET Q_{52} whose drain is connected to the terminal V_{pp} and whose source is connected to the common bit line CB_0 , a level shifting MISFET Q_{51} whose gate and drain are connected to the gate of the output MISFET Q_{52} and whose source is connected to a zener diode ZD , and a MISFET Q_{50} which serves to supply a bias current to the level shifting MISFET Q_{51} .

The unit write-down circuit W_{01} further includes a MISFET Q_{74} for controlling "on" and "off" the output MISFET Q_{52} , MISFETs Q_{72} and Q_{73} constituting a logic gate circuit for controlling "on" and "off" the MISFET Q_{74} , and a load MISFET Q_{71} .

The gate electrode of the MISFET Q_{72} is supplied with the selection signal K_0 delivered from the Y-DEC₂ circuit, whilst the gate electrode of the MISFET Q_{73} is supplied with a signal delivered from the

input circuit W_{05} .

The output MISFET Q_{52} needs to have a drain breakdown voltage with a sufficient margin with respect to a high voltage such as +25 volts supplied to the terminal V_{pp} .

In general, the drain breakdown voltage of a MISFET is limited comparatively greatly in proportion as the width of a depletion layer in the drain junction surface is limited by a field strength applied from the gate electrode as is well known.

The output MISFET Q_{52} is therefore preferably in the form of a stacked gate structure. The first gate electrode thereof is connected to the gate electrode and drain electrode of the MISFET Q_{51} , and the second gate electrode close to the drain region thereof is connected to the terminal V_{pp} .

If necessary, the output MISFET Q_{52} is in the form of the so-called ring gate structure in which the drain region is surrounded with the gate electrode.

If necessary, a channel region underlying the first gate electrode of the output MISFET Q_{52} is implemented in the depletion mode with the view to appropriately increasing the write-down voltage level to be supplied to the common bit line. A channel region underlying the second gate electrode is implemented in the enhancement mode.

The channel region of the depletion mode can be formed by the known impurity ion implantation technology. In this case, the quantity of an impurity to be implanted into the channel region can be

determined at a much higher accuracy than in the case of employing the impurity diffusion technology.

Accordingly, a threshold voltage which is determined by the first gate electrode of the output MISFET Q_{52} is set at a value which is shifted an accurate predetermined value from the threshold voltages of the MISFETs Q_{51} etc.

The biasing MISFET Q_{50} is also formed in the stacked gate structure in similar manner to the MISFET Q_{52} .

Each of the MISFETs Q_{52} and Q_{50} can be replaced with two MISFETs connected in series. However, in the case where the MISFETs Q_{52} and Q_{50} are implemented in the stacked gate structure as described above, they can be made comparatively small in size.

The breakdown voltage of the zener diode ZD is set at a value substantially equal to the drain junction breakdown voltages of the non-volatile memory elements and is allowed to have a dispersion substantially equal to that of the drain junction breakdown voltages of the non-volatile memory elements.

To this end, the well-known feature in the integrated circuit producing technology that a plurality of elements simultaneously formed exhibit comparatively small relative dispersions in characteristics is exploited. As will be better understood from later description, semiconductor regions for forming the zener diode ZD are formed simultaneously with semiconductor regions for forming the non-volatile memory elements.

The bias current of the zener diode ZD is supplied from the terminal V_{pp} through the MISFETs Q_{50} and Q_{51} .

5 The level shifting MISFET Q_{51} is formed so that a voltage substantially equal to its threshold voltage is provided between its drain and source.

Now, letting A denote the size ratio W/L (W: channel width, L: channel length) of the MISFET Q_{50} , and B denote the size ratio W/L of the MISFET Q_{51} , $B \ll A$ is held. Thus, the level shift amount V_1 is evaluated from the following equation where V_{thD} denotes the threshold voltage of the MISFET Q_{50} and V_{thE} denotes the threshold voltage of the MISFET Q_{51} :

$$V_1 = \sqrt{A/B} V_{thD} + V_{thE} \approx V_{thE} \dots (1)$$

15 The unit write-down circuit W_{01} , the input circuit W_{05} and the associated circuits as shown in Figure 2(B) operate as follows.

During the write-down operation of information, the terminal V_{pp} is supplied with a high voltage V_{pp} such as +25 volts, and the gate electrodes of the MISFETs Q_{64} , Q_{67} and Q_{68} of the input circuit W_{05} are supplied with the write-down control signal We at a level which is substantially equal to the power source voltage V_{CC} .

25 If, at this time, the selection signal K_0 delivered from the Y-DEC₂ circuit is at the selection level or at the high level substantially equal to the voltage V_{CC} , the MISFET Q_{72} will be thereby brought into the "on" state.

30 If the data signal supplied to the common

input/output terminal I/O from a circuit (not shown) is at the low level, the drain of the MISFET Q_{70} of the input circuit W_{05} will provide the high level signal, by which the MISFET Q_{73} will be brought into the "on" state.

5 Since the MISFETs Q_{72} and Q_{73} which are connected in series are concurrently brought into the "on" states, a voltage substantially equal to earth potential GND of the circuitry is applied to the gate electrode of the MISFET Q_{74} . As a result, the MISFET Q_{74} is brought into the "off" state.

10 The turn-off of the MISFET Q_{74} results in supplying the gate electrode of the MISFET Q_{52} with a voltage which is substantially equal to the sum $V_Z + V_{thE}$ between the zener voltage V_Z of the zener diode ZD and the threshold voltage V_{thE} of the MISFET Q_{51} .

15 The output MISFET Q_{52} whose threshold voltage is denoted by V_{thD} maintains its "on" state until the potential of its source electrode is increased to $(V_Z + V_{thE}) - V_{thD}$, that is, until it is increased to a potential higher than the zener voltage V_Z by the threshold voltage difference $V_{thE} - V_{thD}$.

20 Accordingly, the write-down voltage of a value approximately equal to the zener voltage V_Z is supplied from the source electrode of the output MISFET Q_{52} to the common bit line CB_0 .

25 The write-down voltage supplied to the common bit line CB_0 is fed to one of the bit lines B_1 to B_8 through the column switch circuit CS_0 .

30 In the case where the memory element to be

selected is the element F_{11} , the write-down voltage is supplied to the bit line B_1 . In this case, the word line driving signal W_1 is brought to the selection level substantially equal to the voltage V_{pp} by the X-DEC circuit.

