



(51) International Patent Classification:

H01L 21/8242 (2006.01) *H01L 27/10* (2006.01)
H01L 21/8234 (2006.01) *H01L 27/108* (2006.01)
H01L 27/08 (2006.01) *H01L 29/786* (2006.01)
H01L 27/088 (2006.01)

(21) International Application Number:

PCT/JP2010/070753

(22) International Filing Date:

16 November 2010 (16.11.2010)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

2009-282139 11 December 2009 (11.12.2009) JP

(71) Applicant (for all designated States except US): SEMI-CONDUCTOR ENERGY LABORATORY CO., LTD.
[JP/JP]; 398, Hase, Atsugi-shi, Kanagawa, 2430036 (JP).

(72) Inventors; and

(75) Inventors/Applicants (for US only): KATO, Kiyoshi
[JP/JP]; c/o SEMICONDUCTOR ENERGY LABORATORY CO., LTD., 398, Hase, Atsugi-shi, Kanagawa, 2430036 (JP). KOYAMA, Jun.

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

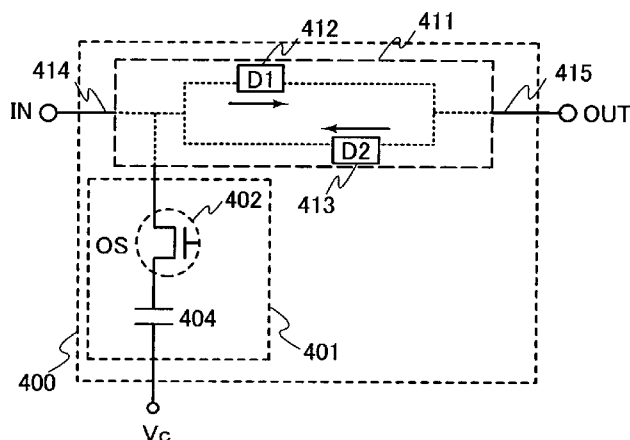
(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— with international search report (Art. 21(3))

(54) Title: NONVOLATILE LATCH CIRCUIT AND LOGIC CIRCUIT, AND SEMICONDUCTOR DEVICE USING THE SAME

FIG. 1



(57) Abstract: To provide a novel nonvolatile latch circuit and a semiconductor device using the nonvolatile latch circuit, a non-volatile latch circuit includes a latch portion having a loop structure where an output of a first element is electrically connected to an input of a second element, and an output of the second element is electrically connected to an input of the first element; and a data holding portion configured to hold data of the latch portion. In the data holding portion, a transistor using an oxide semiconductor as a semiconductor material for forming a channel formation region is used as a switching element. In addition, a capacitor electrically connected to a source electrode or a drain electrode of the transistor is included.

DESCRIPTION

**NONVOLATILE LATCH CIRCUIT AND LOGIC CIRCUIT, AND
SEMICONDUCTOR DEVICE USING THE SAME**

5

TECHNICAL FIELD

[0001]

The invention disclosed herein relates to a nonvolatile logic circuit in which the logical state of storing data is not erased even after power is turned off, and a semiconductor device using the nonvolatile logic circuit. In particular, the invention disclosed herein relates to a nonvolatile latch circuit and a semiconductor device using the nonvolatile latch circuit.

BACKGROUND ART

15

[0002]

An integrated circuit has been proposed in which nonvolatile logic is integrated, where a feature of "nonvolatile" with which data is not erased even when power is turned off is applied to a logic circuit. For example, a nonvolatile latch circuit using a ferroelectric element as nonvolatile logic has been proposed (see Patent Document 1).

20

[Reference]

[Patent Document]

[0003]

[Patent Document 1] PCT International Publication No. 2003/044953

25

DISCLOSURE OF INVENTION

[0004]

However, a nonvolatile latch circuit using a ferroelectric element has problems in terms of reliability of the number of rewrites and reduction in voltage. In addition, a ferroelectric element is polarized by an electric field which is applied to the element, and stores data by remanent polarization. However, when the remanent polarization is small, the following problems might arise: the influence of variation in the amount of

charge becomes large, and a high-accuracy reading circuit is needed.

[0005]

In view of the aforementioned problems, an object of an embodiment of the present invention is to provide a novel nonvolatile latch circuit and a semiconductor device using the nonvolatile latch circuit.

[0006]

A nonvolatile latch circuit according to one embodiment of the present invention includes a latch portion having a loop structure, in which an output of a first element is electrically connected to an input of a second element, and an output of the second element is electrically connected to an input of the first element; and a data holding portion for holding data of the latch portion. In the data holding portion, a transistor using an oxide semiconductor as a semiconductor material for forming a channel formation region is used as a switching element.

[0007]

In addition, the data holding portion includes a capacitor which is electrically connected to a source electrode or a drain electrode of the transistor. With the use of the transistor, data held in the latch portion can be written into the capacitor of the data holding portion. Further, with the use of the transistor, the data written into the capacitor of the data holding portion can be held. Furthermore, with the use of the transistor, the data held in the capacitor of the data holding portion can be read to the latch portion.

[0008]

In other words, a nonvolatile latch circuit according to one embodiment of the present invention includes a latch portion and a data holding portion for holding data of the latch portion. The data holding portion includes a transistor and a capacitor. A channel formation region of the transistor includes an oxide semiconductor layer. One of a source electrode and a drain electrode of the transistor is electrically connected to one of electrodes of the capacitor, and the other of the source electrode and the drain electrode of the transistor is electrically connected to the latch portion.

[0009]

In the aforementioned nonvolatile latch circuit, the latch portion includes a first element and a second element, and has a loop structure in which an output of the first

element is electrically connected to an input of the second element, and an output of the second element is electrically connected to an input of the first element. In addition, the input of the first element is electrically connected to a wiring supplied with an input signal, and the output of the first element is electrically connected to a wiring supplied with an output signal. For example, an inverter can be used for each of the first element and the second element. Alternatively, a NAND can be used for the first element, and a clocked inverter can be used for the second element, for example.

[0010]

In the aforementioned nonvolatile latch circuit, the other of the source electrode and the drain electrode of the transistor is electrically connected to the input of the first element of the latch portion, and the other of the source electrode and the drain electrode of the transistor is electrically connected to the wiring supplied with the input signal.

[0011]

In the aforementioned nonvolatile latch circuit, the transistor has a function of writing data held in the latch portion into the capacitor of the data holding portion. In addition, the transistor has a function of holding the data written into the capacitor of the data holding portion. Moreover, the transistor has a function of reading the data held in the capacitor of the data holding portion to the latch portion.

[0012]

A nonvolatile latch circuit according to another embodiment of the present invention includes a latch portion and a data holding portion for holding data of the latch portion. The data holding portion includes a first transistor, a second transistor, a first capacitor, and a second capacitor. Channel formation regions of the first transistor and the second transistor each include an oxide semiconductor layer. One of a source electrode and a drain electrode of the first transistor is electrically connected to one of electrodes of the first capacitor, and the other of the source electrode and the drain electrode of the first transistor is electrically connected to the latch portion. One of a source electrode and a drain electrode of the second transistor is electrically connected to one of electrodes of the second capacitor, and the other of the source electrode and the drain electrode of the second transistor is electrically connected to the latch portion.

[0013]

In the aforementioned nonvolatile latch circuit, the latch portion includes a first element and a second element, and has a loop structure in which an output of the first element is electrically connected to an input of the second element, and an output of the second element is electrically connected to an input of the first element. In addition, the input of the first element is electrically connected to a wiring supplied with an input signal, and the output of the first element is electrically connected to a wiring supplied with an output signal. For example, an inverter can be used for each of the first element and the second element. Alternatively, a NAND can be used for the first element, and a clocked inverter can be used for the second element, for example.

10 [0014]

In the aforementioned nonvolatile latch circuit, the other of the source electrode and the drain electrode of the first transistor is electrically connected to the input of the first element of the latch portion, and the other of the source electrode and the drain electrode of the first transistor is electrically connected to the wiring supplied with the input signal. The other of the source electrode and the drain electrode of the second transistor is electrically connected to the output of the first element of the latch portion, and the other of the source electrode and the drain electrode of the second transistor is electrically connected to the wiring supplied with the output signal.

[0015]

20 In the aforementioned nonvolatile latch circuit, the first and second transistors each have a function of writing data held in the latch portion into the first and second capacitors of the data holding portion. In addition, the first and second transistors each have a function of holding data written into the first and second capacitors of the data holding portion. Moreover, the first and second transistors each have a function of reading data held in the first and second capacitors of the data holding portion to the latch portion.

[0016]

30 In the aforementioned nonvolatile latch circuit, with a transistor including an oxide semiconductor layer, which is formed with an oxide semiconductor material, for a channel formation region, can obtain the following characteristics even in the case of, for example, an element whose channel width W is $1 \times 10^4 \mu\text{m}$ and channel length L is 3

μm : the off-state current is less than or equal to 1×10^{-13} A at room temperature; and the subthreshold swing (S value) is approximately 0.1 V/dec. (a gate insulating film: 100 nm thickness). In addition, the aforementioned transistor has characteristics of a normally-off transistor (threshold voltage thereof is positive in the case of an n-channel transistor).

[0017]

Therefore, leakage current, that is, an off-state current at a state where a voltage between a gate electrode and a source electrode is approximately 0 V is much smaller than that of a transistor using silicon. For example, in the transistor whose channel width W is $1 \times 10^4 \mu\text{m}$, a leakage current at room temperature per one micrometer of a channel width is lower than or equal to 10 aA (hereinafter, in this specification, this is described as "the leakage current per unit channel width is less than or equal to 10 aA/ μm at room temperature").

[0018]

Accordingly, with the transistor including an oxide semiconductor layer for a channel formation region, which serves as a switching element, a charge accumulated in the capacitor of the data storing portion can be kept stored without any change even after supply of a power source voltage to the latch circuit has stopped. In other words, data written into the data holding portion can be kept held without any change.

[0019]

For example, a refresh time and retention can be made much longer than those of a DRAM including a transistor using silicon for a channel formation region, and memory retention characteristics (a data holding property) which are substantially at the same level as a nonvolatile memory can be realized. In addition, after the supply of the power source voltage to the latch circuit has started again, the data held in the data holding portion can be read to the latch portion with the transistor. Accordingly, the logical state can be restored to the logical state prior to the stop of the supply of the power source voltage.

[0020]

Further, in temperature characteristics, the off-state current can be sufficiently low and the on-state current can be sufficiently high even at a high temperature. For

example, as V_G - I_D characteristics of the transistor including an oxide semiconductor layer for a channel formation region, data is obtained in a range of $-25\text{ }^{\circ}\text{C}$ to $150\text{ }^{\circ}\text{C}$ with low temperature dependence of off-state currents, on-state currents, mobilities, and S values. Furthermore, data is obtained which shows that the off-state current in the

5 aforementioned temperature range is as extremely low as $1 \times 10^{-13}\text{ A}$ or less. One of the reasons is that an i-type or substantially i-type oxide semiconductor which is obtained by sufficiently reducing the hydrogen concentration to be highly purified and has extremely low carrier concentration is used as the oxide semiconductor.

[0021]

10 Note that in this specification, an oxide semiconductor having a carrier density which is less than $1 \times 10^{11}\text{ /cm}^3$ is called an "intrinsic or i-type oxide semiconductor", and an oxide semiconductor having a carrier density greater than or equal to $1 \times 10^{11}\text{ /cm}^3$ but less than $1 \times 10^{12}\text{ /cm}^3$ is called a "substantially-intrinsic oxide semiconductor".

15 [0022]

In this manner, an embodiment of the present invention is to provide a nonvolatile latch circuit which has a wide operating temperature range and operates stably even at a high temperature and in which a logical state of storing data is not erased even after power is turned off.

20 [0023]

In the aforementioned nonvolatile latch circuit, various logic circuits can be provided by using the nonvolatile latch circuit. In addition, various semiconductor devices using the logic circuits can be provided. For example, among a plurality of block circuits of the logic circuit, supply of a power source voltage to one or the

25 plurality of block circuits which is/are not used can be stopped. With the use of the nonvolatile latch circuit, the logical state of the block circuit can be kept stored even after the supply of the power source voltage to the block circuit has stopped. Moreover, the stored logical state can be read after the supply of the power source voltage to the block circuit has started again. Accordingly, the logical state can be

30 restored to the logical state prior to the stop of the supply of the power source voltage.

[0024]

In the aforementioned nonvolatile latch circuit, as the oxide semiconductor layer, any of the following oxide semiconductors can be used: an In-Sn-Ga-Zn-O-based oxide semiconductor which is a four-component metal oxide; an In-Ga-Zn-O-based oxide semiconductor, an In-Sn-Zn-O-based oxide semiconductor, an In-Al-Zn-O-based oxide semiconductor, a Sn-Ga-Zn-O-based oxide semiconductor, an Al-Ga-Zn-O-based oxide semiconductor, or a Sn-Al-Zn-O-based oxide semiconductor which are three-component metal oxides; an In-Zn-O-based oxide semiconductor, a Sn-Zn-O-based oxide semiconductor, an Al-Zn-O-based oxide semiconductor, a Zn-Mg-O-based oxide semiconductor, a Sn-Mg-O-based oxide semiconductor, or an In-Mg-O-based oxide semiconductor which are two-component metal oxides; or an In-O-based oxide semiconductor; a Sn-O-based oxide semiconductor; or a Zn-O-based oxide semiconductor which are one-component metal oxides. In addition, the aforementioned oxide semiconductors may contain SiO₂.

[0025]

Note that in this specification, for example, an In-Sn-Ga-Zn-O-based oxide semiconductor means an oxide semiconductor containing at least In, Sn, Ga, and Zn, where a composition ratio of each metal element is not limited. In addition, a metal element other than In, Sn, Ga, and Zn may be contained therein.

[0026]

As the oxide semiconductor layer, a thin film containing a material represented by InMO₃(ZnO)_m ($m > 0$) can be used. Here, M represents one or more of metal elements selected from Ga, Al, Mn, and Co. For example, M can be Ga, Ga and Al, Ga and Mn, Ga and Co, or the like.

[0027]

In the aforementioned nonvolatile latch circuit, the hydrogen concentration in the oxide semiconductor layer can be set to less than or equal to $5 \times 10^{19} / \text{cm}^3$, preferably less than or equal to $5 \times 10^{18} / \text{cm}^3$, more preferably less than or equal to $5 \times 10^{17} / \text{cm}^3$, and still more preferably less than $1 \times 10^{16} / \text{cm}^3$. Further, the carrier concentration in the oxide semiconductor layer can be set to less than $1 \times 10^{14} / \text{cm}^3$, preferably less than $1 \times 10^{12} / \text{cm}^3$, more preferably less than $1 \times 10^{11} / \text{cm}^3$. The off-state current of the transistor using such an i-type oxide semiconductor or

substantially i-type oxide semiconductor can be set to less than or equal to 1×10^{-17} A, preferably 1×10^{-18} A.

[0028]

In the aforementioned nonvolatile latch circuit, the transistor using an oxide semiconductor may be a bottom-gate type, a top-gate type, a bottom-contact type, or a top-contact type. A bottom-gate transistor includes at least a gate electrode over an insulating surface; a gate insulating film over the gate electrode; and an oxide semiconductor layer to be a channel formation region over the gate electrode, with the gate insulating film interposed therebetween.

[0029]

The top-gate transistor includes at least an oxide semiconductor layer to be a channel formation region over an insulating surface; a gate insulating film over the oxide semiconductor layer; and a gate electrode over the oxide semiconductor layer, with the gate insulating film interposed therebetween. The bottom-contact transistor includes an oxide semiconductor layer to be a channel formation region over a source electrode and a drain electrode. The top-contact transistor includes a source electrode and a drain electrode over an oxide semiconductor layer to be a channel formation region.

[0030]

Note that the term such as "over" or "below" in this specification does not necessarily mean that a component is placed "directly on" or "directly under" another component. For example, the expression "a gate electrode over a gate insulating layer" does not exclude the case where a component is placed between the gate insulating layer and the gate electrode. Moreover, the terms such as "over" and "below" are only used for convenience of description and can include the case where the relation of components is reversed, unless otherwise specified.

[0031]

In this specification, the term of "electrode" or "wiring" does not limit the function of components. For example, an "electrode" can be used as part of "wiring", and the "wiring" can be used as part of the "electrode". In addition, the term of "electrode" or "wiring" can also mean a combination of a plurality of "electrodes" and

"wirings", for example.

[0032]

Note that functions of the "source" and the "drain" may be switched in the case where transistors of different polarities are employed or in the case where the direction
5 of a current flow changes in a circuit operation. Therefore, the terms "source" and "drain" can be replaced with each other in this specification.

[0033]

Note that in this specification, the term "electrically connected" includes the case where components are connected through an object having any electric function.
10 There is no particular limitation on an object having any electric function as long as electric signals can be transmitted and received between components that are connected through the object.

[0034]

Examples of an object having any electric function are a switching element
15 such as a transistor, a resistor, an inductor, a capacitor, and an element with a variety of functions as well as an electrode and a wiring.

[0035]

Note that in general, the term "SOI substrate" means a substrate having a silicon semiconductor layer over its insulating surface. In this specification, the term
20 "SOI substrate" also means a substrate having a semiconductor layer using a material other than silicon over its insulating surface. That is, a semiconductor layer included in the "SOI substrate" is not limited to a silicon semiconductor layer.

[0036]

A substrate in the "SOI substrate" is not limited to a semiconductor substrate
25 such as a silicon wafer and can be a non-semiconductor substrate such as a glass substrate, a quartz substrate, a sapphire substrate, or a metal substrate. That is, an "SOI substrate" also includes a conductive substrate and an insulating substrate over which a layer is formed using a semiconductor material.

[0037]

30 Further, in this specification, the term "semiconductor substrate" means not only a substrate formed using only a semiconductor material but also all substrates including a semiconductor material. That is, in this specification, the "SOI substrate"

is also included in the category of the "semiconductor substrate".

[0038]

According to an embodiment of the present invention, with the use of a transistor using an oxide semiconductor as a semiconductor material for forming a channel formation region, which serves as a switching element of a data holding portion, a nonvolatile latch circuit which has a wide operating temperature range and operates stably even at a high temperature and in which a logical state of storing data is not erased even after power is turned off, or a latch circuit provided with a data holding portion where a refresh period is sufficiently long can be realized. Since data writing is performed by switching of the transistor, the number of rewrites is not substantially limited. In addition, the writing voltage is almost equivalent to the threshold voltage of the transistor; thus, the data writing can be performed at a low voltage. Further, the charge accumulated in the capacitor of the data storing portion can be kept held as data without any change; therefore, the variation in the amount of charge which is stored as data can be suppressed small and data can be read easily as compared to the case where remanent polarization is used as data.

[0039]

Various logic circuits can be provided by using the nonvolatile latch circuit. For example, in the logic circuit using the nonvolatile latch circuit, the power consumption can be reduced by turning off power of the unused block. In addition, since a logical state is stored even when power is turned off, a system can be started when power is turned on or terminated when power is turned off, at high speed and low power.

25 BRIEF DESCRIPTION OF DRAWINGS

[0040]

FIG. 1 illustrates an example of a configuration of a nonvolatile latch circuit.