5 Owing to the selection level of the word line driving signal W_1 , the memory element F_{11} and the switching MISFET S_1 which is arranged between the source electrode of this memory element and earth are brought into the "on" states. As a result, the write-down
10 current is supplied to the memory element F_{11} .

 Thus, the threshold voltage of the memory element F_{11} is changed from its value which is lower than the selection level (for example, the level of 5 volts which is substantially equal to the voltage V_{CC})
15 of the word line driving signal W_1 delivered from the X-DEC circuit during the read-out operation of information, into its value which is higher than the selection level during the read-out.

 Voltage drops which develop in such a manner
20 that the write-down current flows through the ON resistance between the source and drain of the MISFET, e.g., the MISFET Q_{11} in the column switch circuit CS_0 , a resistance existing on the common bit line, the source output resistance of the output MISFET Q_{52} ,
25 etc., are substantially cancelled by the threshold voltage difference $V_{thE} - V_{thD}$ between the MISFETs Q_{51} and Q_{52} .

 In this case, by forming the MISFETs Q_{51} and Q_{52} on an identical semiconductor substrate
30 substantially simultaneously, the threshold voltages

of the respective MISFETs Q_{51} and Q_{52} undergo dispersions and fluctuations in values substantially equal to each other. Accordingly, the threshold voltage difference $V_{thE} - V_{thD}$ is not substantially affected by the absolute values of the dispersions and fluctuations of the threshold voltages of the respective MISFETs. Since the amount of change of the threshold voltage of the output MISFET Q_{52} is determined by the impurity quantity introduced into the channel region by the ion implantation technology as stated before, the threshold voltage difference $V_{thE} - V_{thD}$ can be determined comparatively accurately.

Accordingly, the voltage drops developing across the resistance components can be satisfactorily cancelled irrespective of the dispersions of the threshold voltages of the MISFETs.

Therefore, the bit lines can be supplied with the write-down voltage in a value being very close to the zener voltage V_Z of the zener diode ZD irrespective of the characteristic dispersions and characteristic fluctuations of the integrated circuit. In other words, the bit lines can be supplied with the write-down voltage of a value which is approximately equal to the drain junction breakdown voltage of the non-volatile memory elements.

The write-down voltage to be supplied to the bit line is sufficiently raised, with the result that a sufficient quantity of information charges are injected into the floating gate of the memory element.

The output MISFET Q_{52} shown in Figure 2(B)

can be altered into the enhancement mode. Even in this case, the write-down voltage at a comparatively good level can be supplied to the bit line owing to the zener voltage V_Z delivered from the zener diode ZD.

5 In this case, however, note must be taken of the fact that the write-down voltage is somewhat lowered in comparison with that in the case where the output MISFET Q_{52} is implemented in the depletion mode, so that the information charges to be injected into the floating gate of the memory element decrease.

10

In the circuit of Figure 2(A), in the case where the bit line B_1 has been supplied with the write-down voltage as stated above, the potentials of the floating gates of the memory elements F_{21} to F_{2561} are raised to levels which are not negligible through the parasitic capacitances between the respective floating gates and the respective drain electrodes because the write-down voltage has the comparatively large value.

15 As a result, if the switching MISFETs S_1 to S_{256} are omitted, the non-selected memory elements F_{21} to F_{2561} will be substantially brought into the "on" states in spite of the non-selection level of the word line driving signals W_2 to W_{256} . Between the source and drain of the switching MISFET S_1 in the "on" state, the write-down current is supplied to the selected memory element F_{11} , and hence, a voltage appears. The memory element F_{21} has its source electrode supplied with the drain voltage of the switching MISFET S_1 , so that it gives rise to a comparatively small leakage current or that it is substantially maintained in the "off" state.

20

25

30

In this embodiment, the switching MISFETs S_3 to S_{256} arranged between the source electrodes of the memory elements F_{31} to F_{256} and earth are held in the "off" states because the word line driving signals W_3 to W_{256} are at the non-selection level.

Accordingly, notwithstanding that the memory elements F_{31} to F_{256} are substantially brought into the "on" states as described above, the channel leakage current can be substantially perfectly prevented from flowing between the bit line B_1 and earth.

In the case where the channel leakage currents are supplied to the respective memory elements F_{31} to F_{256} due to the omission of the switching MISFETs S_1 to S_{256} , a substantial leakage current flows even when the respective channel currents are small. In contrast, the leakage current can be reduced to a substantially negligible magnitude with the use of the above described circuit.

Since the leakage currents which flow through the memory arrays are reduced to the substantially negligible extent, the magnitude of the current driving the write-down circuit can be reduced. In other words, the sizes of the output MISFET Q_{52} etc. can be made small. In addition, the write-down current to be provided from the write-down circuit can be set without considering the dispersion and fluctuation of the leakage current, and the design of the circuitry as well as the elements can be facilitated. Accordingly, an enhancement in the production yield can be achieved.

Instead of arranging the reference potential lines corresponding to the word lines as in Figure 2(A), the sources of a plurality of memory elements belonging to one bit line may be connected to earth through a common resistance. In this case, a voltage is formed across the common resistance by a write-down current to be supplied to one selected memory element. The voltage developing across the common resistance is supplied to the sources of the non-selected memory elements. The non-selected memory elements have their source potentials raised, and are thereby brought into the "off" states. Even with this measure, the occurrence of leakage currents can be prevented as disclosed in German Offenlegungsschrift No. 30.37315.

However, in this case, the following fact must be noted. In order to prevent the leakage currents in the non-selected memory elements, the resistance value of the common resistor needs to be made comparatively great. A comparatively large voltage drop is developed across the common resistor of the high resistance by the write-down current. This voltage is also supplied to the source of the memory element to be selected. Accordingly, the write-down current of the memory element to be selected is reduced, and the write-down efficiency of information is decreased.

The switching MISFETs as shown in Figure 2(A) can be formed in a structure which has a sufficiently low ON-resistance value. Accordingly, the efficiency of the information write-down into the memory to be selected is not substantially limited by the switching

MISFETs.

In the circuit shown in Figure 2(B), in the case where the common input/output terminal I/O is supplied with the data signal at the high level, the output signal of the input circuit W_{05} (the drain output of the MISFET Q_{70}) is brought to the low level substantially equal to earth potential, and the MISFET Q_{73} is brought into the "off" state.