FIGS. 2A and 2B illustrate an example of a configuration of part of a nonvolatile latch circuit.

30 FIGS. 3A and 3B are cross-sectional views illustrating an example of elements of a nonvolatile latch circuit.

FIGS. 4A to 4H illustrate an example of a manufacturing method of an element

of a nonvolatile latch circuit.

FIGS. 5A to 5G illustrate an example of a manufacturing method of an element of a nonvolatile latch circuit.

FIGS. 6A to 6D illustrate an example of a manufacturing method of an element
5 of a nonvolatile latch circuit.

FIG. 7 is a cross-sectional view of a transistor using an oxide semiconductor.

FIG. 8 is an energy band diagram (schematic diagram) along an A-A' section in FIG. 7.

FIG. 9A illustrates a state where a positive voltage ($V_G > 0$) is applied to a gate
10 (GE1), and FIG. 9B illustrates a state where a negative voltage ($V_G < 0$) is applied to the gate (GE1).

FIG. 10 illustrates the relationships between the vacuum level and the work function of a metal (ϕ_M) and between the vacuum level and the electron affinity (χ) of an oxide semiconductor.

15 FIG. 11 illustrates energy required for hot carrier injection in silicon (Si).

FIG. 12 illustrates energy required for hot carrier injection in an In-Ga-Zn-O-based oxide semiconductor (IGZO).

FIG. 13 illustrates energy required for hot carrier injection in silicon carbide (4H-SiC).

20 FIG. 14 shows the results of device simulation as to short-channel effect.

FIG. 15 shows the results of device simulation as to short-channel effect.

FIG. 16 illustrates an example of a cross section of elements of a nonvolatile latch circuit.

FIGS. 17A to 17E illustrate an example of a manufacturing method of an
25 element of a nonvolatile latch circuit.

FIGS. 18A to 18D illustrate an example of a manufacturing method of an element of a nonvolatile latch circuit.

FIG. 19A illustrates an example of a configuration of a nonvolatile latch circuit, and FIG. 19B illustrates an example of an operation of the nonvolatile latch circuit.

30 FIGS. 20A and 20B each illustrate an example of an operation of a nonvolatile latch circuit.

FIG. 21 illustrates an example of a configuration of a nonvolatile latch circuit.

FIG. 22 illustrates an example of a configuration of a nonvolatile latch circuit.

FIG. 23 illustrates an example of a configuration of a nonvolatile latch circuit.

FIGS. 24A and 24B each illustrate an example of an operation of a nonvolatile latch circuit.

5 FIG. 25 illustrates an example of an operation of a nonvolatile latch circuit.

FIG. 26 illustrates a configuration of a nonvolatile latch circuit.

FIGS. 27A to 27E illustrate an example of a manufacturing method of an element of a nonvolatile latch circuit.

10 FIGS. 28A to 28E illustrate an example of a manufacturing method of an element of a nonvolatile latch circuit.

FIGS. 29A to 29F each illustrate an example of an electronic device including a semiconductor device using a nonvolatile latch circuit.

BEST MODE FOR CARRYING OUT THE INVENTION

15 [0041]

Hereinafter, embodiments of the present invention are described below with reference to the drawings. However, the present invention is not limited to the following description. It is easily understood by those skilled in the art that the mode and detail can be changed in various ways unless departing from the scope and spirit of the present invention. Therefore, the present invention should not be construed as being limited to the description in the following embodiments. In describing structures of the present invention with reference to the drawings, reference numerals denoting the same components are used in different drawings.

[0042]

25 Note that the size, the thickness of a layer, and a region of each structure illustrated in the drawings and the like in embodiments are exaggerated for simplicity in some cases. Therefore, embodiments of the present invention are not limited to such scales.

[0043]

30 Note that terms with ordinal numbers such as "first", "second", and "third" in this specification are used in order to identify components, and the terms do not limit the components numerically.

[0044]

[Embodiment 1]

In this embodiment, a configuration and an operation of a nonvolatile latch circuit which is an embodiment of the invention disclosed herein; and a structure, a manufacturing method, and the like of elements of the nonvolatile latch circuit will be described with reference to FIG. 1, FIGS. 2A and 2B, FIGS. 3A and 3B, FIGS. 4A to 4H, FIGS. 5A to 5G, FIGS. 6A to 6D, FIG. 7, FIG. 8, FIGS. 9A and 9B, FIG. 10, FIG. 11, FIG. 12, FIG. 13, FIG. 14, and FIG. 15.

[0045]

10 <Configuration and Operation of Nonvolatile Latch Circuit>

FIG. 1 illustrates a configuration of a nonvolatile latch circuit 400 including a latch portion 411 and a data holding portion 401 for holding data of the latch portion.

[0046]

The nonvolatile latch circuit 400 in FIG. 1 includes the latch portion 411 having a loop structure and the data holding portion 401 for holding data of the latch portion. In the latch portion 411 having a loop structure, an output of a first element (D1) 412 is electrically connected to an input of a second element (D2) 413, and an output of the second element (D2) 413 is electrically connected to an input of the first element (D1) 412.

20 [0047]

The input of the first element (D1) 412 is electrically connected to a wiring 414 supplied with an input signal of the latch circuit. The output of the first element (D1) 412 is electrically connected to a wiring 415 supplied with an output signal of the latch circuit.

25 [0048]

When there is a plurality of inputs of the first element (D1) 412, one of the inputs can be electrically connected to the wiring 414 supplied with the input signal of the latch circuit. When there is a plurality of inputs of the second element (D2) 413, one of the inputs can be electrically connected to the output of the first element (D1) 412.

[0049]

As the first element (D1) 412, an element in which inputted signal is inverted

and the resulting signal serves as an output can be used. For example, as the first element (D1) 412, an inverter, a NAND, a NOR, a clocked inverter, or the like can be used. As the second element (D2) 413, an element in which inputted signal is inverted and the resulting signal serves as an output can be used. For example, as the second
5 element (D2) 413, an inverter, a NAND, a NOR, a clocked inverter, or the like can be used.

[0050]

In the data holding portion 401, a transistor 402 using an oxide semiconductor as a semiconductor material for forming a channel formation region is used as a
10 switching element. In addition, the data holding portion 401 includes a capacitor 404 which is electrically connected to a source electrode or a drain electrode of the transistor 402. In other words, one of electrodes of the capacitor 404 is electrically connected to one of the source electrode and the drain electrode of the transistor 402. The other of the source electrode and the drain electrode of the transistor 402 is electrically
15 connected to the input of the first element or a wiring supplied with the input signal of the latch circuit. The other electrode of the capacitor 404 is supplied with a potential V_c .

[0051]

As the data holding portion 401, the configuration illustrated in FIGS. 2A and
20 2B can be employed instead of the configuration illustrated in FIG. 1.

[0052]

In a data holding portion 401 in FIG. 2A, a transistor 402 includes a first gate electrode and a second gate electrode. An oxide semiconductor layer for forming a channel formation region is provided between the first gate electrode and the second
25 gate electrode. The first gate electrode is electrically connected to a wiring supplied with a control signal. The second gate electrode is electrically connected to a wiring supplied with a predetermined potential. For example, the second gate electrode is electrically connected to a wiring supplied with a negative potential or a ground potential (GND).

30 [0053]

In addition, in the data holding portion 401 in FIG. 2A, one of the electrodes of the capacitor 404 is electrically connected to one of the source electrode and the drain

electrode of the transistor 402. The other of the source electrode and the drain electrode of the transistor 402 is electrically connected to the input of the first element or the wiring supplied with the input signal of the latch circuit. The other electrode of the capacitor 404 is supplied with a potential V_c .

5 [0054]

With the nonvolatile latch circuit using the data holding portion 401 in FIG. 2A, an advantageous effect that adjustment of electric characteristics (e.g., a threshold voltage) of the transistor 402 is facilitated can be obtained in addition to the advantageous effect of the nonvolatile latch circuit in FIG. 1. For example, when the
10 second gate electrode of the transistor 402 is supplied with a negative potential, the transistor 402 can be normally off easily.

[0055]

In a data holding portion 401 in FIG. 2B, the transistor 402 includes a first gate electrode and a second gate electrode. An oxide semiconductor layer for forming a
15 channel formation region is provided between the first gate electrode and the second gate electrode. The second gate electrode is electrically connected to the first gate electrode. In addition, in the data holding portion 401 in FIG. 2B, one of the electrodes of the capacitor 404 is electrically connected to one of the source electrode and the drain
20 electrode of the transistor 402. The other of the source electrode and the drain electrode of the transistor 402 is electrically connected to the input of the first element or the wiring supplied with the input signal of the latch circuit. The other electrode of the capacitor 404 is supplied with a potential V_c . With the nonvolatile latch circuit using the data holding portion 401 in FIG. 2B, an advantageous effect that the amount of
25 current in the transistor 402 is increased can be obtained in addition to the advantageous effect of the nonvolatile latch circuit in FIG. 1.

[0056]

In the nonvolatile latch circuit having the configurations illustrated in FIG. 1 and FIGS. 2A and 2B, the following writing, holding, and reading of data can be performed. Note that although the description will be made below with reference to
30 the configuration of FIG. 1, the aforementioned operations can be similarly performed in the case of the other configurations.

[0057]

The transistor 402 using an oxide semiconductor has a function of writing data held in the latch portion 411 into the capacitor 404 of the data holding portion 401. In addition, the transistor 402 has a function of holding the data written into the capacitor 404 of the data holding portion 401. Moreover, the transistor 402 has a function of reading the data held in the capacitor 404 of the data holding portion 401 to the latch portion 411.

[0058]

A writing operation of the data held in the latch portion 411 into the data holding portion 401, a holding operation of the data, a reading operation of the data from the data holding portion 401 to the latch portion 411, and a rewriting operation of the data of the data holding portion 401 will be described. First, the transistor 402 is turned on by supplying a gate electrode of the transistor 402 with a potential at which the transistor 402 is turned on. Accordingly, one of the electrodes of the capacitor 404 is supplied with the data held in the latch portion, that is, a potential of the input of the first element (D1) 412, which is held in the latch portion. As a result, the charge corresponding to the potential of the input of the first element (D1) 412, which is held in the latch portion, is accumulated in one of the electrodes of the capacitor 404 (this operation corresponds to writing).

[0059]

After that, the transistor 402 is turned off in such a manner that a potential of the gate electrode of the transistor 402 is set to a potential at which the transistor 402 is turned off. Accordingly, the charge accumulated in one of the electrodes of the capacitor 404 is held (holding). In addition, after the potential of the input of the first element (D1) 412 is brought into a floating state, the transistor 402 is turned on by supplying the gate electrode of the transistor 402 with a potential at which the transistor 402 is turned on. Accordingly, the charge is distributed to one of the electrodes of the capacitor 404 and the input of the first element (D1) 412. As a result, the input of the first element (D1) 412 is supplied with the potential corresponding to the charge accumulated in one of the electrodes of the capacitor 404. Then, the data is held in the latch portion. As a result, the data can be read (reading). Rewriting of the data can be performed in a manner similar to that of the writing and holding of the data.

[0060]

As the oxide semiconductor layer included in the transistor 402, any of the following oxide semiconductors can be used: an In-Sn-Ga-Zn-O-based oxide semiconductor which is a four-component metal oxide; an In-Ga-Zn-O-based oxide semiconductor, an In-Sn-Zn-O-based oxide semiconductor, an In-Al-Zn-O-based oxide semiconductor, a Sn-Ga-Zn-O-based oxide semiconductor, an Al-Ga-Zn-O-based oxide semiconductor, or a Sn-Al-Zn-O-based oxide semiconductor which are three-component metal oxides; an In-Zn-O-based oxide semiconductor, a Sn-Zn-O-based oxide semiconductor, an Al-Zn-O-based oxide semiconductor, a Zn-Mg-O-based oxide semiconductor, a Sn-Mg-O-based oxide semiconductor, or an In-Mg-O-based oxide semiconductor which are two-component metal oxides; or an In-O-based oxide semiconductor; a Sn-O-based oxide semiconductor; or a Zn-O-based oxide semiconductor which are one-component metal oxides. In addition, the aforementioned oxide semiconductors may contain SiO₂.

[0061]

As the oxide semiconductor layer, a thin film containing a material represented by InMO₃(ZnO)_m ($m > 0$) can be used. Here, M represents one or more of metal elements selected from Ga, Al, Mn, and Co. For example, M can be Ga, Ga and Al, Ga and Mn, Ga and Co, or the like.

[0062]

The oxide semiconductor layer is preferably an oxide semiconductor layer which is highly purified by sufficiently removing an impurity such as hydrogen and supplying oxygen. Specifically, the hydrogen concentration in the oxide semiconductor layer, which is measured by SIMS (secondary ion mass spectroscopy), can be set to less than or equal to $5 \times 10^{19} / \text{cm}^3$, preferably less than or equal to $5 \times 10^{18} / \text{cm}^3$, more preferably less than or equal to $5 \times 10^{17} / \text{cm}^3$, and still more preferably less than $1 \times 10^{16} / \text{cm}^3$.

[0063]

Further, the carrier concentration in the oxide semiconductor layer can be set to less than $1 \times 10^{14} / \text{cm}^3$, preferably less than $1 \times 10^{12} / \text{cm}^3$, more preferably less than $1 \times 10^{11} / \text{cm}^3$. In the oxide semiconductor layer which is highly purified by sufficiently reducing the hydrogen concentration and supplying oxygen, the carrier concentration is

sufficiently low (e.g., less than $1 \times 10^{12} / \text{cm}^3$, preferably less than $1 \times 10^{11} / \text{cm}^3$), as compared to carrier concentration (approximately $1 \times 10^{14} / \text{cm}^3$) in a general silicon wafer (a silicon wafer to which a slight amount of impurity elements such as phosphorus or boron is added).

5 [0064]

In this manner, by using an i-type or a substantially i-type oxide semiconductor which is obtained by sufficiently reducing the hydrogen concentration to be highly purified and have extremely low carrier concentration, the transistor 402 which has extremely favorable off-state current characteristics can be obtained. For example, even in the case of an element whose channel width W is $1 \times 10^4 \mu\text{m}$ and channel length L is $3 \mu\text{m}$, when a drain voltage V_D which is applied to a drain electrode is +1 V or +10 V and a gate voltage V_G which is applied to a gate electrode ranges from -5 V to -20 V, the off-state current is less than or equal to 1×10^{-13} A at room temperature. Moreover, the aforementioned transistor has characteristics of a normally-off transistor. Therefore, leakage current, that is, an off-state current at a state where a voltage between a gate electrode and a source electrode is approximately 0 V is much smaller than that of a transistor using silicon. For example, the leakage current per unit channel width is less than or equal to 10 aA/ μm at room temperature.

[0065]

Further, in temperature characteristics, a transistor in which an off-state current can be sufficiently low and an on-state current can be sufficiently high even at a high temperature can be obtained. For example, as V_G - I_D characteristics of the transistor 402, data is obtained in a range of -25 °C to 150 °C with low temperature dependence of on-state currents, mobilities, and S values. Furthermore, data is obtained which shows that the off-state current in the aforementioned temperature range is as extremely low as 1×10^{-13} A or less (less than or equal to the measurement limit). One of the reasons is that an i-type or substantially i-type oxide semiconductor which is obtained by sufficiently reducing the hydrogen concentration to be highly purified and has extremely low carrier concentration is used as the oxide semiconductor.

30 [0066]

In this manner, by using as a switching element the transistor 402 using an

i-type or substantially i-type oxide semiconductor which is obtained by sufficiently reducing the hydrogen concentration to be highly purified and have extremely low carrier concentration, the charge accumulated in the capacitor 404 of the data holding portion 401 can be kept held for an extremely long time even after supply of a power source voltage to the latch circuit 400 has stopped. In other words, the data written into the data holding portion 401 can be kept held for an extremely long time.

[0067]

For example, in the transistor 402, a refresh time and retention can be made much longer than those of a DRAM including a transistor using silicon for a channel formation region, and memory retention characteristics (a data holding property) which are substantially at the same level as a nonvolatile memory can be realized. In addition, the logical state can be restored to the logical state prior to the stop of the supply of the power source voltage by reading the data held in the data holding portion 401. In this manner, by using as a switching element the transistor 402 using an i-type or substantially i-type oxide semiconductor which is obtained by sufficiently reducing the hydrogen concentration to be highly purified and have extremely low carrier concentration, a novel nonvolatile latch circuit which has a wide operating temperature range and operates stably even at a high temperature and in which a logical state of storing data is not erased even after power is turned off can be realized.

[0068]

<Structure of Elements of Nonvolatile Latch Circuit>

Among elements of the nonvolatile latch circuit 400, a material other than an oxide semiconductor can be used as a semiconductor material for the elements other than the transistor 402 using an oxide semiconductor. As the material other than an oxide semiconductor, single crystal silicon, crystalline silicon, or the like can be used. For example, the element other than the transistor 402 can be provided over a substrate containing a semiconductor material. As the substrate containing a semiconductor material, a silicon wafer, an SOI (silicon on insulator) substrate, a silicon film over an insulating surface, or the like can be used. With the use of the material other than an oxide semiconductor, high-speed operation can be realized. For example, the first element (D1) 412 and the second element (D2) 413 of the latch portion can be formed with a transistor using a material other than an oxide semiconductor.

[0069]

FIGS. 3A and 3B illustrate an example of the structure of the elements of the nonvolatile latch circuit. In FIG. 3A, a transistor 160 using a material other than an oxide semiconductor is provided in a lower portion, and the transistor 402 using an oxide semiconductor is provided in an upper portion. The transistor 160 using a material other than an oxide semiconductor can be used as a transistor included in the first element (D1) 412 and the second element (D2) 413 of the latch portion. Also other elements of the nonvolatile latch circuit can have a structure similar or comparable to that of the transistor 160.

[0070]

The element such as the capacitor 404 of the nonvolatile latch circuit can be formed using a conductive film, a semiconductor film, an insulating film, or the like included in the transistor 402 or the transistor 160. Note that the transistor 160 and the transistor 402 are n-channel transistors here; alternatively, a p-channel transistor may be used. It is easy to use a p-channel transistor as the transistor 160. FIG. 3B illustrates an example where the connection relationship between the transistor 402 and an electrode (or a wiring) in the lower portion differs from that in FIG. 3A. The structure in FIG. 3A is mainly described below.

[0071]

The transistor 160 includes a channel formation region 116 provided in a substrate 100 containing a semiconductor material; impurity regions 114 and high-concentration regions 120 (a combination of the impurity regions 114 and the high-concentration regions 120 can simply be referred to as impurity regions) formed with the channel formation region 116 interposed therebetween; a gate insulating layer 108a over the channel formation region 116; a gate electrode 110a over the gate insulating layer 108a; a source or drain electrode 130a electrically connected to the impurity region 114; and a source or drain electrode 130b electrically connected to the impurity region 114 (see FIG. 3A).

[0072]

Here, sidewall insulating layers 118 are formed on the sides of the gate electrode 110a. Moreover, as seen from a plane, the high-concentration regions 120 are formed in a region of the substrate 100 which does not overlap with the sidewall

insulating layers 118, and metal compound regions 124 is in contact with the high-concentration regions 120. Further, element isolation insulating layers 106 are formed over the substrate 100 so as to surround the transistor 160, and an interlayer insulating layer 126 and an interlayer insulating layer 128 are formed to cover the transistor 160.

[0073]

The source or drain electrode 130a and the source or drain electrode 130b are electrically connected to the metal compound regions 124 through openings formed in the interlayer insulating layer 126 and the interlayer insulating layer 128. In other words, the source or drain electrode 130a and the source or drain electrode 130b are electrically connected to the high-concentration regions 120 and the impurity regions 114 through the metal compound regions 124.

[0074]

The transistor 402 includes a gate electrode 136d over the interlayer insulating layer 128; a gate insulating layer 138 over the gate electrode 136d; an oxide semiconductor layer 140 over the gate insulating layer 138; and a source or drain electrode 142a and a source or drain electrode 142b which are over the oxide semiconductor layer 140 and electrically connected to the oxide semiconductor layer 140 (see FIG. 3A).

[0075]

A protective insulating layer 144 is formed over the transistor 402 so as to be in contact with part of the oxide semiconductor layer 140. An interlayer insulating layer 146 is formed over the protective insulating layer 144. Here, the protective insulating layer 144 and the interlayer insulating layer 146 are provided with openings that reach the source or drain electrode 142a and the source or drain electrode 142b. An electrode 150d and an electrode 150e are in contact with the source or drain electrode 142a and the source or drain electrode 142b through the openings.

[0076]

At the same time as formation of the electrode 150d and the electrode 150e, an electrode 150a, an electrode 150b, and an electrode 150c are formed which are in contact with the electrode 136a, the electrode 136b, and the electrode 136c, respectively, through openings in the gate insulating layer 138, the protective insulating layer 144,

and the interlayer insulating layer 146. Note that although the example of a bottom-gate transistor is shown as the transistor 402, the present invention is not limited thereto. A top-gate transistor may also be employed.

[0077]

5 Here, the oxide semiconductor layer 140 is preferably an oxide semiconductor layer which is highly purified by sufficiently removing an impurity such as hydrogen and supplying oxygen. Specifically, the hydrogen concentration in the oxide semiconductor layer 140, which is measured by SIMS (secondary ion mass spectroscopy), can be set to less than or equal to $5 \times 10^{19}/\text{cm}^3$, preferably less than or
10 equal to $5 \times 10^{18}/\text{cm}^3$, more preferably less than or equal to $5 \times 10^{17}/\text{cm}^3$, and still more preferably less than $1 \times 10^{16}/\text{cm}^3$.

[0078]

Note that in the oxide semiconductor layer 140 which is highly purified by sufficiently reducing the hydrogen concentration and supplying oxygen, the carrier
15 concentration is sufficiently low (e.g., less than $1 \times 10^{12}/\text{cm}^3$, preferably less than $1 \times 10^{11}/\text{cm}^3$), as compared to carrier concentration (approximately $1 \times 10^{14}/\text{cm}^3$) in a general silicon wafer (a silicon wafer to which a slight amount of impurity elements such as phosphorus or boron is added).

[0079]

20 In this manner, by using an i-type or substantially i-type oxide semiconductor, the transistor 402 which has extremely favorable off-state current characteristics can be obtained. For example, when a drain voltage V_D is +1 V or +10 V and a gate voltage V_G ranges from -5 V to -20 V, the off-state current is less than or equal to 1×10^{-13} A at room temperature. Moreover, the aforementioned transistor has characteristics of a
25 normally-off transistor. Therefore, leakage current, that is, an off-state current at a state where a voltage between a gate electrode and a source electrode is approximately 0 V is much smaller than that of a transistor using silicon. For example, the leakage current per unit channel width is less than or equal to 10 aA/ μm at room temperature.

[0080]

30 Further, in temperature characteristics, a transistor in which an off-state current can be sufficiently low and an on-state current can be sufficiently high even at a high

temperature can be obtained. For example, as V_G - I_D characteristics of the transistor 402, data is obtained in a range of $-25\text{ }^{\circ}\text{C}$ to $150\text{ }^{\circ}\text{C}$ with low temperature dependence of on-state currents, mobilities, and S values. Furthermore, data is obtained which shows that the off-state current in the aforementioned temperature range is as extremely
5 low as 1×10^{-13} A or less. One of the reasons is that an i-type or substantially i-type oxide semiconductor which is obtained by sufficiently reducing the hydrogen concentration to be highly purified and has extremely low carrier concentration is used as the oxide semiconductor.

[0081]

10 When the oxide semiconductor layer 140 which is highly purified by sufficiently reducing the hydrogen concentration is used and the off-state current of the transistor 402 is reduced, a semiconductor device having a novel structure can be realized.

[0082]

15 An insulating layer 152 is formed over the interlayer insulating layer 146. An electrode 154a, an electrode 154b, an electrode 154c, and an electrode 154d are formed so as to be embedded in the insulating layer 152. Here, the electrode 154a is in contact with the electrode 150a, the electrode 154b is in contact with the electrode 150b, the electrode 154c is in contact with the electrode 150c and the electrode 150d, and the
20 electrode 154d is in contact with the electrode 150e.

[0083]

That is, the source or drain electrode 142a of the transistor 402 is electrically connected to another element (e.g., the transistor using a material other than an oxide semiconductor) through an electrode 130c, the electrode 136c, the electrode 150c, the
25 electrode 154c, and the electrode 150d (see FIG. 3A). In addition, the source or drain electrode 142b of the transistor 402 is electrically connected to another element through the electrode 150e and the electrode 154d. Note that the structure of connection electrodes (such as the electrode 130c, the electrode 136c, the electrode 150c, the electrode 154c, and the electrode 150d) is not limited to the aforementioned structure,
30 and appropriate addition, omission, or the like is possible.

[0084]

FIG. 3B illustrates the case where the connection relationship of the source or drain electrode 142a of the transistor 402 differs from that in FIG. 3A. Specifically, the source or drain electrode 142a is electrically connected to an electrode 110b through the electrode 130c, the electrode 136c, the electrode 150c, the electrode 154c, and the electrode 150d. Here, the electrode 110b is formed in a manner similar to that of the gate electrode 110a. The electrode 110b may be a component of the transistor or may be part of a wiring or the like. Note that the structure of connection electrodes (such as the electrode 130c, the electrode 136c, the electrode 150c, the electrode 154c, and the electrode 150d) is not limited to the aforementioned structure, and appropriate addition, omission, or the like is possible.

[0085]

Although the two examples of typical connection relationships are given above, an embodiment of the invention disclosed herein is not limited to these examples. For example, the structure illustrated in FIG. 3A and the structure illustrated in FIG. 3B may be combined. Furthermore, the gate electrode 110a of the transistor 160 and the source or drain electrode 142a of the transistor 402 may be electrically connected to each other.

[0086]

<Manufacturing Method of Elements of Nonvolatile Latch Circuit>

Next, an example of a manufacturing method of the elements of the nonvolatile latch circuit will be described. First, a manufacturing method of the transistor 160 will be described below with reference to FIGS. 4A to 4H, and then a manufacturing method of the transistor 402 will be described below with reference to FIGS. 5A to 5G and FIGS. 6A to 6D. With the manufacturing method described below, the elements of the nonvolatile latch circuit can be manufactured. Note that FIGS. 4A to 4H illustrate only a cross section corresponding to A1-A2 in FIG. 3A. In addition, FIGS. 5A to 5G and FIGS. 6A to 6D illustrate cross sections corresponding to A1-A2 and B1-B2 in FIG. 3A.

[0087]

<Manufacturing Method of Transistor in Lower Portion>

First, the substrate 100 containing a semiconductor material is prepared (see FIG. 4A). A single crystal semiconductor substrate or a polycrystalline semiconductor substrate of silicon, carbon silicon, or the like; a compound semiconductor substrate of silicon germanium or the like; an SOI substrate, or the like can be used as the substrate

100 containing a semiconductor material. Here, an example of the case where a single crystal silicon substrate is used as the substrate 100 containing a semiconductor material is described.

[0088]

5 Note that in general, the term "SOI substrate" means a substrate having a silicon semiconductor layer over its insulating surface. In this specification, the term "SOI substrate" also means a substrate having a semiconductor layer using a material other than silicon over its insulating surface. That is, a semiconductor layer included in the "SOI substrate" is not limited to a silicon semiconductor layer. Examples of the
10 SOI substrate include a substrate which has a semiconductor layer over its insulating substrate such as a glass substrate.

[0089]

 A protective layer 102 that serves as a mask for forming an element isolation insulating layer is formed over the substrate 100 (see FIG. 4A). An insulating layer of
15 silicon oxide, silicon nitride, silicon nitride oxide, or the like, for example, can be used as the protective layer 102. Note that before and after this step, an impurity element imparting n-type conductivity or an impurity element imparting p-type conductivity can be added to the substrate 100 in order to control the threshold voltage of the transistor. As the impurity imparting n-type conductivity, phosphorus, arsenic, or the like can be
20 used, for example, when the semiconductor material contained in the substrate 100 is silicon. As the impurity imparting p-type conductivity, boron, aluminum, gallium, or the like can be used, for example.

[0090]

 Next, with the use of the aforementioned protective layer 102 as a mask, part of
25 the substrate 100 in a region which is not covered with the protective layer 102 (an exposed region) is removed by etching. Thus, an isolated semiconductor region 104 is formed (see FIG. 4B). As the etching, dry etching is preferably performed, but wet etching can be performed. An etching gas and an etchant can be selected as appropriate depending on a material of layers to be etched.

30 [0091]

 Next, an insulating layer is formed to cover the semiconductor region 104 and a region of the insulating layer which overlaps with the semiconductor region 104 is

selectively removed, so that the element isolation insulating layers 106 are formed (see FIG. 4B). The insulating layer is formed using silicon oxide, silicon nitride, silicon nitride oxide, or the like. Methods for removing the insulating layer include etching, polishing such as CMP, and the like, and any of these are applicable. Note that after the semiconductor region 104 is formed or after the element isolation insulating layers 106 are formed, the protective layer 102 is removed.

[0092]

Next, an insulating layer is formed over the semiconductor region 104, and a layer containing a conductive material is formed over the insulating layer.

[0093]

Because the insulating layer serves as a gate insulating layer later, the insulating layer preferably has a single-layer structure or a stacked structure using a film containing silicon oxide, silicon nitride oxide, silicon nitride, hafnium oxide, aluminum oxide, tantalum oxide, or the like formed with a CVD method, a sputtering method, or the like. Alternatively, the insulating layer can be formed by oxidizing or nitriding a surface of the semiconductor region 104 by high-density plasma treatment or thermal oxidation treatment. The high-density plasma treatment can be performed using, for example, a mixed gas of a rare gas such as He, Ar, Kr, or Xe and a gas such as oxygen, nitrogen oxide, ammonia, nitrogen, or hydrogen. There is no particular limitation on the thickness of the insulating layer; the insulating layer can be formed in the range of 1 nm to 100 nm inclusive, for example.

[0094]

The layer containing a conductive material can be formed using a metal material such as aluminum, copper, titanium, tantalum, or tungsten. Alternatively, the layer containing a conductive material may be formed using a semiconductor material such as polycrystalline silicon containing an impurity element imparting conductivity. There is no particular limitation on the film formation method of the layer containing a conductive material; a variety of film formation methods such as an evaporation method, a CVD method, a sputtering method, and a spin coating method can be employed. Note that in this embodiment, an example of the case where the layer containing a conductive material is formed using a metal material is described.

[0095]

After that, the insulating layer and the layer containing a conductive material are selectively etched, so that the gate insulating layer 108a and the gate electrode 110a are formed (see FIG. 4C).

[0096]

5 Next, an insulating layer 112 that covers the gate electrode 110a is formed (see FIG. 4C). Then, the impurity regions 114 with a shallow junction depth are formed by adding phosphorus (P), arsenic (As), or the like to the semiconductor region 104 (see FIG. 4C). Note that phosphorus or arsenic is added here in order to form an n-channel transistor; however, an impurity element such as boron (B) or aluminum (Al) may be
10 added in the case of forming a p-channel transistor.

[0097]

With the formation of the impurity regions 114, the channel formation region 116 is formed in the semiconductor region 104 below the gate insulating layer 108a (see FIG. 4C). Here, the concentration of the impurity added can be set as appropriate; the
15 concentration is preferably increased when the size of a semiconductor element is extremely decreased. The step in which the impurity regions 114 are formed after the formation of the insulating layer 112 is employed here; alternatively, the insulating layer 112 may be formed after the formation of the impurity regions 114.

[0098]

20 Next, the sidewall insulating layers 118 are formed (see FIG. 4D). When, as the sidewall insulating layers 118, an insulating layer is formed to cover the insulating layer 112 and then subjected to highly anisotropic etching, the sidewall insulating layers 118 can be formed in a self-aligned manner. At this time, it is preferable to etch the insulating layer 112 partly so that a top surface of the gate electrode 110a and top
25 surfaces of the impurity regions 114 are exposed.

[0099]

Then, an insulating layer is formed to cover the gate electrode 110a, the impurity regions 114, the sidewall insulating layers 118, and the like. Next, phosphorus (P), arsenic (As), or the like is added to regions in which the insulating layer
30 is in contact with the impurity regions 114, so that the high-concentration impurity regions 120 are formed. After that, the insulating layer is removed, and a metal layer 122 is formed to cover the gate electrode 110a, the sidewall insulating layers 118, the

high-concentration impurity regions 120, and the like (see FIG. 4E).

[0100]

A variety of film formation methods such as a vacuum evaporation method, a sputtering method, or a spin coating method can be employed for forming the metal layer 122. The metal layer 122 is preferably formed using a metal material that reacts with a semiconductor material included in the semiconductor region 104 to be a low-resistance metal compound. Examples of such a metal material are titanium, tantalum, tungsten, nickel, cobalt, and platinum.

[0101]

Next, heat treatment is performed so that the metal layer 122 reacts with the semiconductor material. Thus, the metal compound regions 124 that are in contact with the high-concentration impurity regions 120 are formed (see FIG. 4F). Note that when the gate electrode 110a is formed using polycrystalline silicon or the like, a metal compound region is also formed in a region of the gate electrode 110a in contact with the metal layer 122.

[0102]

As the heat treatment, irradiation with a flash lamp can be employed, for example. Although it is needless to say that another heat treatment method may be used, a method with which heat treatment for an extremely short time can be achieved is preferably used in order to improve the controllability of chemical reaction in formation of the metal compound. Note that the metal compound regions are formed by reaction of the metal material and the semiconductor material and have sufficiently high conductivity. The formation of the metal compound regions can properly reduce the electric resistance and improve element characteristics. Note that the metal layer 122 is removed after the metal compound regions 124 are formed.

[0103]

Then, the interlayer insulating layer 126 and the interlayer insulating layer 128 are formed to cover the components formed in the aforementioned steps (see FIG. 4G). The interlayer insulating layer 126 and the interlayer insulating layer 128 can be formed using a material including an inorganic insulating material such as silicon oxide, silicon nitride oxide, silicon nitride, hafnium oxide, aluminum oxide, or tantalum oxide. Moreover, the interlayer insulating layer 126 and the interlayer insulating layer 128 can

also be formed using an organic insulating material such as polyimide or acrylic. Note that a two-layer structure of the interlayer insulating layer 126 and the interlayer insulating layer 128 is employed here; however, the structure of an interlayer insulating layer is not limited to this structure. After the formation of the interlayer insulating layer 128, a surface of the interlayer insulating layer 128 is preferably planarized with
5 CMP, etching, or the like.

[0104]

Then, openings that reach the metal compound regions 124 are formed in the interlayer insulating layers, and the source or drain electrode 130a and the source or drain electrode 130b are formed in the openings (see FIG. 4H). The source or drain electrode 130a and the source or drain electrode 130b can be formed in such a manner that, for example, a conductive layer is formed in a region including the openings with a PVD method, a CVD method, or the like and then part of the conductive layer is removed by etching, CMP, or the like.

15 [0105]

Note that in the case where the source or drain electrode 130a and the source or drain electrode 130b are formed by removing part of the conductive layer, the process is preferably performed so that the surfaces are planarized. For example, when a thin titanium film or a thin titanium nitride film is formed in a region including the openings and then a tungsten film is formed to be embedded in the openings, excess tungsten, titanium, titanium nitride, or the like is removed and the planarity of the surface can be improved by subsequent CMP. When the surface including the source or drain electrode 130a and the source or drain electrode 130b is planarized in such a manner, an electrode, a wiring, an insulating layer, a semiconductor layer, and the like can be
20 favorably formed in later steps.

[0106]

There is no particular limitation on a material used for the source or drain electrode 130a and the source or drain electrode 130b, and a variety of conductive materials can be used. For example, a conductive material such as molybdenum, titanium, chromium, tantalum, tungsten, aluminum, copper, neodymium, or scandium
30 can be used. In addition, although only the source or drain electrode 130a and the source or drain electrode 130b which are in contact with the metal compound regions

124 are illustrated here, the electrode 130c and the like in FIGS. 3A and 3B can also be formed together in this step.

[0107]

Specifically, it is possible to employ a method, for example, in which a thin
5 titanium film is formed in a region including the openings with a PVD method and a thin titanium nitride film is formed with a CVD method, and then, a tungsten film is formed to be embedded in the openings. Here, the titanium film formed with a PVD method has a function of reducing an oxide film which might be formed on the surface of the metal compound regions to decrease the contact resistance with the metal
10 compound regions. The titanium nitride film formed after the formation of the titanium film has a barrier function of preventing diffusion of the conductive material. A copper film may be formed with a plating method after the formation of the barrier film of titanium, titanium nitride, or the like. Note that not only a so-called single damascene method but also a dual damascene method may be employed.

15 [0108]

Through the aforementioned steps, the transistor 160 using the substrate 100 containing a semiconductor material is formed. Note that an electrode, a wiring, an insulating layer, or the like may be further formed after the aforementioned steps. When the wirings have a multi-layer structure of a stacked structure including an
20 interlayer insulating layer and a conductive layer, a highly integrated semiconductor device can be provided.

[0109]

<Manufacturing Method of Transistor in Upper Portion>

Next, steps for manufacturing the transistor 402 over the interlayer insulating
25 layer 128 will be described with reference to FIGS. 5A to 5G and FIGS. 6A to 6D. Note that FIGS. 5A to 5G and FIGS. 6A to 6D illustrate steps for manufacturing electrodes, the transistor 402, and the like over the interlayer insulating layer 128; therefore, the transistor 160 and the like placed below the transistor 402 are omitted.

[0110]

30 First, the insulating layer 132 is formed over the interlayer insulating layer 128, the source or drain electrode 130a, the source or drain electrode 130b, and the electrode 130c (see FIG. 5A). Next, openings that reach the source or drain electrode 130a, the

source or drain electrode 130b, and the electrode 130c are formed in the insulating layer 132. Then, a conductive layer 134 is formed to be embedded in the openings (see FIG. 5B). After that, part of the conductive layer 134 is removed by etching, CMP, or the like, so that the insulating layer 132 is exposed and the electrode 136a, the electrode 136b, the electrode 136c, and the gate electrode 136d are formed (see FIG. 5C).