In accordance with the "off" state of the MISFET Q_{73} , the MISFET Q_{74} is brought into the "on" state.

The gate electrode of the MISFET Q_{52} is substantially brought to earth potential by the MISFET Q_{73} .

In this case, the common bit line CB_0 is supplied with only a voltage of a comparatively small value through the MISFET Q_{52} .

Accordingly, no charge is injected into the floating gate of the selected memory element F_{11} .

In the case where the selection signal K_0 provided from the Y-DEC₂ circuit is at the non-selection level, and in the case where the write-down control signal W_e is at the low level, at least one of the MISFETs Q_{72} and Q_{73} responds thereto to be rendered into the "off" state, so that the MISFET Q_{74} is similarly rendered into the "on" state.

The unit write-down circuits W_{02} to W_{04} have the same arrangement as that of the unit write-down circuit W_{01} as shown in the figure.

It will be noted that although the zener diode ZD is common to the unit write-down circuits W_{01}

to W_{04} , in an alternative form separate zener diodes may be provided for each write-down circuit.

According to the arrangement shown in Figures 2(A) and 2(B), the bias current of the zener diode ZD is supplied through one of the unit write-down circuits W_{01} to W_{04} . The bias current is supplied through the MISFETs Q_{50} and Q_{51} when the MISFET Q_{74} is brought into the "off" state. In this case, the drain electrodes of the MISFETs Q_{54} , Q_{57} and Q_{60} are respectively brought to earth potential by MISFETs Q_{78} , Q_{82} and Q_{86} and function substantially as the source electrodes. Accordingly, the MISFETs Q_{54} , Q_{57} and Q_{60} are brought into the "off" states. Since the MISFETs Q_{54} , Q_{57} and Q_{60} are brought into the "off" states, the normal bias current is supplied to the zener diode ZD.

In this embodiment, the read-out circuit R_0 consists of sense amplifiers SA_0 to SA_3 which correspond to the common bit lines CB_0 to CB_3 one to one, and an output buffer circuit OUT.

The sense amplifiers SA_0 to SA_3 are selectively operated by the selection signals K_0 to K_3 which are provided from the Y-DEC₂ circuit.

The output buffer circuit OUT supplies the common input/output terminal I/O with a signal of a level corresponding to an output signal of one selected sense amplifier.

The output terminal of the output buffer circuit OUT is brought into the floating state in such a way that the control signal $\overline{cs}$ is at the low level as during the chip non-selection.

Figure 3 shows a practicable circuit arrangement of the sense amplifier SA_0 and circuits associated therewith. The common bit line CB_0 is connected to the source electrode of an amplifying MISFET Q_{99} and the drain electrode of a constant current MISFET Q_{101} . The gate of the MISFET Q_{99} is connected to the junction between the MISFETs Q_{96} and Q_{97} which are connected in series between the power source terminal and earth so as to form a predetermined bias voltage.

The common bit line CB_0 is also connected to the source electrode of a MISFET Q_{98} whose gate is connected to the junction between the MISFETs Q_{96} and Q_{97} and whose drain is connected to the power source terminal.

Accordingly, during the read-out operation of information, the bias voltage is applied through the MISFET Q_{98} in the sense circuit SA_0 to one memory element selected by the X-DEC and Y-DEC₁ circuits.

In accordance with an information written down previously, the selected memory element has the threshold voltage higher or lower than the selection level of the word line driving signal as stated before.

In the case where the selected memory element is held in the "off" state irrespective of the selection level of the word line driving signal, the common bit line CB_0 is brought to a comparatively high level by the MISFET Q_{98} .

In the case where the selected memory element is held in the "on" state by the selection level of the

47

word line driving signal, the common bit line CB_0 is brought to a comparatively low level.

In this case, the high level of the common bit line CB_0 is brought to a comparatively low level because the gate voltage of the MISFET Q_{98} is held comparatively low.

The low level of the common bit line CB_0 is made to be at a comparatively high level by properly setting the size ratios of the MISFETs Q_{98} and Q_{99} and the memory element.

When the high level and low level of the common bit line CB_0 are limited in this manner, the speed of read-out can be rendered high in spite of the fact that capacitances such as stray capacitances limiting the signal changing speed exist in the common bit line CB_0 . In other words, in such a case where data from a plurality of memory elements are read out in succession, the period of time in which one level of the common bit line CB_0 is changed to the other level can be made short.

The amplifying MISFET Q_{99} has the fixed bias voltage applied to its gate, and the read-out signal from the common bit line CB_0 applied to its source. Therefore, it constitutes a common-gate amplifier together with a load MISFET Q_{100} . An amplified output signal provided at the drain of the MISFET Q_{99} is applied to the gate of one differential MISFET Q_{90} .

The gate of the other differential MISFET Q_{91} is supplied with a reference voltage V_{ref} which is

formed by a reference voltage generating circuit at the succeeding stage.

The reference voltage generating circuit includes a dummy memory element F. The gate of the dummy memory element F is connected to the junction between MISFETs Q_{104} and Q_{105} which are connected in series between the power source terminal and earth so as to form a predetermined intermediate voltage.

Connected to the drain electrode of the dummy memory element F is a MISFET Q_{106} for a dummy which corresponds to the MISFETs of the column switch circuit.

The same circuit as that formed by the MISFETs Q_{96} to Q_{101} is formed by MISFETs Q_{102} , Q_{103} and Q_{107} to Q_{110} .

As shown, the dummy memory element F and the MISFET Q_{106} are connected to the source electrodes of the MISFETs Q_{107} and Q_{108} and the drain of the constant current MISFET Q_{110} .

A potential of an intermediate value between the high level and low level of the common bit line CB_0 is applied to the source electrode of the MISFET Q_{108} by the dummy memory element F and the MISFET Q_{106} .

As a result, the reference voltage V_{ref} at the intermediate level between the read-out high level and low level which are delivered to the drain of the MISFET Q_{99} is provided at the drain of the amplifying MISFET Q_{108} .

Since the reference voltage generating circuit of this form has an arrangement similar to the circuit formed by the MISFETs Q_{96} to Q_{101} , the column

switch circuit and the memory elements, it provides the reference voltage V_{ref} of good level irrespective of the fluctuations of the threshold voltages of the MISFETs.