[0111]

The insulating layer 132 can be formed with a PVD method, a CVD method, or the like. The insulating layer 132 can be formed using a material containing an inorganic insulating material such as silicon oxide, silicon nitride oxide, silicon nitride, hafnium oxide, aluminum oxide, or tantalum oxide.

[0112]

The openings can be formed in the insulating layer 132 with a method such as etching using a mask. The mask can be formed with a method such as light exposure using a photomask. Either wet etching or dry etching may be used as the etching; dry etching is preferably used in terms of microfabrication.

[0113]

The conductive layer 134 can be formed with a film formation method such as a PVD method or a CVD method. The conductive layer 134 can be formed using a conductive material such as molybdenum, titanium, chromium, tantalum, tungsten, aluminum, copper, neodymium, or scandium or an alloy or a compound (e.g., a nitride) of any of these materials, for example (see FIG. 5B).

[0114]

More specifically, it is possible to employ a method, for example, in which a thin titanium film is formed in a region including the openings with a PVD method and a thin titanium nitride film is formed with a CVD method, and then, a tungsten film is formed to be embedded in the openings. Here, the titanium film formed with a PVD method has a function of reducing an oxide film which might be formed on the surface of lower electrodes (here, the source or drain electrode 130a, the source or drain electrode 130b, the electrode 130c, and the like) to decrease the contact resistance with lower electrodes.

[0115]

The titanium nitride film formed after the formation of the titanium film has a

barrier function of preventing diffusion of the conductive material. A copper film may be formed with a plating method after the formation of the barrier film of titanium, titanium nitride, or the like. Note that not only a so-called single damascene method but also a dual damascene method may be employed.

5 [0116]

After the conductive layer 134 is formed, part of the conductive layer 134 is removed by etching, CMP, or the like, so that the insulating layer 132 is exposed and the electrode 136a, the electrode 136b, the electrode 136c, and the gate electrode 136d can be formed (see FIG. 5C). Note that when the electrode 136a, the electrode 136b, the electrode 136c, and the gate electrode 136d are formed by removing part of the conductive layer 134, the process is preferably performed so that the surfaces are planarized. When the surfaces of the insulating layer 132, the electrode 136a, the electrode 136b, the electrode 136c, and the gate electrode 136d are planarized in such a manner, an electrode, a wiring, an insulating layer, a semiconductor layer, and the like can be favorably formed in later steps.

15 [0117]

Next, the gate insulating layer 138 is formed to cover the insulating layer 132, the electrode 136a, the electrode 136b, the electrode 136c, and the gate electrode 136d (see FIG. 5D). The gate insulating layer 138 can be formed with a CVD method, a sputtering method, or the like. The gate insulating layer 138 is preferably formed using silicon oxide, silicon nitride, silicon oxynitride, silicon nitride oxide, aluminum oxide, hafnium oxide, tantalum oxide, or the like. Note that the gate insulating layer 138 may have a single-layer structure or a stacked structure.

20 [0118]

For example, the gate insulating layer 138 can be formed using silicon oxynitride with a plasma CVD method using silane (SiH_4), oxygen, and nitrogen as a source gas. There is no particular limitation on the thickness of the gate insulating layer 138; the gate insulating layer 138 can have a thickness of 10 nm to 500 nm inclusive, for example. In the case of employing a stacked structure, for example, the gate insulating layer 138 is preferably a stack of a first gate insulating layer with a thickness of 50 nm to 200 nm inclusive, and a second gate insulating layer with a thickness of 5 nm to 300 nm inclusive over the first gate insulating layer.

[0119]

If hydrogen, water, or the like is contained in the gate insulating layer 138, hydrogen may enter the oxide semiconductor layer or extract oxygen from the oxide semiconductor layer, whereby characteristics of the transistor might be degraded.

5 Therefore, it is preferable to form the gate insulating layer 138 so as to contain hydrogen or water as little as possible.

[0120]

In the case where a sputtering method or the like is employed, for example, it is preferable that the gate insulating layer 138 be formed in a state where moisture in the treatment chamber is removed. In order to remove moisture in the treatment chamber, an entrapment vacuum pump such as a cryopump, an ion pump, or a titanium sublimation pump is preferably used. Alternatively, a turbo pump provided with a cold trap may be used. Since hydrogen, water, or the like is sufficiently removed from the treatment chamber evacuated with a cryopump or the like, the concentration of an impurity contained in the gate insulating layer 138 can be reduced.

[0121]

When the gate insulating layer 138 is formed, it is preferable to use a high-purity gas in which an impurity such as hydrogen or water is reduced to a concentration of a few ppm or less (preferably, a few ppb or less).

20 [0122]

Note that an i-type or substantially i-type oxide semiconductor which is obtained by removing an impurity (a highly purified oxide semiconductor) is quite susceptible to the interface level and the interface charge; therefore, when such an oxide semiconductor is used for an oxide semiconductor layer, the interface with the gate insulating layer is important. In other words, the gate insulating layer 138 that is to be in contact with a highly purified oxide semiconductor layer needs to have high quality.

[0123]

For example, the gate insulating layer 138 is preferably formed with a high-density plasma CVD method using a microwave (frequency: 2.45 GHz) because the gate insulating layer 138 can be dense and have high withstand voltage and high quality. When a highly purified oxide semiconductor layer and a high-quality gate insulating layer are in close contact with each other, the interface level can be reduced

and favorable interface characteristics can be obtained.

[0124]

It is needless to say that, even when a highly purified oxide semiconductor layer is used, another method such as a sputtering method or a plasma CVD method can be employed as long as a high-quality insulating layer can be formed as the gate insulating layer 138. Moreover, it is possible to use an insulating layer whose film quality and interface characteristics with the oxide semiconductor layer are improved with heat treatment performed after the formation of the gate insulating layer 138. In any case, an insulating layer that has favorable film quality as the gate insulating layer 138 and can reduce interface level density with an oxide semiconductor layer to form a favorable interface is formed as the gate insulating layer 138.

[0125]

Next, an oxide semiconductor layer is formed over the gate insulating layer 138 and processed with a method such as etching using a mask, so that the island-shaped oxide semiconductor layer 140 is formed (see FIG. 5E).

[0126]

As the oxide semiconductor layer, any of the following oxide semiconductors can be used: an In-Sn-Ga-Zn-O-based oxide semiconductor which is a four-component metal oxide; an In-Ga-Zn-O-based oxide semiconductor, an In-Sn-Zn-O-based oxide semiconductor, an In-Al-Zn-O-based oxide semiconductor, a Sn-Ga-Zn-O-based oxide semiconductor, an Al-Ga-Zn-O-based oxide semiconductor, or a Sn-Al-Zn-O-based oxide semiconductor which are three-component metal oxides; an In-Zn-O-based oxide semiconductor, a Sn-Zn-O-based oxide semiconductor, an Al-Zn-O-based oxide semiconductor, a Zn-Mg-O-based oxide semiconductor, a Sn-Mg-O-based oxide semiconductor, or an In-Mg-O-based oxide semiconductor which are two-component metal oxides; or an In-O-based oxide semiconductor; a Sn-O-based oxide semiconductor; or a Zn-O-based oxide semiconductor which are one-component metal oxides. In addition, the aforementioned oxide semiconductors may contain SiO₂.

[0127]

As the oxide semiconductor layer, a thin film containing a material represented by $\text{InMO}_3(\text{ZnO})_m$ ($m > 0$) can be used. Here, M represents one or more of metal elements selected from Ga, Al, Mn, and Co. For example, M can be Ga, Ga and Al,

Ga and Mn, Ga and Co, or the like.

[0128]

In this embodiment, as the oxide semiconductor layer, an amorphous oxide semiconductor layer is formed with a sputtering method using an In-Ga-Zn-O-based metal oxide target. Note that since crystallization of an amorphous oxide semiconductor layer can be suppressed by adding silicon to the amorphous oxide semiconductor layer, an oxide semiconductor layer may be formed, for example, using a target containing SiO₂ of 2 wt% to 10 wt% inclusive.

[0129]

As the metal oxide target used for forming an oxide semiconductor layer with a sputtering method, for example, a metal oxide target having a composition ratio of In₂O₃:Ga₂O₃:ZnO = 1:1:1 [molar ratio] can be used. Furthermore, a metal oxide target having a composition ratio of In₂O₃:Ga₂O₃:ZnO = 1:1:2 [molar ratio] or a metal oxide target having a composition ratio of In₂O₃:Ga₂O₃:ZnO = 1:1:4 [molar ratio] can also be used. The filling rate of a metal oxide target is 90 % to 100 % inclusive, preferably greater than or equal to 95 % (e.g., 99.9 %). A dense oxide semiconductor layer is formed by using a metal oxide target with a high filling rate.

[0130]

The atmosphere in which the oxide semiconductor layer is formed is preferably a rare gas (typically argon) atmosphere, an oxygen atmosphere, or a mixed atmosphere containing a rare gas (typically argon) and oxygen. Specifically, it is preferable to use, for example, a high-purity gas atmosphere from which an impurity such as hydrogen, water, a hydroxyl group, or hydride is removed to a concentration of a few ppm or less (preferably, a few ppb or less).

[0131]

At the time of forming the oxide semiconductor layer, the substrate is held in a treatment chamber kept under reduced pressure and the substrate is heated to a temperature of 100 °C to 600 °C inclusive, preferably 200 °C to 400 °C inclusive. The oxide semiconductor layer is formed while the substrate is heated, so that the concentration of an impurity contained in the oxide semiconductor layer can be reduced. Moreover, damage of the semiconductor layer due to sputtering is reduced. Then, a

sputtering gas from which hydrogen and water are removed is introduced into the treatment chamber while moisture in the treatment chamber is removed, whereby the oxide semiconductor layer is formed using metal oxide as a target.

[0132]

5 An entrapment vacuum pump is preferably used in order to remove moisture in the treatment chamber. For example, a cryopump, an ion pump, or a titanium sublimation pump can be used. An evacuation unit may be a turbo pump provided with a cold trap. In the film formation chamber that is evacuated with the cryopump, a hydrogen atom, a compound containing a hydrogen atom such as water (H₂O) (and
10 more preferably also a compound containing a carbon atom), and the like are removed, whereby the concentration of an impurity contained in the oxide semiconductor layer formed in the film formation chamber can be reduced.

[0133]

 The oxide semiconductor layer can be formed under the following conditions,
15 for example: the distance between the substrate and the target is 100 mm; the pressure is 0.6 Pa; the direct-current (DC) power supply is 0.5 kW; and the atmosphere is oxygen (the proportion of the oxygen flow is 100 %). Note that it is preferable to use a pulsed direct-current (DC) power source because dust can be reduced and the thickness distribution can be reduced. The thickness of the oxide semiconductor layer is 2 nm to
20 200 nm inclusive, preferably 5 nm to 30 nm inclusive. Note that the appropriate thickness of the oxide semiconductor layer differs depending on the oxide semiconductor material to be used; therefore, the thickness may be determined in accordance with the material to be used.

[0134]

25 Note that before the oxide semiconductor layer is formed with a sputtering method, reverse sputtering is preferably performed in which plasma is generated with an argon gas introduced, so that dust on the surface of the gate insulating layer 138 is removed. Here, the reverse sputtering is a method in which ions collide with a surface to be processed so that the surface is modified, in contrast to normal sputtering in which
30 ions collide with a sputtering target. An example of a method for making ions collide with a surface to be processed is a method in which a high-frequency voltage is applied to the surface to be processed under an argon atmosphere so that plasma is generated

near a substrate. Note that an atmosphere of nitrogen, helium, oxygen, or the like may be used instead of an argon atmosphere.

[0135]

As an etching method for the oxide semiconductor layer, either dry etching or wet etching may be employed. It is needless to say that dry etching and wet etching can be used in combination. The etching conditions (e.g., an etching gas or an etchant, etching time, and temperature) are set as appropriate depending on the material so that the oxide semiconductor layer can be etched into a desired shape.

[0136]

An example of an etching gas used for dry etching is a gas containing chlorine (a chlorine-based gas such as chlorine (Cl_2), boron trichloride (BCl_3), silicon tetrachloride (SiCl_4), or carbon tetrachloride (CCl_4)) or the like. Moreover, a gas containing fluorine (a fluorine-based gas such as carbon tetrafluoride (CF_4), sulfur hexafluoride (SF_6), nitrogen trifluoride (NF_3), or trifluoromethane (CHF_3)), hydrogen bromide (HBr), oxygen (O_2), any of these gases to which a rare gas such as helium (He) or argon (Ar) is added, or the like may be used.

[0137]

As the dry etching method, a parallel plate RIE (reactive ion etching) method or an ICP (inductively coupled plasma) etching method can be used. In order to etch the oxide semiconductor layer into a desired shape, etching conditions (e.g., the amount of electric power applied to a coiled electrode, the amount of electric power applied to an electrode on the substrate side, and the electrode temperature on the substrate side) are set as appropriate.

[0138]

As an etchant used for wet etching, a mixed solution of phosphoric acid, acetic acid, and nitric acid; ammonia peroxide mixture (a mixture of ammonia, water, and a hydrogen peroxide solution); or the like can be used. An etchant such as ITO07N (produced by KANTO CHEMICAL CO., INC.) may also be used.

[0139]

Then, first heat treatment is preferably performed on the oxide semiconductor layer. The oxide semiconductor layer can be dehydrated or dehydrogenated through the first heat treatment. The temperature of the first heat treatment is 300 °C to 800 °C

inclusive, preferably 400 °C to 700 °C inclusive, more preferably 450 °C to 700 °C inclusive, and still more preferably 550 °C to 700 °C inclusive.

[0140]

The oxide semiconductor layer can be dehydrated or dehydrogenated when the first heat treatment is performed at a temperature of 350 °C or higher, so that the hydrogen concentration in the oxide semiconductor layer can be reduced. In addition, when the first heat treatment is performed at a temperature of 450 °C or higher, the hydrogen concentration in the oxide semiconductor layer can be further reduced. Moreover, when the first heat treatment is performed at a temperature of 550 °C or higher, the hydrogen concentration in the oxide semiconductor layer can be still further reduced. For example, the substrate is introduced into an electric furnace in which a resistance heating element or the like is used and the oxide semiconductor layer 140 is subjected to heat treatment at 450 °C for 1 hour under a nitrogen atmosphere. The oxide semiconductor layer 140 is not exposed to the air during the heat treatment so that the entry of water or hydrogen can be prevented.

[0141]

The heat treatment apparatus is not limited to the electric furnace and can be an apparatus for heating an object to be processed by thermal conduction or thermal radiation from a medium such as a heated gas. For example, a rapid thermal annealing (RTA) apparatus such as a gas rapid thermal annealing (GRTA) apparatus or a lamp rapid thermal annealing (LRTA) apparatus can be used.

[0142]

An LRTA apparatus is an apparatus for heating an object to be processed by radiation of light (an electromagnetic wave) emitted from a lamp such as a halogen lamp, a metal halide lamp, a xenon arc lamp, a carbon arc lamp, a high-pressure sodium lamp, or a high-pressure mercury lamp. A GRTA apparatus is an apparatus for performing heat treatment using a high-temperature gas. As the gas, an inert gas that does not react with an object to be processed by heat treatment, for example, nitrogen or a rare gas such as argon is used.

[0143]

For example, as the first heat treatment, a GRTA process may be performed as

follows. The substrate is put in an inert gas atmosphere that has been heated to a high temperature of 650 °C to 700 °C, heated for several minutes, and taken out of the inert gas atmosphere. The GRTA process enables high-temperature heat treatment for a short time. Moreover, the GRTA process can be employed even when the temperature exceeds the strain point of the substrate because it is heat treatment for a short time. For example, in the case where an SOI substrate including a substrate with relatively low heat resistance, such as a glass substrate, is used, shrinkage of the substrate becomes a problem at a temperature higher than the upper temperature limit (strain point) but does not in the case where heat treatment is performed for a short time.

10 [0144]

Note that as the inert gas atmosphere under which the first heat treatment is performed, it is preferable to employ an atmosphere that contains nitrogen or a rare gas (e.g., helium, neon, or argon) as its main component and does not contain water, hydrogen, or the like. For example, the purity of nitrogen or a rare gas such as helium, neon, or argon introduced into the heat treatment apparatus is greater than or equal to 6 N (99.9999 %), preferably greater than or equal to 7 N (99.99999 %) (i.e., the impurity concentration is less than or equal to 1 ppm, preferably less than or equal to 0.1 ppm).

[0145]

Note that the inert gas atmosphere may be changed during the process to an atmosphere including oxygen. For example, in the case where an electrical furnace is used in the first heat treatment, an atmosphere can be changed when a heat treatment temperature falls. For example, the heat treatment can be performed (at a constant temperature) under an atmosphere of an inert gas such as a rare gas (e.g., helium, neon, or argon) or nitrogen, and the atmosphere can be switched to an atmosphere containing oxygen when the heat treatment temperature falls. As the atmosphere containing oxygen, an oxygen gas or a mixed gas of an oxygen gas and a nitrogen gas can be used. Also in the case where the atmosphere containing oxygen is employed, it is preferable that the atmosphere do not contain water, hydrogen, or the like. Alternatively, the purity of the oxygen gas or the nitrogen used is preferably greater than or equal to 6N (99.9999 %), more preferably greater than or equal to 7N (99.99999 %) (that is, the impurity concentration is less than or equal to 1 ppm, preferably less than or equal to

0.1 ppm). This is because defects caused by oxygen deficiency can be reduced by performing the first heat treatment under an atmosphere containing oxygen.

[0146]

In some cases, the oxide semiconductor layer is crystallized to be microcrystalline or polycrystalline, which depends on the conditions of the first heat treatment or the material of the oxide semiconductor layer. For example, in some cases, the oxide semiconductor layer becomes a microcrystalline oxide semiconductor layer having a degree of crystallization of 90 % or more, or 80 % or more. Further, in some cases, the oxide semiconductor layer may be an amorphous oxide semiconductor layer containing no crystalline component, which depends on the conditions of the first heat treatment or the material of the oxide semiconductor layer.

[0147]

Furthermore, in some cases, the oxide semiconductor layer becomes a layer in which a microcrystal (with a grain size of 1 nm to 20 nm inclusive, typically 2 nm to 4 nm inclusive) is mixed in an amorphous oxide semiconductor (e.g., a surface of the oxide semiconductor layer).

[0148]

The electric characteristics of the oxide semiconductor layer can be changed by aligning microcrystals in an amorphous semiconductor. For example, when the oxide semiconductor layer is formed using an In-Ga-Zn-O-based metal oxide target, the electric characteristics of the oxide semiconductor layer can be changed by formation of a microcrystalline region in which crystal grains of $\text{In}_2\text{Ga}_2\text{ZnO}_7$ with electrical anisotropy are aligned.

[0149]

For example, when the crystal grains are arranged so that the c-axis of $\text{In}_2\text{Ga}_2\text{ZnO}_7$ is perpendicular to a surface of the oxide semiconductor layer, the conductivity in the direction parallel to the surface of the oxide semiconductor layer can be improved and insulating properties in the direction perpendicular to the surface of the oxide semiconductor layer can be improved. Furthermore, such a microcrystalline region has a function of suppressing entry of an impurity such as water or hydrogen into the oxide semiconductor layer.

[0150]

Note that the oxide semiconductor layer including the microcrystalline region can be formed by heating the surface of the oxide semiconductor layer by a GRTA process. Further, the oxide semiconductor layer can be formed in a more preferred manner by using a sputtering target in which the amount of Zn is smaller than that of In or Ga.

[0151]

The first heat treatment for the oxide semiconductor layer 140 can be performed on the oxide semiconductor layer that has not yet been processed into the island-shaped oxide semiconductor layer 140. In that case, after the first heat treatment, the substrate is taken out of the heating apparatus and a photolithography step is performed.

[0152]

Note that the aforementioned heat treatment, which has an effect of dehydration or dehydrogenation on the oxide semiconductor layer 140, can also be referred to as dehydration treatment, dehydrogenation treatment, or the like. The dehydration treatment or dehydrogenation treatment can be performed, for example, after the oxide semiconductor layer is formed, after a source or drain electrode is stacked over the oxide semiconductor layer 140, or after a protective insulating layer is formed over the source or drain electrode. Such dehydration treatment or dehydrogenation treatment may be performed once or plural times.

[0153]

Next, the source or drain electrode 142a and the source or drain electrode 142b are formed so as to be in contact with the oxide semiconductor layer 140 (see FIG. 5F). The source or drain electrode 142a and the source or drain electrode 142b can be formed in such a manner that a conductive layer is formed to cover the oxide semiconductor layer 140 and then is selectively etched.

[0154]

The conductive layer can be formed with a PVD (Physical Vapor Deposition) method such as a sputtering method or a CVD (Chemical Vapor Deposition) method such as a plasma CVD method. As a material for the conductive layer, an element selected from aluminum, chromium, copper, tantalum, titanium, molybdenum, and tungsten; an alloy containing any of these elements as its component; or the like can be

used. Moreover, one or more of materials selected from manganese, magnesium, zirconium, beryllium, and yttrium may be used. It is also possible to use aluminum combined with one or more of elements selected from titanium, tantalum, tungsten, molybdenum, chromium, neodymium, and scandium.

5 [0155]

The conductive layer may be formed using an oxide conductive film. As the oxide conductive film, indium oxide (In_2O_3), tin oxide (SnO_2), zinc oxide (ZnO), an indium oxide-tin oxide alloy ($\text{In}_2\text{O}_3\text{-SnO}_2$, which is abbreviated to ITO in some cases), an indium oxide-zinc oxide alloy ($\text{In}_2\text{O}_3\text{-ZnO}$), or any of these metal oxide materials in
10 which silicon or silicon oxide is included can be used.

[0156]

In that case, as compared to a material for the oxide semiconductor layer 140, a material whose conductivity is high or whose resistivity is low is preferably used for the oxide conductive film. The conductivity of the oxide conductive film can be increased
15 by an increase in the carrier concentration. The carrier concentration in the oxide conductive film can be increased by an increase in the hydrogen concentration. Further, the carrier concentration in the oxide conductive film can be increased by an increase in oxygen deficiency.

[0157]

20 The conductive layer can have a single-layer structure or a stacked structure including two or more layers. For example, the conductive layer can have a single-layer structure of an aluminum film containing silicon, a two-layer structure in which a titanium film is stacked over an aluminum film, or a three-layer structure in which a titanium film, an aluminum film, and a titanium film are stacked in this order.
25 Here, a three-layer structure of a titanium film, an aluminum film, and a titanium film is employed.

[0158]

Note that an oxide conductive layer may be formed between the oxide semiconductor layer 140 and the conductive layer. The oxide conductive layer and the
30 conductive layer can be successively formed. By providing such an oxide conductive layer, the resistance of the source region or the drain region can be reduced, so that the transistor can operate at high speed.

[0159]

Next, the source or drain electrode 142a and the source or drain electrode 142b are formed by selectively etching the conductive layer (see FIG. 5F). Ultraviolet light, KrF laser light, or ArF laser light is preferably used for light exposure at the time of forming a mask used for etching.

[0160]

The channel length (L) of the transistor is determined by a distance between a lower edge portion of the source or drain electrode 142a and a lower edge portion of the source or drain electrode 142b. Note that in the case where light exposure is performed so that the channel length (L) is less than 25 nm, light exposure for forming a mask is performed with extreme ultraviolet rays whose wavelength is extremely short of several nanometers to several tens of nanometers. The resolution of light exposure with extreme ultraviolet rays is high and the depth of focus is large. For these reasons, it is possible to design a mask so that the channel length (L) of the transistor to be formed later is less than 25 nm, that is, in the range of 10 nm to 1000 nm inclusive, and the circuit can operate at higher speed. Moreover, the off-state current is extremely low, which prevents an increase in power consumption.

[0161]

The materials and etching conditions of the conductive layer and the oxide semiconductor layer 140 are adjusted as appropriate so that the oxide semiconductor layer 140 is not removed in etching of the conductive layer. Note that in some cases, the oxide semiconductor layer 140 is partly etched in the etching step and thus has a groove portion (a recessed portion) depending on the materials and the etching conditions.

[0162]

In order to reduce the number of masks to be used and reduce the number of steps, an etching step may be performed with the use of a resist mask formed using a multi-tone mask which is a light-exposure mask through which light is transmitted to have a plurality of intensities. A resist mask formed with the use of a multi-tone mask has a plurality of thicknesses (has a stair-like shape) and further can be changed in shape by ashing; therefore, the resist mask can be used in a plurality of etching steps for processing into different patterns. That is, a resist mask corresponding to at least two

kinds of different patterns can be formed by using a multi-tone mask. Thus, the number of light-exposure masks can be reduced and the number of corresponding photolithography steps can also be reduced, whereby a process can be simplified.

[0163]

5 Note that plasma treatment is preferably performed with the use of a gas such as N_2O , N_2 , or Ar after the aforementioned step. This plasma treatment removes water or the like attached to an exposed surface of the oxide semiconductor layer. Plasma treatment may be performed using a mixed gas of oxygen and argon.

[0164]

10 Next, the protective insulating layer 144 is formed in contact with part of the oxide semiconductor layer 140 without exposure to the air (see FIG. 5G).

[0165]

 The protective insulating layer 144 can be formed with a method such as a sputtering method, with which impurities such as water and hydrogen are prevented
15 from being mixed to the protective insulating layer 144, as appropriate. The protective insulating layer 144 has a thickness of at least 1 nm. The protective insulating layer 144 can be formed using silicon oxide, silicon nitride, silicon oxynitride, silicon nitride oxide, or the like. The protective insulating layer 144 can have a single-layer structure or a stacked structure. The substrate temperature at the time of forming the protective
20 insulating layer 144 is preferably room temperature to 300 °C inclusive. The atmosphere for forming the protective insulating layer 144 is preferably a rare gas (typically argon) atmosphere, an oxygen atmosphere, or a mixed atmosphere containing a rare gas (typically argon) and oxygen.

[0166]

25 If hydrogen is contained in the protective insulating layer 144, the hydrogen may enter the oxide semiconductor layer or extract oxygen in the oxide semiconductor layer, whereby the resistance of the oxide semiconductor layer on the backchannel side might be decreased and a parasitic channel might be formed. Therefore, it is important not to use hydrogen in forming the protective insulating layer 144 so that the oxide
30 insulating layer 144 contains hydrogen as little as possible.

[0167]

Moreover, the protective insulating layer 144 is preferably formed while water

in the treatment chamber is removed, in order that hydrogen, a compound containing a hydroxyl group, or moisture is not contained in the oxide semiconductor layer 140 and the protective insulating layer 144.

[0168]

5 An entrapment vacuum pump is preferably used in order to remove moisture in the treatment chamber. For example, a cryopump, an ion pump, or a titanium sublimation pump is preferably used. An evacuation unit may be a turbo pump provided with a cold trap. In the film formation chamber that is evacuated with the cryopump, a hydrogen atom and a compound containing a hydrogen atom, such as
10 water (H₂O), are removed, for example; thus, the concentration of an impurity contained in the protective insulating layer 144 formed in the film formation chamber can be reduced.

[0169]

 As a sputtering gas used at the time of forming the protective insulating layer
15 144, it is preferable to use a high-purity gas from which an impurity such as hydrogen, water, a compound containing a hydroxyl group, or hydride is removed to a concentration of 1 ppm or less (preferably, 1 ppb or less).

[0170]

 Next, second heat treatment is preferably performed under an inert gas
20 atmosphere or an oxygen gas atmosphere (preferably, at 200 °C to 400 °C inclusive, for example, at 250 °C to 350 °C inclusive). For example, the second heat treatment is performed at 250 °C for 1 hour under a nitrogen atmosphere. The second heat treatment can reduce variation in electric characteristics of the transistor.

[0171]

25 Furthermore, heat treatment may be performed at 100 °C to 200 °C inclusive for 1 hour to 30 hours inclusive in the air. This heat treatment may be performed at a fixed heating temperature; alternatively, the following change in the heating temperature may be repeatedly performed plural times: the heating temperature is increased from room temperature to a temperature of 100 °C to 200 °C inclusive and then decreased to
30 room temperature. This heat treatment may be performed under a reduced pressure before the protective insulating layer is formed. The heat treatment time can be

shortened under the reduced pressure. This heat treatment may be performed instead of the second heat treatment or may be performed before or after the second heat treatment.

[0172]

5 Next, the interlayer insulating layer 146 is formed over the protective insulating layer 144 (see FIG. 6A). The interlayer insulating layer 146 can be formed with a PVD method, a CVD method, or the like. The interlayer insulating layer 146 can be formed using a material including an inorganic insulating material such as silicon oxide, silicon nitride oxide, silicon nitride, hafnium oxide, aluminum oxide, or tantalum
10 oxide. After the formation of the interlayer insulating layer 146, a surface of the interlayer insulating layer 146 is preferably planarized with a method such as CMP or etching.

[0173]

 Next, openings that reach the electrode 136a, the electrode 136b, the electrode
15 136c, the source or drain electrode 142a, and the source or drain electrode 142b are formed in the interlayer insulating layer 146, the protective insulating layer 144, and the gate insulating layer 138. Then, a conductive layer 148 is formed to be embedded in the openings (see FIG. 6B). The openings can be formed with a method such as etching using a mask. The mask can be formed with a method such as light exposure
20 using a photomask.

[0174]

 Either wet etching or dry etching may be used as the etching; dry etching is preferably used in terms of microfabrication. The conductive layer 148 can be formed with a film formation method such as a PVD method or a CVD method. The
25 conductive layer 148 can be formed using a conductive material such as molybdenum, titanium, chromium, tantalum, tungsten, aluminum, copper, neodymium, or scandium or an alloy or a compound (e.g., a nitride) of any of these materials, for example.

[0175]

 Specifically, it is possible to employ a method, for example, in which a thin
30 titanium film is formed in a region including the openings with a PVD method and a thin titanium nitride film is formed with a CVD method, and then, a tungsten film is formed to be embedded in the openings. Here, the titanium film formed with a PVD

method has a function of reducing an oxide film at the interface with the interlayer insulating layer 146 to decrease the contact resistance with lower electrodes (here, the electrode 136a, the electrode 136b, the electrode 136c, the source or drain electrode 142a, and the source or drain electrode 142b). The titanium nitride film formed after the formation of the titanium film has a barrier function of preventing diffusion of the conductive material. A copper film may be formed with a plating method after the formation of the barrier film of titanium, titanium nitride, or the like.

[0176]

After the conductive layer 148 is formed, part of the conductive layer 148 is removed with a method such as etching or CMP, so that the interlayer insulating layer 146 is exposed and the electrode 150a, the electrode 150b, the electrode 150c, the electrode 150d, and the electrode 150e are formed (see FIG. 6C). Note that when the electrode 150a, the electrode 150b, the electrode 150c, the electrode 150d, and the electrode 150e are formed by removing part of the conductive layer 148, the process is preferably performed so that the surfaces are planarized. When the surfaces of the interlayer insulating layer 146, the electrode 150a, the electrode 150b, the electrode 150c, the electrode 150d, and the electrode 150e are planarized in such a manner, an electrode, a wiring, an insulating layer, a semiconductor layer, and the like can be favorably formed in later steps.

[0177]

Then, the insulating layer 152 is further formed, and openings that reach the electrode 150a, the electrode 150b, the electrode 150c, the electrode 150d, and the electrode 150e are formed in the insulating layer 152. After a conductive layer is formed to be embedded in the openings, part of the conductive layer is removed with a method such as etching or CMP. Thus, the insulating layer 152 is exposed and the electrode 154a, the electrode 154b, the electrode 154c, and the electrode 154d are formed (see FIG. 6D). This step is similar to the step of forming the electrode 150a and the like; therefore, the detailed description is omitted.

[0178]

In the case where the transistor 402 is formed with the aforementioned method, the hydrogen concentration in the oxide semiconductor layer 140 is $5 \times 10^{19} / \text{cm}^3$ or less

and the off-state current of the transistor 402 is 1×10^{-13} A or less at room temperature. The carrier concentration in the oxide semiconductor layer is less than 1×10^{14} /cm³. The transistor 402 with excellent characteristics can be obtained by the application of the oxide semiconductor layer 140 that is highly purified by sufficiently reducing the hydrogen concentration and supplying oxygen as described above. Moreover, since the nonvolatile latch circuit includes the transistor 160 formed using a material other than an oxide semiconductor in the lower portion and the transistor 402 formed using an oxide semiconductor in the upper portion, it is possible to manufacture an excellent nonvolatile latch circuit having characteristics of both the transistors and a semiconductor device using the nonvolatile latch circuit.

[0179]

Note that it is preferable that oxygen be supplied to the oxide semiconductor layer 140 shortly after the hydrogen concentration is reduced because there is no possibility that hydrogen, water, or the like enters the oxide semiconductor layer and thus an oxide semiconductor layer with extremely favorable characteristics can be realized. It is needless to say that treatment for reducing the hydrogen concentration and treatment for supplying oxygen do not need to be performed successively as long as an oxide semiconductor layer with favorable characteristics can be realized. For example, another treatment may be performed between the treatment for reducing the hydrogen concentration and treatment for supplying oxygen. Alternatively, both of the treatment may be performed at the same time.

[0180]

Note that silicon carbide (e.g., 4H-SiC) is given as a semiconductor material which can be compared to an oxide semiconductor. An oxide semiconductor and 4H-SiC have some things in common. The carrier density is one of them. In accordance with Fermi-Dirac distribution, the density of minority carriers in an oxide semiconductor is estimated to be approximately 1×10^{-7} /cm³. This value of the minority carrier density is extremely small similarly to that in 4H-SiC, 6.7×10^{-11} /cm³. When the minority carrier density of an oxide semiconductor is compared to the intrinsic carrier density of silicon (approximately 1.4×10^{10} /cm³), it can be understood well that the minority carrier density of an oxide semiconductor is significantly low.

[0181]

Further, the energy band gap of an oxide semiconductor is 3.0 eV to 3.5 eV and the energy band gap of 4H-SiC is 3.26 eV. Thus, an oxide semiconductor and silicon carbide are similar in that they are both wide-gap semiconductors.

5 [0182]

On the other hand, there is a major difference between an oxide semiconductor and silicon carbide, that is, the process temperature. Since heat treatment at 1500 °C to 2000 °C is generally needed in a semiconductor process using silicon carbide, it is difficult to form a stack of silicon carbide and a semiconductor element formed using a semiconductor material other than silicon carbide. This is because a semiconductor substrate, the semiconductor element, or the like is damaged at such high temperatures. Meanwhile, an oxide semiconductor can be formed with heat treatment at 300 °C to 500 °C (the glass transition temperature or lower, up to about 700 °C); therefore, it is possible to form an integrated circuit with the use of a semiconductor material other than an oxide semiconductor and then to form a semiconductor element including an oxide semiconductor.

10
15

[0183]

In addition, in contrast to silicon carbide, an oxide semiconductor is advantageous because a low heat-resistant substrate such as a glass substrate can be used. Moreover, an oxide semiconductor does not need to be subjected to heat treatment at high temperature, so that energy cost can be reduced sufficiently as compared to silicon carbide, which is another advantage.

20

[0184]

Although a lot of researches on properties of an oxide semiconductor such as density of state (DOS) have been conducted, they do not include the idea of sufficiently reducing localized states themselves. According to an embodiment of the invention disclosed herein, a highly purified oxide semiconductor is formed by removing water or hydrogen which might affect the localized states. This is based on the idea that the localized states themselves are sufficiently reduced. Such a highly purified oxide semiconductor enables fabrication of very excellent industrial products.

25
30

[0185]

Further, it is also possible to form a more highly purified (i-type) oxide semiconductor by supplying oxygen to a dangling bond of metal which is generated by oxygen vacancy and reducing the localized states due to the oxygen vacancy. For example, an oxide film containing excessive oxygen is formed in close contact with a channel formation region and then oxygen is supplied to the channel formation region from the oxide film, so that the localized states due to oxygen vacancy can be reduced.

[0186]

A defect of an oxide semiconductor is said to be attributed to a level of 0.1 eV to 0.2 eV under the conduction band due to excessive hydrogen, a deep level due to shortage of oxygen, or the like. Thorough removal of hydrogen and sufficient supply of oxygen for elimination of such a defect would be right as a technological thought.

[0187]

An oxide semiconductor is generally considered as an n-type semiconductor; however, according to an embodiment of the invention disclosed herein, an i-type semiconductor is realized by removing an impurity, particularly water and hydrogen. In this respect, it can be said that an embodiment of the invention disclosed herein includes a novel technical idea because it is different from an i-type semiconductor such as silicon added with an impurity.

[0188]

The example is described above in which, among elements of the nonvolatile latch circuit 400, a material other than an oxide semiconductor is used as a semiconductor material for the elements other than the transistor 402 using an oxide semiconductor. However, the invention disclosed herein is not limited thereto. Among the elements of the nonvolatile latch circuit 400, an oxide semiconductor can also be used as a semiconductor material for the elements other than the transistor 402.

[0189]

<Electrical Conduction Mechanism of Transistor Using Oxide Semiconductor>

An electrical conduction mechanism of a transistor using an oxide semiconductor will be described here with reference to FIG. 7, FIG. 8, FIGS. 9A and 9B, and FIG. 10. Note that the following description is based on the assumption of an ideal situation for easy understanding and does not necessarily reflect a real situation. Note also that the following description is just a consideration and does not affect the validity

of the invention.

[0190]

FIG. 7 is a cross-sectional view of a transistor (thin film transistor) using an oxide semiconductor. An oxide semiconductor layer (OS) is provided over a gate electrode (GE1) with a gate insulating layer (GI) interposed therebetween, and a source electrode (S) and a drain electrode (D) are provided over the oxide semiconductor layer. An insulating layer is provided so as to cover the source electrode (S) and the drain electrode (D).