5 The common source of the differential MISFETs Q_{90} and Q_{91} is provided with a MISFET Q_{92} which functions as a constant current element, and a MISFET Q_{93} whose gate electrode is supplied with the selection signal K_0 from the $Y-DEC_2$ circuit in order to select the sense amplifier.

10 On the other hand, the drains of the differential MISFETs Q_{90} and Q_{91} are respectively provided with load MISFETs Q_{94} and Q_{95} . The amplified output signal of this differential amplifier circuit is transmitted to the output buffer circuit OUT through a latch circuit LC.

5 The latch circuit LC has the selection signal K_0 applied thereto, and is selected together with the sense amplifier SA_0 .

20 The latch circuit LC is arranged so that when selected by the signal K_0 , it provides its output terminal with a signal corresponding to the signal delivered from the differential amplifier circuit. When not selected by the signal K_0 , the output terminal of the latch circuit LC is in the floating state.

25 Accordingly, the input terminal of the output buffer circuit OUT is supplied with only the output signal of that one of the sense amplifiers SA_0 to SA_3 which has been selected by the selection signals K_0 to K_3 .

 Figures 4 and 5 each show different forms of the write-down circuit shown in Figure 2(B).

30 In each of the circuits shown in Figures 4 and 5, an output circuit is formed by MISFETs Q_{111} to

Q₁₁₄. In the preceding stage of the output circuit, there is arranged a suitable logic gate circuit which is supplied with the write-down control signal We, the write-down data and the selection signal. Signals at a level substantially equal to V_{pp} are provided at the output terminals N₁ and N₂ of the output circuit during the write-down operation if the data supplied to the common input/output terminal is "0" or at the low level. A MISFET Q₁₁₆ is arranged between the output terminal N₂ of the output circuit and the corresponding common bit line CB₀. A zener diode ZD similar to that in the foregoing embodiment is connected to the output terminal N₁ through a current clamping MISFET Q₁₁₅. The gate of the MISFET Q₁₁₆ is connected to the zener diode ZD directly or through a level shifting MISFET Q_{115'}. Accordingly, the voltage of the common bit line CB₀ is determined by the MISFET Q₁₁₆.

In the circuit shown in Figure 4, a zener voltage is directly applied to the gate of the clamping MISFET Q₁₁₆. In response thereto, the voltage of the common bit line CB₀ is lowered by an amount equal to the threshold voltage of the MISFET Q₁₁₆. Accordingly, the avalanche breakdown does not occur in the non-volatile memory element F₁₁. Accordingly, the voltage which is applied to the drain of the memory element F₁₁ during the write-down becomes comparatively low, so that the circuit does not operate properly in the case where a comparatively high write-down speed is required.

In the circuit shown in Figure 5, the level shifting MISFET Q_{115'} is arranged as in the embodiment shown in Figure 2(B).

In accordance with the circuit of Figure 5, therefore, the voltage which is applied to the gate of the MISFET Q_{116} is raised approximately to the threshold voltage V_{thE} of the MISFET Q_{115} with respect to the voltage of the zener diode ZD as in the circuit of the embodiment in Figure 2(B). As a result, the potential of the common bit line can be set substantially at the zener voltage as in the circuit of the embodiment shown in Figure 2(B).

Thus, the write-down voltage to be applied to the bit line can be accurately set to a magnitude which is necessary for preventing the avalanche breakdown in the non-volatile memory element and which is close to the maximum voltage. Accordingly, the non-volatile memory element is permitted to perform a high-speed write-down operation.

As a constant voltage element for determining the clamp voltage, it is convenient to utilize the semiconductor regions which are formed simultaneously with the formation of the semiconductor regions of the non-volatile memory element as stated before.

In order to attain a high-speed write-down operation, the non-volatile memory elements F_{11} are fabricated in a structure as shown in Figures 6(A) to 6(C).

Referring to Figures 6(A) to 6(C), a semiconductor substrate 1 is made of a p-type Si (silicon) single crystal in which a conductivity type determining impurity is contained at, for example, 10^5 atoms/cm³. One major surface of the substrate 1 is formed with a

thick field SiO_2 (silicon oxide) film 2. A portion surrounded with the field SiO_2 film 2 is made an active region. The substrate surface which constitutes the active region is formed with a thin gate SiO_2 film 3. The gate SiO_2 film 3 is overlaid with a floating gate FG which is made of a first conductive polycrystalline Si layer. The floating gate electrode FG is overlaid through an intermediate insulating film 5a with a control gate electrode CG which is made of a second conductive polycrystalline Si layer substantially forming a word line.

The floating gate FG has a width equal to that of the control gate CG and extends onto the field SiO_2 film 2. In Figure 6(A), the end parts of the floating gate FG are indicated by broken lines E_{d1} and E_{d2} .

In that surface of the semiconductor substrate 1 which is not surrounded with the field SiO_2 film 2 and the gate electrodes, n^+ -type semiconductor regions R_1 , 4_{S1} , 4_D , 4_{S2} and R_2 formed by introducing an n-type impurity at a comparatively high concentration are formed.

The n^+ -type semiconductor region 4_{S1} is used as the source region of one non-volatile memory element. The n^+ -type semiconductor region R_1 which is extended in the lateral direction on the drawing as shown in Figure 6(A) is used as a region for forming a reference potential line.

The source region 4_{S1} and the wiring region R_1 are integrally formed as shown in the drawing.

The n^+ -type semiconductor region 4_{S2} is

similarly used as the source region of another non-volatile memory element, and the n^+ -type semiconductor region R_2 is used as a wiring region.

5 The n^+ -type semiconductor region 4_D is used as the common drain regions of the pair of non-volatile memory elements.

10 Over the control gate electrode CG, a wiring layer 6 which forms a bit line (made of, for example, an evaporated aluminium layer) is formed through an intermediate insulating film $5b$ so as to intersect orthogonally with the control gate electrode.

The wiring layer 6 is held in ohmic contact with the drain region 4_D in a contact portion CNT.

15 As illustrated in Figure 6(C), in the surface of the semiconductor substrate 1 underneath the field SiO_2 film 2, a p^+ -type region 7 which has the same conductivity type as that of the substrate and an impurity concentration higher than that of the substrate is formed as a channel stopper. Further, if necessary,
20 a p^{++} -type region $7a$ having a still higher impurity concentration is formed in that part of the channel portion of the semiconductor non-volatile memory element which is close to the field SiO_2 .

25 In the non-volatile memory element constructed as above when the write-down voltage is applied to the drain region 4_D and the word line driving signal of the selection level is applied to the control gate CG_1 , the spread of a depletion layer to spread to the p^{++} -type region $7a$ from the junction
30 portion between the drain region 4_D and the p^{++} -type

region 7a is limited, and the drain field strength is intensified accordingly. Electrons which form a channel current are greatly accelerated by the high drain field strength. Accordingly, high energy is afforded to electron-hole pairs which are created when the electrons have collided against semiconductor crystal lattices. In other words, the generation efficiency of hot electrons is improved, and the write-down speed is greatly enhanced.

In the non-volatile memory element, in the case where electrons are injected into the floating gate FG through the gate SiO_2 film 3 as is well known, this floating gate FG becomes a negative potential, so that a threshold voltage for the control gate CG is made a comparatively great value such as about 7 volts. In contrast, in the case where electrons are not injected into the floating gate FG, a threshold voltage for a voltage to be applied to the control gate CG is set at a comparatively low value such as about 2 volts.

Accordingly, when the selection level of the word line driving signal, i.e., the level of the signal to be applied to the control gate CG is set at an intermediate value between the high and low threshold voltages of the memory element, e.g., 5 volts in the read-out operation, the memory element is rendered into the "on" state or "off" state, depending upon whether or not the electrons are injected into the floating gate FG in advance. Therefore, the read-out can be carried out.

Although this is not especially restrictive,

the non-volatile memory element as stated above can be fabricated in accordance with a method of manufacture to be described below. Figures 7(A) to 7(F) show sectional views of a semiconductor substrate in
5
respective manufacturing steps. Accordingly, the method of manufacturing an embodiment will now be described with reference to Figures 7(A) to 7(F).

The left half parts I of Figures 7(A) to 7(F) show a non-volatile memory element, whilst the right
10
half parts II show a MISFET for constituting a peripheral circuit.

(a) First, a semiconductor substrate 1 made of single-crystal p-type silicon whose major surface is parallel to the (100) plane, which is about 200 μ m thick and which has a resistivity of about 20 ohm-cm is prepared.

A thin thermally-oxidized SiO₂ film 3' having a thickness of, e.g. several hundred Å is formed on the surface of the substrate 1 by the known thermal
20
oxidation technology in order to prevent crystal defects from being induced in the surface of the substrate 1 due to a silicon nitride film to be formed later.

Subsequently, the silicon nitride (Si₃N₄) film
25
8 which is about 1400 Å thick is formed on the SiO₂ film 3' by the known chemical vapour deposition technology which uses silane gas and ammonia gas.

A photoresist film (not shown) for an etching mask is formed on the Si₃N₄ film 8, whereupon
30
the photoresist film is selectively exposed to light

and is developed. As a result, the photoresist film is left on parts of the substrate 1 to be made active regions.

5 Subsequently, using the remaining photoresist film as the etching mask, the Si_3N_4 film 8 is etched and removed. For the etching of the Si_3N_4 film 8, the plasma etching process which uses freon gas as a reaction gas can be employed.

10 After removing the photoresist film, a photoresist film 9 is formed on the substrate 1 anew.

The photoresist film 9 is selectively exposed to light and is developed in order that a window for forming the p^{++} -type region 7a as shown in Figures 6(A) to 6(C) may be formed by the combination between the Si_3N_4 film 8 and the photoresist film 9.

15 Using the remaining photoresist film 9 and Si_3N_4 film 8 as an ion implantation mask, a p-type impurity is introduced into the surface of the substrate 1 through the SiO_2 film 3' by the ion implantation technology as illustrated in Figure 7(A).

20 In this case, BF_2^+ (boron fluoride) is used as impurity ions though no special restriction is intended. The impurity ions are given energy of, e.g. 75 keV and are implanted into the surface of the substrate 1 at a concentration of, e.g., about 10^{14} atoms/cm².

25 Since BF_2 as the impurity ions has a mass greater than that of the impurity such as B (boron) (mass of BF_2 : 49, mass of B: 11), it is difficult to penetrate through the Si_3N_4 film 8 and the underlying SiO_2 film 3'. Accordingly, even when the Si_3N_4 film 8

57

is made comparatively thin as described above, the impurity is not introduced into the surface of the substrate 1 under the Si_3N_4 film 8.

(b) Subsequently, after removing the photoresist film 9, the substrate 1 is annealed at a temperature of, e.g., 1100°C in an inert gas atmosphere such as nitrogen gas. On the basis of the annealing treatment, the impurity introduced in the surface of the substrate 1 is diffused by the drive-in diffusion. As a result, the p^{++} -type region 7a which is extended into the substrate surface under the Si_3N_4 film 8 is formed. After the annealing treatment, using the Si_3N_4 film 8 as a mask, BF_2 ions at the second time are implanted into the whole surface of the substrate 1 under conditions establishing, for example, 75 keV and 2×10^{13} atoms/cm². The impurity for forming the p^+ -type layer 7 as the channel stopper is introduced into the surface of the substrate 1 by the second ion implantation.

In an alternative method, the implanting operations of the impurity ions may be carried out as follows.

First, the Si_3N_4 film 8 formed on the SiO_2 film 3' is coated with a photoresist film. The photoresist film on the part in which the p^{++} -type layer 7a is to be formed is removed by exposure and developing. Using the remaining photoresist film as a mask, the Si_3N_4 film is selectively etched and removed.

Using the remaining Si_3N_4 film 8 and the overlying photoresist film as an ion implantation mask, p-type impurity ions at the first time are implanted

into the surface of the substrate 1.

After removing the photoresist film, the substrate 1 is annealed as in the foregoing.

5 The whole surface of the substrate is coated with a photoresist film 9 again. The photoresist film 9 is selectively exposed to light and is developed so as to remain on the part to be made the active region.

10 Using the remaining photoresist film 9 as an etching mask, the Si_3N_4 film 8 is selectively etched and removed.

Using the remaining photoresist film 9 and Si_3N_4 film 8 as an ion implantation mask, p-type impurity ions at the second time are implanted into the surface of the substrate 1.