[0191]

FIG. 8 is an energy band diagram (schematic diagram) of the cross section A-A' in FIG. 7. In FIG. 8, a black circle (●) and a white circle (○) represent an electron and a hole and have electric charges ($-q$, $+q$), respectively. With a positive voltage ($V_D > 0$) applied to the drain electrode, the dashed line shows the case where no voltage is applied to the gate electrode ($V_G = 0$) and the solid line shows the case where a positive voltage is applied to the gate electrode ($V_G > 0$). In the case where no voltage is applied to the gate electrode, carriers (electrons) are not injected to the oxide semiconductor side from an electrode because of high potential barrier, so that a current does not flow, which means an off state. On the other hand, when a positive voltage is applied to the gate electrode, potential barrier is lowered, and thus a current flows, which means an on state.

[0192]

FIGS. 9A and 9B are energy band diagrams (schematic diagrams) along B-B' in FIG. 7. FIG. 9A illustrates a state in which a positive voltage ($V_G > 0$) is applied to the gate electrode (GE1) and carriers (electrons) flow between the source electrode and the drain electrode. FIG. 9B illustrates an off state in which a negative voltage ($V_G < 0$) is applied to the gate electrode (GE1) and minority carriers do not flow.

[0193]

FIG. 10 illustrates the relationships between the vacuum level and the work function of a metal (ϕ_M) and between the vacuum level and the electron affinity (χ) of an oxide semiconductor. At normal temperature, electrons in the metal are degenerated and the Fermi level is located in the conduction band. On the other hand,

a conventional oxide semiconductor is an n-type semiconductor, in which the Fermi level (E_F) is away from the intrinsic Fermi level (E_i) located in the middle of a band gap and is located closer to the conduction band. Note that it is known that part of hydrogen is a donor in an oxide semiconductor and is one factor causing an oxide semiconductor to be an n-type semiconductor.

[0194]

On the other hand, an oxide semiconductor according to an embodiment of the invention disclosed herein is an intrinsic (i-type) or a substantially intrinsic oxide semiconductor which is obtained in the following manner: hydrogen that is a factor for an n-type oxide semiconductor from an oxide semiconductor and purifying the oxide semiconductor such that an element other than its main component of the oxide semiconductor (i.e., an impurity element) is prevented from being contained therein as much as possible. That is, a feature is that a purified i-type (intrinsic) semiconductor or a semiconductor close thereto is obtained not by adding an impurity element but by removing an impurity such as hydrogen and water as much as possible. Thus, the Fermi level (E_F) can be comparable with the intrinsic Fermi level (E_i).

[0195]

It is said that the band gap (E_g) of an oxide semiconductor is 3.15 eV and the electron affinity (χ) thereof is 4.3 eV. The work function of titanium (Ti) contained in the source electrode and the drain electrode is substantially equal to the electron affinity (χ) of the oxide semiconductor. In this case, a Schottky barrier is not formed for an electron is not formed at the interface between metal and the oxide semiconductor.

[0196]

At that time, as illustrated in FIG. 9A, the electron moves in the vicinity of the interface between the gate insulating layer and the highly purified oxide semiconductor (the lowest portion of the oxide semiconductor which is stable in terms of energy).

[0197]

In addition, as illustrated in FIG. 9B, when a negative potential is supplied to the gate electrode (GE1), the value of current is extremely close to zero because holes that are minority carriers are substantially zero.

[0198]

In such a manner, an intrinsic (i-type) or substantially intrinsic oxide semiconductor is obtained by being highly purified such that an element other than its main component (i.e., an impurity element) is contained as little as possible. Thus, characteristics of the interface between the oxide semiconductor and the gate insulating layer become important. For that reason, the gate insulating layer needs to be able to form a favorable interface with the oxide semiconductor. Specifically, it is preferable to use the following insulating layer, for example: an insulating layer formed with a CVD method using high-density plasma generated with a power source frequency in the range of the VHF band to the microwave band, or an insulating layer formed with a sputtering method.

[0199]

When the oxide semiconductor is highly purified and the interface between the oxide semiconductor and the gate insulating layer is made favorable, in the case where the transistor has a channel width (W) of $1 \times 10^4 \mu\text{m}$ and a channel length (L) of $3 \mu\text{m}$, for example, it is possible to realize an off-state current of $1 \times 10^{-13} \text{ A}$ or less and a subthreshold swing (S value) of 0.1 V/dec. (a gate insulating layer: 100 nm thickness).

[0200]

When the oxide semiconductor is highly purified as described above so as to contain an element other than its main component (i.e., an impurity element) as little as possible, the transistor can operate in a favorable manner.

[0201]

<Resistance of transistor using oxide semiconductor to hot carrier degradation>

Next, the resistance of a transistor using an oxide semiconductor to hot carrier degradation will be described with reference to FIG. 11, FIG. 12, and FIG. 13. Note that the following description is based on the assumption of an ideal situation for easy understanding and does not necessarily reflect a real situation. Note also that the following description is just a consideration.

[0202]

Main causes of hot carrier degradation are channel hot electron injection (CHE injection) and drain avalanche hot carrier injection (DAHC injection). Note that only electrons are considered below for simplicity.

[0203]

CHE injection refers to a phenomenon in which electrons having gained energy higher than the barrier of a gate insulating layer in a semiconductor layer are injected into the gate insulating layer or the like. Electrons gain energy by being accelerated by a low electric field.

[0204]

DAHC injection refers to a phenomenon in which electrons generated by collision of electrons accelerated by a high electric field are injected into a gate insulating layer or the like. A difference between DAHC injection and CHE injection is whether or not they involve avalanche breakdown caused by impact ionization. Note that DAHC injection requires electrons having a kinetic energy higher than a band gap of a semiconductor.

[0205]

FIG. 11 illustrates energy required for each hot carrier injection which is estimated from the band structure of silicon (Si), and FIG. 12 illustrates energy required for each hot carrier injection which is estimated from the band structure of an In-Ga-Zn-O-based oxide semiconductor (IGZO). The left of each of FIG. 11 and FIG. 12 show CHE injection, and the right of each of FIG. 11 and FIG. 12 show DAHC injection.

[0206]

Regarding silicon, degradation caused by DAHC injection is more serious than that caused by CHE injection. This results from the fact that carriers (e.g., electrons) which are accelerated without colliding are very few in silicon whereas silicon has a narrow band gap and avalanche breakdown readily occurs therein. The avalanche breakdown sharply increases the number of electrons capable of crossing over the barrier of the gate insulating layer (i.e., electrons injected into the gate insulating layer), which causes degradation.

[0207]

Regarding an In-Ga-Zn-O-based oxide semiconductor, the energy required for CHE injection does not greatly differ from that in the case of silicon, and the probability of CHE injection is still low. On the other hand, since the band gap of the In-Ga-Zn-O-based oxide semiconductor is wider than that of silicon, the energy

required for DAHC injection is increased and thus avalanche breakdown unlikely occurs. In other words, the probabilities of both CHE injection and DAHC injection are low, and hot carrier degradation unlikely occurs as compared to the case of using silicon.

5 [0208]

Meanwhile, the band gap of an In-Ga-Zn-O-based oxide semiconductor is comparable to that of silicon carbide (SiC) which attracts attention as a material having a high withstand voltage. FIG. 13 illustrates energy required for each hot carrier injection regarding 4H-SiC. Regarding CHE injection, an In-Ga-Zn-O-based oxide semiconductor has a slightly higher threshold and can be said to have an advantage.

[0209]

As described above, it can be seen that an In-Ga-Zn-O-based oxide semiconductor has significantly higher resistance to hot carrier degradation and higher resistance to source-drain breakdown than silicon. It can also be said that a withstand voltage comparable to that of silicon carbide can be obtained.

[0210]

<Short-channel effect in transistor using oxide semiconductor>

Next, a short-channel effect in a transistor using an oxide semiconductor will be described with reference to FIG. 14 and FIG. 15. Note that the following description is based on the assumption of an ideal situation for easy understanding and does not necessarily reflect a real situation. Note also that the following description is just a consideration.

[0211]

The short-channel effect refers to degradation of electric characteristics which becomes obvious with miniaturization of a transistor (a reduction in channel length (L)). The short-channel effect results from the effect of a drain on a source. Specific examples of the short-channel effect are a decrease in threshold voltage, an increase in subthreshold swing (S value), an increase in leakage current, and the like.

[0212]

Here, a structure capable of suppressing a short-channel effect is examined by device simulation. Specifically, four kinds of models each having a different carrier concentration and a different thickness of an oxide semiconductor layer were prepared,

and the relationship between a channel length (L) and a threshold voltage (V_{th}) was checked. As the models, bottom-gate transistors were employed, in each of which an oxide semiconductor had a carrier concentration of $1.7 \times 10^{-8} / \text{cm}^3$ or $1.0 \times 10^{15} / \text{cm}^3$ and an oxide semiconductor layer with a thickness of 1 μm or 30 nm. Note that an In-Ga-Zn-O-based oxide semiconductor was used for the oxide semiconductor layer, and a silicon oxynitride film with a thickness of 100 nm was used as a gate insulating layer. It was assumed that, in the oxide semiconductor, the band gap was 3.15 eV, the electron affinity was 4.3 eV, the relative permittivity was 15, and the electron mobility was $10 \text{ cm}^2 / \text{Vs}$. The relative permittivity of the silicon oxynitride film was assumed to be 4.0. The calculation was performed using device simulation software "ATLAS" produced by Silvaco Inc.

[0213]

Note that there is no significant difference in calculation results between a top-gate transistor and a bottom-gate transistor. FIG. 14 and FIG. 15 show calculation results. FIG. 14 shows the case where the carrier concentration is $1.7 \times 10^{-8} / \text{cm}^3$, and FIG. 15 shows the case where the carrier concentration is $1.0 \times 10^{15} / \text{cm}^3$. FIG. 14 and FIG. 15 each show the amount of change (ΔV_{th}) in threshold voltage (V_{th}) when a transistor whose channel length (L) is 10 μm is used as a reference and channel lengths (L) vary from 10 μm to 1 μm . As shown in FIG. 14, in the case where the carrier concentration in the oxide semiconductor was $1.7 \times 10^{-8} / \text{cm}^3$ and the thickness of the oxide semiconductor layer was 1 μm , the amount of change (ΔV_{th}) in threshold voltage was -3.6 V . Moreover, as shown in FIG. 14, in the case where the carrier concentration in the oxide semiconductor was $1.7 \times 10^{-8} / \text{cm}^3$ and the thickness of the oxide semiconductor layer was 30 nm, the amount of change (ΔV_{th}) in threshold voltage was -0.2 V . In addition, as shown in FIG. 15, in the case where the carrier concentration in the oxide semiconductor was $1.0 \times 10^{15} / \text{cm}^3$ and the thickness of the oxide semiconductor layer was 1 μm , the amount of change (ΔV_{th}) in threshold voltage was -3.6 V . Moreover, as shown in FIG. 15, in the case where the carrier concentration in the oxide semiconductor was $1.0 \times 10^{15} / \text{cm}^3$ and the thickness of the oxide semiconductor layer was 30 nm, the amount of change (ΔV_{th}) in threshold voltage was -0.2 V . The results show that a short-channel effect can be suppressed in a transistor

using an oxide semiconductor by a reduction in thickness of an oxide semiconductor layer. For example, in the case where the channel length (L) is approximately 1 μm , even with an oxide semiconductor layer having sufficiently high carrier concentration, it can be understood that a short-channel effect can be sufficiently suppressed when the
5 thickness of the oxide semiconductor layer is set to approximately 30 nm.

[0214]

With the use of a transistor using an oxide semiconductor as a semiconductor material for forming a channel formation region, which serves as a switching element of a data holding portion, for the nonvolatile latch circuit according to this embodiment, a
10 nonvolatile latch circuit which has a wide operating temperature range and operates stably even at a high temperature and in which a logical state of storing data is not erased even after power is turned off or a latch circuit provided with a data holding portion where a refresh period is sufficiently long can be realized.

[0215]

15 Since data writing is performed by switching of the transistor, the number of rewrites is not substantially limited. In addition, the writing voltage is almost equivalent to the threshold voltage of the transistor; thus, the data writing can be performed at a low voltage. Further, a potential is directly supplied to the data holding portion; therefore, the variation in the amount of charge which is stored as data can be
20 suppressed small and data can be read easily.

[0216]

Various logic circuits can be provided by using the nonvolatile latch circuit. For example, the power consumption can be reduced by turning off power of the unused block. In addition, since a logical state is stored even when power is turned off, a
25 system can be started when power is turned on or terminated when power is turned off, at high speed and low power.

[0217]

The structures, methods, and the like described in this embodiment can be combined as appropriate with any of the structures, methods, and the like described in
30 the other embodiments.

[0218]

[Embodiment 2]

In this embodiment, another example of a structure, a manufacturing method, and the like of elements of the nonvolatile latch circuit which is an embodiment of the invention disclosed herein will be described with reference to FIG. 16, FIGS. 17A to 17E, and FIGS. 18A to 18D. In this embodiment, the configuration of the nonvolatile latch circuit is similar to that in FIG. 1.

[0219]

FIG. 16 is a cross-sectional view illustrating an example of structures of elements of a nonvolatile latch circuit. FIG. 16 is an example in which, among the elements of the nonvolatile latch circuit, the structure of a transistor 402 formed using an oxide semiconductor in the upper portion is different from that in FIGS. 3A and 3B. In other words, FIG. 16 is an example in which a top-gate transistor is applied to the transistor 402 formed using an oxide semiconductor in the upper portion. The structures of the other elements (the structure of the transistor in the lower portion, and the like) are similar to those in FIGS. 3A and 3B.

[0220]

<Structure of Elements of Nonvolatile Latch Circuit>

In FIG. 16, the transistor 160 using a material other than an oxide semiconductor is provided in a lower portion, and the transistor 402 using an oxide semiconductor is provided in an upper portion. The transistor 160 using a material other than an oxide semiconductor can be used as a transistor included in the first element (D1) 412 and the second element (D2) 413 of the latch portion. With the use of the material other than an oxide semiconductor, high-speed operation can be realized. Also other elements of the nonvolatile latch circuit can have a structure similar or comparable to that of the transistor 160.

[0221]

The element such as the capacitor 404 of the nonvolatile latch circuit can be formed using a conductive film, a semiconductor film, an insulating film, or the like included in the transistor 402 or the transistor 160. Note that the transistor 160 and the transistor 402 are n-channel transistors here; alternatively, a p-channel transistor may be used. It is easy to use a p-channel transistor as the transistor 160.

[0222]

The transistor 160 includes the channel formation region 116 provided in the

substrate 100 containing a semiconductor material; the impurity regions 114 and the high-concentration regions 120 (a combination of the impurity regions 114 and the high-concentration regions 120 can simply be referred to as impurity regions) formed with the channel formation region 116 interposed therebetween; the gate insulating layer 108a over the channel formation region 116; the gate electrode 110a over the gate insulating layer 108a; the source or drain electrode 130a electrically connected to the impurity region 114; and the source or drain electrode 130b electrically connected to the impurity region 114.

[0223]

Here, the sidewall insulating layers 118 are formed on the sides of the gate electrode 110a. Moreover, as seen from a plane, the high-concentration regions 120 are formed in a region of the substrate 100 which does not overlap with the sidewall insulating layers 118, and the metal compound regions 124 is in contact with the high-concentration regions 120. Further, the element isolation insulating layers 106 are formed over the substrate 100 so as to surround the transistor 160, and the interlayer insulating layer 126 and the interlayer insulating layer 128 are formed to cover the transistor 160.

[0224]

The source or drain electrode 130a and the source or drain electrode 130b are electrically connected to the metal compound regions 124 through the openings formed in the interlayer insulating layer 126 and the interlayer insulating layer 128. In other words, the source or drain electrode 130a and the source or drain electrode 130b are electrically connected to the high-concentration regions 120 and the impurity regions 114 through the metal compound regions 124.

[0225]

The transistor 402 includes an oxide semiconductor layer 140 provided over an insulating layer 168, a source or drain electrode 142a and a source or drain electrode 142b provided over the oxide semiconductor layer 140 and electrically connected to the oxide semiconductor layer 140, a gate insulating layer 166 provided to cover the oxide semiconductor layer 140, the source or drain electrode 142a, and the source or drain electrode 142b, and a gate electrode 178 provided over the gate insulating layer 166 in a region overlapping with the oxide semiconductor layer 140 (see FIG. 16).

[0226]

Here, the oxide semiconductor layer 140 is preferably an oxide semiconductor layer which is highly purified by sufficiently removing an impurity such as hydrogen and supplying oxygen. Specifically, the hydrogen concentration in the oxide semiconductor layer 140, which is measured by SIMS (secondary ion mass spectroscopy), can be set to less than or equal to $5 \times 10^{19}/\text{cm}^3$, preferably less than or equal to $5 \times 10^{18}/\text{cm}^3$, more preferably less than or equal to $5 \times 10^{17}/\text{cm}^3$, and still more preferably less than $1 \times 10^{16}/\text{cm}^3$.

[0227]

Note that in the oxide semiconductor layer 140 which is highly purified by sufficiently reducing the hydrogen concentration and supplying oxygen, the carrier concentration is sufficiently low (e.g., less than $1 \times 10^{12}/\text{cm}^3$, preferably less than $1 \times 10^{11}/\text{cm}^3$), as compared to carrier concentration (approximately $1 \times 10^{14}/\text{cm}^3$) in a general silicon wafer (a silicon wafer to which a slight amount of impurity elements such as phosphorus or boron is added).

[0228]

In this manner, by using an i-type or substantially i-type oxide semiconductor, the transistor 402 which has extremely favorable off-state current characteristics can be obtained. For example, when a drain voltage V_D is +1 V or +10 V and a gate voltage V_G ranges from -5 V to -20 V, the off-state current is less than or equal to 1×10^{-13} A at room temperature. Moreover, the aforementioned transistor has characteristics of a normally-off transistor. Therefore, leakage current, that is, an off-state current at a state where a voltage between a gate electrode and a source electrode is approximately 0 V is much smaller than that of a transistor using silicon. For example, the leakage current per unit channel width is less than or equal to 10 aA/ μm at room temperature.

[0229]

Further, in temperature characteristics, the off-state current can be sufficiently low and the on-state current can be sufficiently high even at a high temperature. For example, as V_G - I_D characteristics of the transistor 402, data is obtained in a range of -25 °C to 150 °C with low temperature dependence of off-state currents, on-state currents, mobilities, and S values. Furthermore, data is obtained which shows that the off-state

current in the aforementioned temperature range is as extremely low as 1×10^{-13} A or less. One of the reasons is that an i-type or substantially i-type oxide semiconductor which is obtained by sufficiently reducing the hydrogen concentration to be highly purified and has extremely low carrier concentration is used as the oxide semiconductor.

5 [0230]

When the oxide semiconductor layer 140 which is highly purified by sufficiently reducing the hydrogen concentration is used and the off-state current of the transistor 402 is reduced, a semiconductor device having a novel structure can be realized.

10 [0231]

In addition, over the transistor 402, an interlayer insulating layer 170 and an interlayer insulating layer 172 are provided. Here, the gate insulating layer 166, the interlayer insulating layer 170, and the interlayer insulating layer 172 are provided with openings that reach the source or drain electrode 142a and the source or drain electrode 142b. An electrode 154d and an electrode 154e are in contact with the source or drain electrode 142a and the source or drain electrode 142b through the openings. At the same time as formation of the electrode 154d and the electrode 154e, an electrode 154a, an electrode 154b, and an electrode 154c are formed which are in contact with the electrode 136a, the electrode 136b, and the electrode 136c, respectively, through openings in the gate insulating layer 166, the interlayer insulating layer 170, and the interlayer insulating layer 172.