15 It should be noted that in the alternative method, in the first ion-implanting operation, that surface of the substrate 1 which is not to be implanted with the ions is covered with the stacked Si_3N_4 film 8 and photoresist film. In the second ion-implanting operation, that surface of the substrate 1 which is not to be implanted with the ions, especially the surface which is to be finally made the channel region of the field-effect element, is similarly covered with the stacked Si_3N_4 film 8 and photoresist film 9.

25 As a result, it can be satisfactorily avoided that the impurity ions are undesirably introduced into the surface part of the substrate 1 to be finally made the channel region of the field-effect element.

30 (c) After the second operation of implanting the impurity ions, the substrate 1 is

heated at about 1000°C in, for example, an oxygen atmosphere containing steam. Thus, a comparatively thick field oxide film 2 having a thickness of, e.g. about 1μm is formed on the surface of the substrate 1. Since, in this case, the Si₃N₄ film 8 acts as an oxidation mask, the surface of the substrate 1 covered with the Si₃N₄ film 8 is not formed with the field oxide film 2 as seen from Figure 7(C). In this heat treatment for oxidation, the p-type impurity having been introduced in the substrate surface by the ion implantation in advance is diffused into the substrate. Accordingly, the p⁺⁺-type region 7a is stretched deep under the field oxide film 2 close to the memory element forming region. A part of the p⁺⁺-type region 7a is extended to the active region (under the Si₃N₄ film 8). At the same time, the p⁺-type region 7 at the concentration as the channel stopper is formed under the field oxide film 2. Since the p-type impurity which is introduced by the second operation of implanting the impurity ions is at the comparatively low concentration as described above, the p⁺-type region 7 is hardly stretched to the active region.

In accordance with the manufacturing method as stated above, the self-alignment between the field SiO₂ film 2 and the p⁺⁺-type region 7a is possible, and very accurate and small-sized memory cells can be manufactured at a good yield. In addition, the memory portion and the Si₃N₄ film in the peripheral part can be worked at one stroke, so that the deviation of mask alignment need not be considered.

(d) After the field oxidation, the Si_3N_4 film 8 is etched and removed by the known etching process which uses heated phosphoric acid.

5 If necessary, the SiO_2 film 3' is removed by a method such as wash-out method, whereupon the exposed surface of the substrate 1 is formed with a thin SiO_2 film (not shown) having a thickness of, e.g. about 400 Å by the thermal oxidation method.

10 The whole surface of the substrate 1 is coated with a photoresist film again. The photoresist film is selectively exposed to light and is developed so that the part to be made the channel region of the field effect transistor later may be exposed.

15 Using the photoresist film and the field oxide film 2 as an ion implantation mask, As (arsenic) ions as an n-type impurity are implanted into the surface of the substrate 1 under conditions establishing, for example, 150 keV and 3×10^{12} atoms/cm².

20 After removing the photoresist film, the thin SiO_2 film is removed.

The substrate 1 is heated at, for example, 1000°C in an oxygen atmosphere, thereby to form a gate SiO_2 film 3 being about 500 Å thick on the exposed surface of the substrate 1.

25 In such a case where insulated gate field-effect transistors of comparatively high threshold voltages are required in peripheral circuits of, for example, the 25 volt type, a comparatively thick gate SiO_2 film of, e.g. 800 Å is formed by the thermal oxidation before the formation of the gate SiO_2 film, and it is subsequently removed selectively from the surface part of the substrate 1. Thereafter, the

30

comparatively thin gate SiO_2 film can be formed.

After forming the gate SiO_2 film, boron ions given energy of, e.g. 70 keV can be introduced into the substrate surface through the gate SiO_2 film 3 if it is necessary to control the threshold voltage of an enhancement mode transistor to a desired value by way of example.

After forming the gate SiO_2 film 3, a polycrystalline Si layer 10 having a thickness of about 3500 Å is formed on the surface of the substrate 1 by the CVD (Chemical Vapour Deposition) process in order to form the floating gate electrode of the memory transistor, the gate electrode of the peripheral transistor and the required wiring layer. Subsequently, a photoresist film 9' is formed on the polycrystalline Si layer 10, and the photoresist film 9' is selectively exposed to light and is developed. Using the remaining photoresist film 9' as a mask, the polycrystalline Si layer 10 is selectively etched (patterned) as shown in Figure 7(D), thereby to form the floating gate FG of the memory transistor, the gate electrode G of the peripheral transistor and the necessary wiring.

(e) After removing the photoresist film 9', P (phosphorus) ions are implanted into the polycrystalline Si layer by the ion implantation method. Thus, the polycrystalline Si layer to be made the gate electrodes FG and G and the wiring is turned into the n-type and is given a low resistance characteristic.

Thereafter, the substrate 1 is heated in an oxidizing atmosphere so as to form an SiO_2 film 5a

being about 800 Å thick on the surface of each of the gate electrodes FG and G and the wiring layer as illustrated in Figure 7(E).

5 The SiO₂ films 5a function as an intermediate insulating film.

(f) On the substrate surface subjected to the above step (e), the second layer of polycrystalline Si layer CG being about 3500 Å thick is formed by the CVD process. Subsequently, phosphorus is introduced
10 into the surface of the second layer of polycrystalline Si layer CG as a conductivity type determining impurity by an appropriate impurity deposition method.

Thereafter, a photoresist film is formed on the surface of the substrate 1. The width of the floating gate FG in the memory element is preferably
15 made equal to that of the control gate CG. To this end, the photoresist film is selectively exposed to light and is developed so as to determine the patterns of the control gates and word lines in the memory array
20 portion. Accordingly, except for the part to be made the memory array, the photoresist film is left substantially whole by the exposure and developing.

After the exposure and developing, using the remaining photoresist film as an etching mask, the
25 second layer of polycrystalline Si^{layer} is selectively etched and removed.

Using the remaining polycrystalline Si layer CG as a mask, the SiO₂ films 5a on the surface of the first layer of polycrystalline Si layer FG are etched
30 and removed. Subsequently, using the second layer of

polycrystalline Si layer CG and the SiO_2 films 5a as a mask, the first layer of polycrystalline Si layer FG is etched and removed. By this etching, the first layer of polycrystalline Si layer FG has its width determined by the width of the second layer of polycrystalline Si layer CG.

After removing the photoresist film, a photoresist film is formed on the surface of the substrate 1 anew.