[0232]

An insulating layer 156 is formed over the interlayer insulating layer 172. An electrode 158a, an electrode 158b, an electrode 158c, and an electrode 158d are formed so as to be embedded in the insulating layer 156. Here, the electrode 158a is in contact with the electrode 154a, the electrode 158b is in contact with the electrode 154b, the electrode 158c is in contact with the electrode 154c and the electrode 154d, and the electrode 158d is in contact with the electrode 154e.

[0233]

30 That is, the source or drain electrode 142a of the transistor 402 is electrically connected to another element (e.g., the transistor using a material other than an oxide semiconductor) through the electrode 130c, the electrode 136c, the electrode 154c, the

electrode 158c, and the electrode 154d (see FIG. 16). In addition, the source or drain electrode 142b of the transistor 402 is electrically connected to another element through the electrode 154e and the electrode 158d. Note that the structure of connection electrodes (such as the electrode 130c, the electrode 136c, the electrode 154c, the electrode 158c, and the electrode 154d) is not limited to the aforementioned structure, and appropriate addition, omission, or the like is possible.

[0234]

<Manufacturing Method of Elements of Nonvolatile Latch Circuit>

Next, an example of a manufacturing method of the elements of the nonvolatile latch circuit will be described. With the manufacturing method described below, the elements of the nonvolatile latch circuit can be manufactured. Note that the manufacturing method of the transistor 160 is similar to that in FIGS. 4A to 4H; therefore, description thereof is omitted. The manufacturing method of the transistor 402 will be described with reference to FIGS. 17A to 17E and FIGS. 18A to 18D.

[0235]

<Manufacturing Method of Transistor in Upper Portion>

Next, steps for manufacturing the transistor 402 over the interlayer insulating layer 128 will be described with reference to FIGS. 17A to 17E and FIGS. 18A to 18D. Note that FIGS. 17A to 17E and FIGS. 18A to 18D illustrate steps for manufacturing electrodes, the transistor 402, and the like over the interlayer insulating layer 128; therefore, the transistor 160 and the like placed below the transistor 402 are omitted.

[0236]

First, the insulating layer 132 is formed over the interlayer insulating layer 128, the source or drain electrode 130a, the source or drain electrode 130b, and the electrode 130c. Next, openings that reach the source or drain electrode 130a, the source or drain electrode 130b, and the electrode 130c are formed in the insulating layer 132. Then, a conductive layer is formed to be embedded in the openings. After that, part of the conductive layer is removed by etching, CMP, or the like, so that the insulating layer 132 is exposed and the electrode 136a, the electrode 136b, and the electrode 136c are formed (see FIG. 17A).

[0237]

The insulating layer 132 can be formed with a PVD method, a CVD method, or

the like. The insulating layer 132 can be formed using a material containing an inorganic insulating material such as silicon oxide, silicon nitride oxide, silicon nitride, hafnium oxide, aluminum oxide, or tantalum oxide.

[0238]

5 The openings can be formed in the insulating layer 132 with a method such as etching using a mask. The mask can be formed with a method such as light exposure using a photomask. Either wet etching or dry etching may be used as the etching; dry etching is preferably used in terms of microfabrication.

[0239]

10 The conductive layer can be formed with a film formation method such as a PVD method or a CVD method. The conductive layer can be formed using a conductive material such as molybdenum, titanium, chromium, tantalum, tungsten, aluminum, copper, neodymium, or scandium or an alloy or a compound (e.g., a nitride) of any of these materials, for example.

15 [0240]

More specifically, it is possible to employ a method, for example, in which a thin titanium film is formed in a region including the openings with a PVD method and a thin titanium nitride film is formed with a CVD method, and then, a tungsten film is formed to be embedded in the openings. Here, the titanium film formed with a PVD
20 method has a function of reducing an oxide film which might be formed on the surface of lower electrodes (here, the source or drain electrode 130a, the source or drain electrode 130b, the electrode 130c, and the like) to decrease the contact resistance with the lower electrodes.

[0241]

25 The titanium nitride film formed after the formation of the titanium film has a barrier function of preventing diffusion of the conductive material. A copper film may be formed with a plating method after the formation of the barrier film of titanium, titanium nitride, or the like. Note that not only a so-called single damascene method but also a dual damascene method may be employed.

30 [0242]

When the electrode 136a, the electrode 136b, and the electrode 136c are formed, the process is preferably performed by CMP or the like so that the surfaces are

planarized. When the surfaces of the insulating layer 132, the electrode 136a, the electrode 136b, and the electrode 136c are planarized in such a manner, an electrode, a wiring, an insulating layer, a semiconductor layer, and the like can be favorably formed in later steps.

5 [0243]

Next, the insulating layer 168 is formed to cover the insulating layer 132, the electrode 136a, the electrode 136b, and the electrode 136c. Next, an oxide semiconductor layer is formed over the insulating layer 168 and processed with a method such as etching using a mask, so that the island-shaped oxide semiconductor layer 140 is formed (see FIG. 17B).

10 [0244]

The insulating layer 168 functions as a base and can be formed with a CVD method, a sputtering method, or the like. The insulating layer 168 is preferably formed using silicon oxide, silicon nitride, silicon oxynitride, silicon nitride oxide, aluminum oxide, hafnium oxide, tantalum oxide, or the like. Note that the insulating layer 168 may have a single-layer structure or a stacked structure. There is no particular limitation on the thickness of the insulating layer 168; the insulating layer 168 can be formed in the range of 10 nm to 500 nm inclusive, for example. Here, the insulating layer 168 is not an essential component; therefore, a structure in which the insulating layer 168 is not provided is also possible.

20 [0245]

If hydrogen, water, or the like is contained in the insulating layer 168, hydrogen may enter the oxide semiconductor layer or extract oxygen from the oxide semiconductor layer, whereby characteristics of the transistor might be degraded. Therefore, it is preferable to form the insulating layer 168 so as to contain hydrogen or water as little as possible.

25 [0246]

In the case where a sputtering method or the like is employed, for example, it is preferable that the insulating layer 168 be formed in a state where moisture in the treatment chamber is removed. In order to remove moisture in the treatment chamber, an entrapment vacuum pump such as a cryopump, an ion pump, or a titanium sublimation pump is preferably used. Alternatively, a turbo pump provided with a cold

trap may be used. Since hydrogen, water, or the like is sufficiently removed from the treatment chamber evacuated with a cryopump or the like, the concentration of an impurity contained in the insulating layer 168 can be reduced.

[0247]

5 When the insulating layer 168 is formed, it is preferable to use a high-purity gas in which an impurity such as hydrogen or water is reduced to a concentration of a few ppm or less (preferably, 10 ppb or less).

[0248]

10 As the oxide semiconductor layer, any of the following oxide semiconductors can be used: an In-Sn-Ga-Zn-O-based oxide semiconductor which is a four-component metal oxide; an In-Ga-Zn-O-based oxide semiconductor, an In-Sn-Zn-O-based oxide semiconductor, an In-Al-Zn-O-based oxide semiconductor, a Sn-Ga-Zn-O-based oxide semiconductor, an Al-Ga-Zn-O-based oxide semiconductor, or a Sn-Al-Zn-O-based oxide semiconductor which are three-component metal oxides; an In-Zn-O-based oxide
15 semiconductor, a Sn-Zn-O-based oxide semiconductor, an Al-Zn-O-based oxide semiconductor, a Zn-Mg-O-based oxide semiconductor, a Sn-Mg-O-based oxide semiconductor, or an In-Mg-O-based oxide semiconductor which are two-component metal oxides; or an In-O-based oxide semiconductor; a Sn-O-based oxide semiconductor; or a Zn-O-based oxide semiconductor which are one-component metal
20 oxides. In addition, the aforementioned oxide semiconductors may contain SiO₂.

[0249]

As the oxide semiconductor layer, a thin film containing a material represented by InMO₃(ZnO)_m ($m > 0$) can be used. Here, M represents one or more of metal elements selected from Ga, Al, Mn, and Co. For example, M can be Ga, Ga and Al,
25 Ga and Mn, Ga and Co, or the like.

[0250]

In this embodiment, as the oxide semiconductor layer, an amorphous oxide semiconductor layer is formed with a sputtering method using an In-Ga-Zn-O-based metal oxide target. Note that since crystallization of an amorphous oxide
30 semiconductor layer can be suppressed by adding silicon to the amorphous oxide semiconductor layer, an oxide semiconductor layer may be formed, for example, using a

target containing SiO₂ of 2 wt% to 10 wt% inclusive.

[0251]

As the metal oxide target used for forming an oxide semiconductor layer with a sputtering method, for example, a metal oxide target having a composition ratio of
5 In₂O₃:Ga₂O₃:ZnO = 1:1:1 [molar ratio] can be used. Furthermore, a metal oxide target having a composition ratio of In₂O₃:Ga₂O₃:ZnO = 1:1:2 [molar ratio] or a metal oxide target having a composition ratio of In₂O₃:Ga₂O₃:ZnO = 1:1:4 [molar ratio] can also be used. The filling rate of a metal oxide target is 90 % to 100 % inclusive, preferably greater than or equal to 95 % (e.g., 99.9 %). A dense oxide semiconductor layer is
10 formed by using a metal oxide target with a high filling rate.

[0252]

The atmosphere in which the oxide semiconductor layer is formed is preferably a rare gas (typically argon) atmosphere, an oxygen atmosphere, or a mixed atmosphere containing a rare gas (typically argon) and oxygen. Specifically, it is preferable to use,
15 for example, a high-purity gas atmosphere from which an impurity such as hydrogen, water, a hydroxyl group, or hydride is removed to a concentration of a few ppm or less (preferably, a few ppb or less).

[0253]

At the time of forming the oxide semiconductor layer, the substrate is held in a
20 treatment chamber kept under reduced pressure and the substrate is heated to a temperature of 100 °C to 600 °C inclusive, preferably 200 °C to 400 °C inclusive. Then, a sputtering gas from which hydrogen and water are removed is introduced into the treatment chamber while moisture in the treatment chamber is removed, whereby the oxide semiconductor layer is formed using the metal oxide as a target. The oxide
25 semiconductor layer is formed while the substrate is heated, so that the concentration of an impurity contained in the oxide semiconductor layer can be reduced. Moreover, damage of the oxide semiconductor layer due to sputtering is reduced.

[0254]

An entrapment vacuum pump is preferably used in order to remove moisture in
30 the treatment chamber. For example, a cryopump, an ion pump, or a titanium sublimation pump can be used. Alternatively, a turbo pump provided with a cold trap

may be used. Since hydrogen, water, or the like is removed from the treatment chamber evacuated with a cryopump, the concentration of an impurity in the oxide semiconductor layer can be reduced.

[0255]

5 The oxide semiconductor layer can be formed under the following conditions, for example: the distance between the substrate and the target is 100 mm; the pressure is 0.6 Pa; the direct-current (DC) power supply is 0.5 kW; and the atmosphere is oxygen (the proportion of the oxygen flow is 100 %). Note that it is preferable to use a pulsed direct-current (DC) power source because dust can be reduced and the thickness
10 distribution is can be reduced. The thickness of the oxide semiconductor layer is 2 nm to 200 nm inclusive, preferably 5 nm to 30 nm inclusive. Note that the appropriate thickness of the oxide semiconductor layer differs depending on the oxide semiconductor material to be used, the intended purpose of a semiconductor device, or the like; therefore, the thickness may be determined in accordance with the material to
15 be used, the intended purpose, or the like.

[0256]

 Note that before the oxide semiconductor layer is formed with a sputtering method, reverse sputtering is preferably performed in which plasma is generated with an argon gas introduced, so that dust on the surface of the insulating layer 168 is removed.
20 Here, the reverse sputtering is a method in which ions collide with a surface to be processed so that the surface is modified, in contrast to normal sputtering in which ions collide with a sputtering target. An example of a method for making ions collide with a surface to be processed is a method in which a high-frequency voltage is applied to the surface to be processed under an argon atmosphere so that plasma is generated near a
25 substrate. Note that an atmosphere of nitrogen, helium, oxygen, or the like may be used instead of an argon atmosphere.

[0257]

 As an etching method for the oxide semiconductor layer, either dry etching or wet etching may be employed. It is needless to say that dry etching and wet etching
30 can be used in combination. The etching conditions (e.g., an etching gas or an etchant, etching time, and temperature) are set as appropriate depending on the material so that the oxide semiconductor layer can be etched into a desired shape.

[0258]

As the dry etching method, a parallel plate RIE (reactive ion etching) method or an ICP (inductively coupled plasma) etching method can be used. Also in this case, it is necessary to set etching conditions (e.g., the amount of electric power applied to a coiled electrode, the amount of electric power applied to an electrode on the substrate side, and the electrode temperature on the substrate side) as appropriate.

[0259]

An example of an etching gas used for dry etching is a gas containing chlorine (a chlorine-based gas such as chlorine (Cl_2), boron trichloride (BCl_3), silicon tetrachloride (SiCl_4), or carbon tetrachloride (CCl_4)) or the like. Moreover, a gas containing fluorine (a fluorine-based gas such as carbon tetrafluoride (CF_4), sulfur hexafluoride (SF_6), nitrogen trifluoride (NF_3), or trifluoromethane (CHF_3)), hydrogen bromide (HBr), oxygen (O_2), any of these gases to which a rare gas such as helium (He) or argon (Ar) is added, or the like may be used.

15 [0260]

As an etchant used for wet etching, a mixed solution of phosphoric acid, acetic acid, and nitric acid; ammonia peroxide mixture (a mixture of ammonia, water, and a hydrogen peroxide solution); or the like can be used. An etchant such as ITO07N (produced by KANTO CHEMICAL CO., INC.) may also be used.

20 [0261]

Then, first heat treatment is preferably performed on the oxide semiconductor layer. Water (including a hydroxyl group), hydrogen, or the like contained in the oxide semiconductor layer can be removed through the first heat treatment. The temperature of the first heat treatment is 300 °C to 800 °C inclusive, preferably 400 °C to 700 °C inclusive, more preferably 450 °C to 700 °C inclusive, and still more preferably 550 °C to 700 °C inclusive.

[0262]

The oxide semiconductor layer can be dehydrated or dehydrogenated when the first heat treatment is performed at a temperature of 350 °C or higher, so that the hydrogen concentration in the oxide semiconductor layer can be reduced. In addition, when the first heat treatment is performed at a temperature of 450 °C or higher, the

hydrogen concentration in the oxide semiconductor layer can be further reduced. Moreover, when the first heat treatment is performed at a temperature of 550 °C or higher, the hydrogen concentration in the oxide semiconductor layer can be still further reduced. For example, the substrate is introduced into an electric furnace in which a
5 resistance heating element or the like is used and the oxide semiconductor layer 140 is subjected to heat treatment at 450 °C for 1 hour under a nitrogen atmosphere. The oxide semiconductor layer 140 is not exposed to the air during the heat treatment so that the entry of water or hydrogen can be prevented.

[0263]

10 The heat treatment apparatus is not limited to the electric furnace and can be an apparatus for heating an object to be processed by thermal conduction or thermal radiation from a medium such as a heated gas. For example, a rapid thermal annealing (RTA) apparatus such as a gas rapid thermal annealing (GRTA) apparatus or a lamp
rapid thermal annealing (LRTA) apparatus can be used.

15 [0264]

An LRTA apparatus is an apparatus for heating an object to be processed by radiation of light (an electromagnetic wave) emitted from a lamp such as a halogen lamp, a metal halide lamp, a xenon arc lamp, a carbon arc lamp, a high-pressure sodium lamp, or a high-pressure mercury lamp. A GRTA apparatus is an apparatus for
20 performing heat treatment using a high-temperature gas. As the gas, an inert gas which does not react with an object to be processed by heat treatment, for example, nitrogen or a rare gas such as argon is used.

[0265]

For example, as the first heat treatment, a GRTA process may be performed as
25 follows. The substrate is put in an inert gas atmosphere that has been heated to a high temperature of 650 °C to 700 °C, heated for several minutes, and taken out of the inert gas atmosphere. The GRTA process enables high-temperature heat treatment for a short time. Moreover, the GRTA process can be employed even when the temperature exceeds the upper temperature limit of the substrate because it is heat treatment for a
30 short time. For example, in the case where an SOI substrate including a substrate with relatively low heat resistance, such as a glass substrate, is used, shrinkage of the

substrate becomes a problem at a temperature higher than the upper temperature limit (strain point) but does not in the case where heat treatment is performed for a short time.

[0266]

Note that as the inert gas atmosphere under which the first heat treatment is performed, it is preferable to employ an atmosphere that contains nitrogen or a rare gas (e.g., helium, neon, or argon) as its main component and does not contain water, hydrogen, or the like. For example, the purity of nitrogen or a rare gas such as helium, neon, or argon introduced into the heat treatment apparatus is greater than or equal to 6 N (99.9999 %), preferably greater than or equal to 7 N (99.99999 %) (i.e., the impurity concentration is less than or equal to 1 ppm, preferably less than or equal to 0.1 ppm).

[0267]

Note that the inert gas atmosphere may be changed during the process to an atmosphere including oxygen. For example, in the case where an electrical furnace is used in the first heat treatment, an atmosphere can be changed when a heat treatment temperature falls. For example, the heat treatment can be performed (at a constant temperature) under an atmosphere of an inert gas such as a rare gas (e.g., helium, neon, or argon) or nitrogen, and the atmosphere can be switched to an atmosphere containing oxygen when the heat treatment temperature falls. As the atmosphere containing oxygen, an oxygen gas or a mixed gas of an oxygen gas and a nitrogen gas can be used.

[0268]

Also in the case where the atmosphere containing oxygen is employed, it is preferable that the atmosphere do not contain water, hydrogen, or the like. Alternatively, the purity of the oxygen gas or the nitrogen used is preferably greater than or equal to 6N (99.9999 %), more preferably greater than or equal to 7N (99.99999 %) (that is, the impurity concentration is less than or equal to 1 ppm, preferably less than or equal to 0.1 ppm). This is because defects caused by oxygen deficiency can be reduced by performing the first heat treatment under an atmosphere containing oxygen.

[0269]

In some cases, the oxide semiconductor layer is crystallized to be microcrystalline or polycrystalline, which depends on the conditions of the first heat treatment or the material of the oxide semiconductor layer. For example, in some

cases, the oxide semiconductor layer becomes a microcrystalline oxide semiconductor layer having a degree of crystallization of 90 % or more, or 80 % or more. Further, in some cases, the oxide semiconductor layer may be an amorphous oxide semiconductor layer containing no crystalline component, which depends on the conditions of the first heat treatment or the material of the oxide semiconductor layer.

[0270]

Furthermore, in some cases, the oxide semiconductor layer becomes a layer in which a microcrystal (with a grain size of 1 nm to 20 nm inclusive, typically 2 nm to 4 nm inclusive) is mixed in an amorphous oxide semiconductor (e.g., a surface of the oxide semiconductor layer). The electric characteristics of the oxide semiconductor layer can be changed by aligning microcrystals in an amorphous semiconductor in the aforementioned manner.