The photoresist film formed anew is selectively exposed to light and it developed in a pattern in which the unnecessary parts of the second layer of polycrystalline Si layer as overlies the gate electrodes of the insulated gate field-effect transistors constituting the peripheral circuits are removed.

The second layer of polycrystalline Si layer is also utilized as the second gate electrodes in stacked gate type transistors constituting the peripheral circuits. Accordingly, the new photoresist film is also left on parts to be made the second gate electrodes of the stacked gate type transistors.

Using the photoresist film as an etching mask, the second layer of polycrystalline Si layer is selectively etched and removed.

Subsequently, the photoresist film is removed. At the stage at which the photoresist film has been removed, only the comparatively thin SiO_2 film is left on the surface parts of the substrate 1 not covered with the field oxide film 2 and the first and

second layers of polycrystalline Si layers, i.e. the surface parts to form the source and drain regions of the transistors and the surface parts to form the semiconductor wiring regions.

5 If necessary, the comparatively thin SiO_2 film is removed by an etching method such as the wash-out method. Subsequently, an SiO_2 film being about $350 \text{ \AA}$ thick is formed on the exposed surface of the substrate 1 by the thermal oxidation method. The SiO_2 film is
10 also formed on the exposed surfaces of the first and second layers of polycrystalline Si layers by thermal oxidation.

 Subsequently, using the field oxide film 2 and the polycrystalline Si layers as an ion
15 implantation mask, phosphorus ions are introduced into the surface of the substrate 1 so as to establish, for example, a depth of $1 \text{ \mu m}$ and a concentration of 10^{16} atoms/ cm^2 .

 As a result, the n-type source and drain
20 regions of the transistors and the semiconductor wiring regions are formed in the surface of the substrate 1.

 At the next step, if necessary, the
substrate 1 is heated at a comparatively low temperature of, e.g. 800°C in, e.g. an oxidation atmosphere,
25 thereby to form the insulating films 5b on the exposed surfaces of the polycrystalline Si layers. The structure so far obtained is shown in Figure 7(F).

 Thereafter, using as a mask a photoresist
film for leading out electrodes, the SiO_2 film on the
30 source and drain regions is selectively etched and

removed.

After removing the photoresist film, a PSG (phosphosilicate glass) film being, e.g. about 8000 Å thick is subsequently formed on the substrate surface.

5 Using a photoresist film as a mask, the PSG film is selectively etched to form contact holes.

After removing the photoresist film, Al (aluminium) is evaporated on the substrate, and the Al layer is patterned to form a wiring layer. The
10 aluminium wiring 6 is shown in, for example, Figure 6(B) and 6(C).

In the non-volatile memory element of the above construction, the P^{++} -type region 7a lies in contact with the drain region 4_D, so that the
15 pn-junction break-down takes place under a comparatively low write-down voltage.

Accordingly, the write-down voltage to be applied to the bit line must be limited to a comparatively low value.

20 When one selected bit line, for example, the bit line B₁ in Figure 2(A) is supplied with a write-down voltage exceeding the drain breakdown voltage of the memory element, the drain junctions of the plurality of memory elements F₁₂ - F₁₂₅₆
25 belonging to the bit line B₁ undergo the avalanche breakdown respectively. In the case where the avalanche breakdown has occurred in this manner, holes generated by the avalanche breakdown are injected into the floating gates of the memory
30 elements, and the re-write-down efficiency is substantially reduced.

The write-

down voltage to be applied to the drain of the memory element is formed by the constant voltage diode which has a breakdown voltage substantially equal to the breakdown voltage of the drain junction of the memory element.

5 The constant voltage diode is formed on a single semiconductor substrate by the integrated circuit producing technology together with the transistors for constituting a plurality of memory elements and peripheral circuits.

10 The constant voltage diode is formed in an appropriate part of the peripheral portion of the semiconductor chip lest the carriers injected into the substrate by the breakdown operation of its junction should adversely affect the memory elements.

15 A constant voltage diode of an appropriate structure which can be formed on the semiconductor substrate together with the memory elements can be obtained, as shown in Figures 8A and 8B.

20 As shown in the figures, the constant voltage diode is constructed of an n^+ -type semiconductor region 4a and a p^{++} -type region 7a. An electrode and wiring K such as an evaporated aluminium layer is held in non-rectifying contact with the n^+ -type semiconductor region 4a as the cathode of the diode through an
25 aperture provided in an SiO_2 film 3. Although not illustrated, the wiring K is extended onto a field oxide film 2 and is connected with the sources of the transistors Q_{51} etc. as shown in Figure 2(B).

30 An electrode BE such as of gold is preferably arranged on the lower surface side of the semiconductor

substrate 1. The anode electrode of the constant voltage diode is constituted by the electrode BE.

5 The semiconductor region 7a is formed simultaneously with the p^{++} -type region 7a in Figures 6(A) to 6(C) and Figure 7, while the semiconductor region 4a is similarly formed simultaneously with the source and drain regions 4_{S1} , 4_{S2} and 4_D of the non-volatile memory elements.

10 In Figures 8(A) and 8(B), an n^+ -type region 4b is formed in a manner to surround the n^+ -type region 4a lest the carriers injected into the substrate 1 by the operation of the constant voltage diode should exert any adverse effect on the non-volatile memory elements.

15 The carriers injected into the substrate 1 have difficulty in spreading outside the n^+ -type region 4b.

20 In the drawing, the n^+ -type region 4b is shown as being maintained at a predetermined potential of the circuit. The n^+ -type region 4b, may preferably be held at the level of the power source voltage V_{CC} of the circuit.

25 The p^{++} - n^+ junction in the non-volatile memory element and the p^{++} - n^+ junction of the zener diode ZD are formed by the same manufacturing step, so that the breakdown voltages of both the junctions equalize following the dispersion of the manufacturing process. It is therefore possible to attain a clamp voltage for accurately preventing the avalanche breakdown in the non-volatile memory element.

30

As shown in Figure 2(B), the write-down circuit is of the type wherein the output MISFET Q_{52} etc. arranged on the high voltage side are controlled by the logic circuit which operates under the low voltage V_{CC} . Therefore, the write-down circuit can also be simplified and as a result has a low power dissipation.

Figure 9(A) shows a pattern diagram of a memory array, whilst Figure 9(B) shows a structural sectional view taken along the line A-A' in Figure 9(A).

Referring to Figure 9(A), a field oxide film 2 in a pattern indicated by the combination of oblique lines ascending rightward and solid lines is formed on the surface of a p-type semiconductor substrate 1.

Word lines W_1 to W_5 made of polycrystalline silicon layers are also extended on the surface of the substrate 1 in the lateral direction of the drawing. Predetermined parts of the word lines W_1 to W_5 form the control gates of semiconductor non-volatile memory elements F_{n1} to F_{n+21} , F_{n2} to F_{n+22} , and F_{n3} to F_{n+23} , respectively.

Floating gates as apparent from Figures 6(A) to 6(C) and Figures 7(A) to 7(F) are arranged under the control gates. In Figure 9(A), the pattern of the floating gate is indicated by the combination of broken lines and oblique lines descending rightward.

In that surface of the substrate 1 which is not covered with the field oxide film 2 and the word lines W_1 to W_5 , n^+ -type semiconductor regions to be used as the source and drain regions of the semiconductor

non-volatile memory elements and semiconductor wiring regions are formed.

5 A semiconductor region R_1 which is held between the word lines W_1 and W_2 and which is extended in parallel with these word lines W_1 and W_2 is used as the common source region of the memory elements F_{n1} to F_{n+21} and F_{n2} to F_{n+22} . The semiconductor region R_1 constitutes the reference potential line R_1 as shown in Figure 2(A).

10 The semiconductor substrate 1 is overlaid with a plurality of conductor layers B_n to B_{n+21} which are made of, e.g. aluminium layers and which are extended in the vertical direction in the drawing. These conductor layers constitute bit lines, respectively.

15 The respective conductor layers are connected with the drain regions of the memory elements through contact holes.

By way of example, the conductor layer B_n is connected with the drain regions of the memory elements F_{n1} , F_{n2} , F_{n3} , F_{n4} , F_{n5} etc. through the contact holes CH_{n1} , CH_{n2} and CH_{n3} .

20 In the memory array of Figure 9(A), a plurality of switching MISFETs S_1 to S_5 are formed together with the plurality of memory elements.

25 By way of example, the switching MISFET S_1 is formed from a source region 10_{S1} , the word line W_1 and the semiconductor region R_1 , and the switching MISFET S_2 is formed from the semiconductor region R_1 , the word line W_2 and a source region 10_{S2} . Further, 30 the switching element S_3 is formed from the source

region 10_{S2}, the word line W_3 and a semiconductor region R_2 .

5 The switching MISFETs, for example, S_2 and S_3 belonging to one set of word lines W_2 and W_3 have their source regions made common.

10 The drains of the switching MISFETs are formed from the semiconductor regions R_1 , R_2 , R_3 etc. The sources of the switching MISFETs are connected to an earth line GND through contact holes. The earth line GND can be formed in parallel with the bit lines, and can be made of the aluminium wiring layers similar to the bit lines. Accordingly, the gates of the switching MISFETs can be formed by the patterning step of the word lines, and the earth line for the switching MISFETs by the patterning step of the bit lines.

15 The semiconductor regions constituting the switching MISFETs can be formed without affecting the regularity of the pattern of the memory array. Therefore, this embodiment has the advantage that the switching MISFETs can be inserted in any desired parts of the memory array portion. Moreover, the density of integration is not affected.

20 These semiconductor devices can be formed by the semiconductor manufacturing method described before.

25 The reason why such switching MISFETs are arranged in the memory array portion is to prevent the channel leakage currents as stated previously. Therefore, the switching MISFETs of this embodiment are not provided with regions such as p^{++} -type regions

30

for high speed write-down which are arranged in the substrate surface in the channel direction of the memory elements. The switching MISFETs preventive of the channel leakage currents as stated above
5 can also be extensively applied to EPROMs which employ memory elements provided with no p^{++} -type region.

Attention is drawn to our Application No. 82.08623 in which there is claimed in claim 1,
10 an electrically programmable read only memory including:

- (a) at least one non-volatile memory element which has a floating gate electrode and a control gate electrode; and
- 15 (b) a pn-junction element having a junction which is similar to the drain junction of the non-volatile memory element, the level of a write-down voltage to be applied to the drain of the non-volatile memory element being determined
20 on the basis of the breakdown voltage of the pn-junction element.



CLAIMS.

1. An electrically programable read only memory including:
 - (a) a first word line;
 - (b) a plurality of non-volatile memory elements, each having source, drain and control electrodes, the control electrodes being connected in common to the first word line;
 - (c) a first switch MISFET having source, drain and gate electrodes; and
 - (d) a source of reference potential;
at least one of the source electrodes of the plurality of non-volatile memory elements being connected to the source of reference potential through the drain-source path of the first switch MISFET, the gate electrode of the first switch MISFET being connected to the first word line.
2. An electrically programable read only memory according to Claim 1, wherein said source electrodes of the plurality of non-volatile memory elements are connected in common to the source of reference potential through the drain-source path of the first switch MISFET.
3. An electrically programable read only memory according to Claim 2, further including:
 - (e) a second word line;
 - (f) a second plurality of non-volatile memory elements, the control gates of which are connected in common to the second word line; and
 - (g) a second switch MISFET; wherein the



source electrodes of the second plurality of non-volatile memory elements are connected in common to the source of reference potential through the drain-source path of the second switch MISFET, the gate electrode of the second switch MISFET being connected to the second word line, and the source electrodes of the first and the second plurality of non-volatile memory elements and drain electrodes of said first and said second switch MISFET being connected in common.

4. An electrically programable read only memory according to Claim 3, wherein the source of reference potential is earth potential.

5. An electrically programable read only memory according to Claim 2, further including:-

(h) n additional word lines;

(i) n additional plurality of non-volatile memory elements, the control gates of n respective sets of the plurality being connected in common to the respective n additional word lines; and

(j) n additional switch MISFETs; wherein the source electrodes of n respective sets of the non-volatile memory elements are connected in common to the source of reference potential through the respective source-drain paths of the respective n switch MISFETs, the gate electrodes of the respective n switch MISFETs being connected to the respective n word lines.



- 57 -

6. An electrically programable read only memory constructed and arranged to operate substantially as herein described with reference to and as illustrated in the accompanying drawings.