[0271]

For example, when the oxide semiconductor layer is formed using an In-Ga-Zn-O-based metal oxide target, the electric characteristics of the oxide semiconductor layer can be changed by formation of a microcrystalline region in which crystal grains of $\text{In}_2\text{Ga}_2\text{ZnO}_7$ with electrical anisotropy are aligned. The microcrystalline region is preferably a region in which the crystal grains are arranged so that the c-axis of $\text{In}_2\text{Ga}_2\text{ZnO}_7$ is perpendicular to a surface of the oxide semiconductor layer, for example.

[0272]

By forming a region in which crystal grains are arranged in such a manner, the conductivity in the direction parallel to the surface of the oxide semiconductor layer can be improved and insulating properties in the direction perpendicular to the surface of the oxide semiconductor layer can be improved. Furthermore, such a microcrystalline region has a function of suppressing entry of an impurity such as water or hydrogen into the oxide semiconductor layer.

[0273]

Note that the oxide semiconductor layer including the microcrystalline region can be formed by heating the surface of the oxide semiconductor layer by a GRTA process. Further, the oxide semiconductor layer can be formed in a more preferred manner by using a sputtering target in which the amount of Zn is smaller than that of In

or Ga.

[0274]

The first heat treatment for the oxide semiconductor layer 140 can be performed on the oxide semiconductor layer that has not yet been processed into the island-shaped oxide semiconductor layer 140. In that case, after the first heat treatment, the substrate is taken out of the heating apparatus and a photolithography step is performed.

[0275]

Note that the aforementioned first heat treatment can also be referred to as dehydration treatment, dehydrogenation treatment, or the like. The dehydration treatment or dehydrogenation treatment can be performed, for example, after the oxide semiconductor layer is formed, after a source or drain electrode is stacked over the oxide semiconductor layer 140, or after a gate insulating layer is formed over the source or drain electrode. Such dehydration treatment or dehydrogenation treatment may be performed once or plural times.

[0276]

Next, after a conductive layer 142 is formed so as to be in contact with the oxide semiconductor layer 140, an insulating layer 164 is formed over the conductive layer 142 (see FIG. 17C). Note that the insulating layer 164 is not necessarily formed.

[0277]

The conductive layer 142 can be formed with a PVD method such as a sputtering method or a CVD method such as a plasma CVD method. As a material for the conductive layer 142, an element selected from aluminum, chromium, copper, tantalum, titanium, molybdenum, and tungsten; an alloy containing any of these elements as its component; or the like can be used. Moreover, one or more of materials selected from manganese, magnesium, zirconium, beryllium, and yttrium may be used. It is also possible to use aluminum combined with one or more of elements selected from titanium, tantalum, tungsten, molybdenum, chromium, neodymium, and scandium.

[0278]

The conductive layer 142 may be formed using conductive metal oxide. As the conductive metal oxide, indium oxide (In_2O_3), tin oxide (SnO_2), zinc oxide (ZnO),

an indium oxide-tin oxide alloy ($\text{In}_2\text{O}_3\text{-SnO}_2$, which is abbreviated to ITO in some cases), an indium oxide-zinc oxide alloy ($\text{In}_2\text{O}_3\text{-ZnO}$), or any of these metal oxide materials in which silicon or silicon oxide is included can be used.

[0279]

5 The conductive layer 142 can have a single-layer structure or a stacked structure including two or more layers. For example, the conductive layer 142 can have a single-layer structure of an aluminum film containing silicon, a two-layer structure in which a titanium film is stacked over an aluminum film, or a three-layer structure in which a titanium film, an aluminum film, and a titanium film are stacked in
10 this order. Here, a three-layer structure of a titanium film, an aluminum film, and a titanium film is employed.

[0280]

Note that an oxide conductive layer may be formed between the oxide semiconductor layer 140 and the conductive layer 142. The oxide conductive layer
15 and the conductive layer 142 can be successively formed. By providing such an oxide conductive layer, the resistance of the source region or the drain region can be reduced, so that the transistor can operate at high speed.

[0281]

The insulating layer 164 can be formed with a CVD method, a sputtering
20 method, or the like. The insulating layer 164 is preferably formed using silicon oxide, silicon nitride, silicon oxynitride, silicon nitride oxide, aluminum oxide, hafnium oxide, tantalum oxide, or the like. Note that the insulating layer 164 may have a single-layer structure or a stacked structure. There is no particular limitation on the thickness of the insulating layer 164; the insulating layer 164 can be formed in the range of 10 nm to
25 500 nm inclusive, for example.

[0282]

Next, the source or drain electrode 142a, the source or drain electrode 142b, an insulating layer 164a, and an insulating layer 164b are formed by selectively etching the conductive layer 142 and the insulating layer 164 (see FIG. 17D).

30 [0283]

Ultraviolet light, KrF laser light, or ArF laser light is preferably used for light exposure at the time of forming a mask used for etching. In particular, in the case

where light exposure is performed so that the channel length (L) is less than 25 nm, light exposure for forming a mask is preferably performed with extreme ultraviolet rays whose wavelength is extremely short of several nanometers to several tens of nanometers. The resolution of light exposure with extreme ultraviolet rays is high and the depth of focus is large. For these reasons, it is possible to design a mask so that the channel length (L) of the transistor to be formed later is less than 25 nm, that is, in the range of 10 nm to 1000 nm inclusive. By a decrease in channel length with such a method, operation speed can be improved. In addition, the off-state current of the transistor using an oxide semiconductor is small; thus, an increase in power consumption due to miniaturization can be suppressed.

[0284]

The materials and etching conditions of the conductive layer 142 and the oxide semiconductor layer 140 are adjusted as appropriate so that the oxide semiconductor layer 140 is not removed in etching of the conductive layer 142. Note that in some cases, the oxide semiconductor layer 140 is partly etched in the etching step and thus has a groove portion (a recessed portion) depending on the materials and the etching conditions.

[0285]

In order to reduce the number of masks to be used and reduce the number of steps, an etching step may be performed with the use of a resist mask formed using a multi-tone mask which is a light-exposure mask through which light is transmitted to have a plurality of intensities. A resist mask formed with the use of a multi-tone mask has a plurality of thicknesses (has a stair-like shape) and further can be changed in shape by ashing; therefore, the resist mask can be used in a plurality of etching steps for processing into different patterns. That is, a resist mask corresponding to at least two kinds of different patterns can be formed by using a multi-tone mask. Thus, the number of light-exposure masks can be reduced and the number of corresponding photolithography steps can also be reduced, whereby a process can be simplified.

[0286]

Next, the gate insulating layer 166 is formed in contact with part of the oxide semiconductor layer 140 without exposure to the air (see FIG. 17E). The gate insulating layer 166 can be formed with a CVD method, a sputtering method, or the like.

The gate insulating layer 166 is preferably formed using silicon oxide, silicon nitride, silicon oxynitride, silicon nitride oxide, aluminum oxide, hafnium oxide, tantalum oxide, or the like. Note that the gate insulating layer 166 may have a single-layer structure or a stacked structure. There is no particular limitation on the thickness of the gate insulating layer 166; the gate insulating layer 166 can be formed in the range of 10 nm to 500 nm inclusive, for example.

[0287]

Note that an i-type or substantially i-type oxide semiconductor obtained by removing an impurity (a purified oxide semiconductor) is highly susceptible to interface states or interface charges; therefore, the gate insulating layer 166 needs to have high quality.

[0288]

For example, the gate insulating layer 166 is preferably formed with a high-density plasma CVD method using a microwave (frequency: 2.45 GHz) because the gate insulating layer 166 can be dense and have high withstand voltage and high quality. When a highly purified oxide semiconductor layer and a high-quality gate insulating layer are in close contact with each other, the interface level can be reduced and favorable interface characteristics can be obtained.

[0289]

It is needless to say that another method such as a sputtering method or a plasma CVD method can be employed as long as a high-quality insulating layer can be formed as the gate insulating layer 166. Moreover, it is possible to use an insulating layer whose film quality and interface characteristics are improved with heat treatment performed after the formation of the gate insulating layer 166. In any case, an insulating layer that has favorable film quality as the gate insulating layer 166 and can reduce interface level density with an oxide semiconductor layer to form a favorable interface is formed as the gate insulating layer 166.

[0290]

By thus improving characteristics of the interface with a gate insulating layer and eliminating an impurity, particularly hydrogen, water, or the like, from an oxide semiconductor, it is possible to obtain a stable transistor whose threshold voltage (V_{th}) does not change with a gate bias-temperature stress test (BT test, e.g., at 85 °C and $2 \times$

10⁶ V/cm for 12 hours).

[0291]

After that, second heat treatment is performed under an inert gas atmosphere or an oxygen atmosphere. The heat treatment is performed at a temperature of 200 °C to 400 °C inclusive, preferably 250 °C to 350 °C inclusive. For example, the second heat treatment may be performed at 250 °C for 1 hour under a nitrogen atmosphere. The second heat treatment can reduce variation in electric characteristics of the transistor. Note that although the second heat treatment is performed in this embodiment after the gate insulating layer 166 is formed, there is no particular limitation on the timing of the second heat treatment as long as it is performed after the first heat treatment.

[0292]

Next, the gate electrode 178 is formed over the gate insulating layer 166 in a region overlapping with the oxide semiconductor layer 140 (see FIG. 18A). The gate electrode 178 can be formed by forming a conductive layer over the gate insulating layer 166 and then selectively patterning the conductive layer.

[0293]

The conductive layer can be formed with a PVD method such as a sputtering method or a CVD method such as a plasma CVD method. As a material for the conductive layer, an element selected from aluminum, chromium, copper, tantalum, titanium, molybdenum, and tungsten; an alloy containing any of these elements as its component; or the like can be used. Moreover, one or more of materials selected from manganese, magnesium, zirconium, beryllium, and yttrium may be used. It is also possible to use aluminum combined with one or more of elements selected from titanium, tantalum, tungsten, molybdenum, chromium, neodymium, and scandium.

[0294]

The conductive layer may be formed using conductive metal oxide. As the conductive metal oxide, indium oxide (In₂O₃), tin oxide (SnO₂), zinc oxide (ZnO), an indium oxide-tin oxide alloy (In₂O₃-SnO₂, which is abbreviated to ITO in some cases), an indium oxide-zinc oxide alloy (In₂O₃-ZnO), or any of these metal oxide materials in which silicon or silicon oxide is included can be used.

[0295]

The conductive layer can have a single-layer structure or a stacked structure including two or more layers. For example, the conductive layer can have a single-layer structure of an aluminum film containing silicon, a two-layer structure in which a titanium film is stacked over an aluminum film, or a three-layer structure in which a titanium film, an aluminum film, and a titanium film are stacked in this order. Here, the conductive layer is formed using a material including titanium and then processed into the gate electrode 178.

[0296]

Next, the interlayer insulating layer 170 and the interlayer insulating layer 172 are formed over the gate insulating layer 166 and the gate electrode 178 (see FIG. 18B). The interlayer insulating layer 170 and the interlayer insulating layer 172 can be formed with a PVD method, a CVD method, or the like. The interlayer insulating layer 170 and the interlayer insulating layer 172 can be formed using a material including an inorganic insulating material such as silicon oxide, silicon nitride oxide, silicon nitride, hafnium oxide, aluminum oxide, or tantalum oxide. Note that although a stacked structure of the interlayer insulating layer 170 and the interlayer insulating layer 172 is used in this embodiment, an embodiment of the invention disclosed herein is not limited thereto. A single-layer structure or a stacked structure including three or more layers can also be used.

[0297]

Note that the interlayer insulating layer 172 is preferably formed so as to have a planarized surface. This is because an electrode, a wiring, or the like can be favorably formed over the interlayer insulating layer 172 when the interlayer insulating layer 172 is formed so as to have a planarized surface.

[0298]

Next, openings that reach the electrode 136a, the electrode 136b, the electrode 136c, the source or drain electrode 142a, and the source or drain electrode 142b are formed in the gate insulating layer 166, the interlayer insulating layer 170, and the interlayer insulating layer 172. Then, a conductive layer is formed so as to be embedded in the openings. Then, part of the conductive layer is removed with a method such as etching or CMP, so that the interlayer insulating layer 172 is exposed and the electrode 154a, the electrode 154b, the electrode 154c, the electrode 154d, and

the electrode 154e are formed (see FIG. 18C).

[0299]

The openings can be formed with a method such as etching using a mask. The mask can be formed with a method such as light exposure using a photomask.
5 Either wet etching or dry etching may be used as the etching; dry etching is preferably used in terms of microfabrication.

[0300]

The conductive layer can be formed with a film formation method such as a PVD method or a CVD method. The conductive layer can be formed using a
10 conductive material such as molybdenum, titanium, chromium, tantalum, tungsten, aluminum, copper, neodymium, or scandium or an alloy or a compound (e.g., a nitride) of any of these materials, for example.

[0301]

Specifically, it is possible to employ a method, for example, in which a thin
15 titanium film is formed in a region including the openings with a PVD method and a thin titanium nitride film is formed with a CVD method, and then, a tungsten film is formed to be embedded in the openings. Here, the titanium film formed with a PVD method has a function of reducing an oxide film which might be formed on the surface of lower electrodes (here, the electrode 136a, the electrode 136b, the electrode 136c, the
20 source or drain electrode 142a, the source or drain electrode 142b, and the like) to decrease the contact resistance with the lower electrodes. The titanium nitride film formed after the formation of the titanium film has a barrier function of preventing diffusion of the conductive material. A copper film may be formed with a plating method after the formation of the barrier film of titanium, titanium nitride, or the like.
25 Note that not only a so-called single damascene method but also a dual damascene method may be employed.

[0302]

When part of the conductive layer is removed, the process is preferably performed so that the exposed surface of the interlayer insulating layer 172; the surfaces
30 of the electrode 154a, the electrode 154b, the electrode 154c, the electrode 154d, and the electrode 154e; and the like are planarized. The surfaces are planarized in this manner, whereby an electrode, a wiring, or the like can be favorably formed in a later step.

[0303]

Then, the insulating layer 156 is further formed, and openings that reach the electrode 154a, the electrode 154b, the electrode 154c, the electrode 154d, and the electrode 154e are formed in the insulating layer 156. After a conductive layer is formed to be embedded in the openings, part of the conductive layer is removed with a method such as etching or CMP. Thus, the insulating layer 156 is exposed and the electrode 158a, the electrode 158b, the electrode 158c, and the electrode 158d are formed (see FIG. 18D). This step is similar to the step of forming the electrode 154a and the like; therefore, the detailed description is omitted.

[0304]

In the case where the transistor 402 is formed with the aforementioned method, the hydrogen concentration in the oxide semiconductor layer 140 is $5 \times 10^{19} / \text{cm}^3$ or less and the off-state current of the transistor 402 is 1×10^{-13} A or less. The transistor 402 with excellent characteristics can be obtained by the application of the oxide semiconductor layer 140 that is highly purified by sufficiently reducing the hydrogen concentration and supplying oxygen as described above.

[0305]

Note that it is preferable that oxygen be supplied to the oxide semiconductor layer 140 shortly after the hydrogen concentration is reduced because there is no possibility that hydrogen, water, or the like enters the oxide semiconductor layer and thus an oxide semiconductor layer with extremely favorable characteristics can be realized. It is needless to say that treatment for reducing the hydrogen concentration and treatment for supplying oxygen do not need to be performed successively as long as an oxide semiconductor layer with favorable characteristics can be realized. For example, another treatment may be performed between both of the treatment. Alternatively, both of the treatment may be performed at the same time.

[0306]

Since the nonvolatile latch circuit includes the transistor 160 formed using a material other than an oxide semiconductor in the lower portion and the transistor 402 formed using an oxide semiconductor in the upper portion, it is possible to manufacture an excellent nonvolatile latch circuit having characteristics of both the transistors, and a

semiconductor device using the nonvolatile latch circuit.

[0307]

Although a lot of researches on properties of an oxide semiconductor such as density of state (DOS) have been conducted, they do not include the idea of sufficiently
5 reducing localized states themselves. According to an embodiment of the invention disclosed herein, a highly purified oxide semiconductor is formed by removing water or hydrogen which might affect the localized states. This is based on the idea that the localized states themselves are sufficiently reduced. Such a highly purified oxide semiconductor enables fabrication of very excellent industrial products.

10 [0308]

Note that when hydrogen, water, or the like is removed, oxygen is also removed in some cases. Therefore, it is preferable to obtain more highly purified (i-type) oxide semiconductor by supplying oxygen to dangling bonds of metal which are generated by oxygen deficiency so that localized states resulted from oxygen deficiency
15 are reduced. For example, localized states resulted from oxygen deficiency can be reduced in the following manner: an oxide film having excessive oxygen is formed in a close contact with a channel formation region; and heat treatment at 200 °C to 400 °C, typically approximately 250 °C, is performed so that oxygen is supplied to an oxide semiconductor from the oxide film. An inert gas may be switched to a gas including
20 oxygen during the second heat treatment. In addition, the aforementioned atmosphere may be switched; the oxide semiconductor layer can also be supplied with oxygen by being subjected to a temperature fall process in an oxygen atmosphere or an atmosphere from which hydrogen or water is sufficiently removed, successively after the second heat treatment.

25 [0309]

A defect of an oxide semiconductor is said to be attributed to a level of 0.1 eV to 0.2 eV under the conduction band due to excessive hydrogen, a deep level due to shortage of oxygen, or the like. Thorough removal of hydrogen and sufficient supply of oxygen for elimination of such a defect would be right as a technological thought.

30 [0310]

An oxide semiconductor is generally considered as an n-type semiconductor;

however, according to an embodiment of the invention disclosed herein, an i-type semiconductor is realized by removing an impurity such as water and hydrogen and supplying oxygen that is a component of an oxide semiconductor. In this respect, it can be said that an embodiment of the invention disclosed herein includes a novel technical idea because it is different from an i-type semiconductor such as silicon added with an impurity.

[0311]

With the use of a transistor using an oxide semiconductor as a semiconductor material for forming a channel formation region, which serves as a switching element of a data holding portion, for the nonvolatile latch circuit according to this embodiment, a nonvolatile latch circuit which has a wide operating temperature range and operates stably even at a high temperature and in which a logical state of storing data is not erased even after power is turned off or a latch circuit provided with a data holding portion where a refresh period is sufficiently long can be realized. Since data writing is performed by switching of the transistor, the number of rewrites is not substantially limited. In addition, the writing voltage is almost equivalent to the threshold voltage of the transistor; thus, the data writing can be performed at a low voltage. Further, a potential is directly supplied to the data holding portion; therefore, the variation in the amount of charge which is stored as data can be suppressed small and data can be read easily.

[0312]

Various logic circuits can be provided by using the nonvolatile latch circuit. For example, the power consumption can be reduced by turning off power of the unused block. In addition, since a logical state is stored even when power is turned off, a system can be started when power is turned on or terminated when power is turned off, at high speed and low power.

[0313]

The structures, methods, and the like described in this embodiment can be combined as appropriate with any of the structures, methods, and the like described in the other embodiments.

[0314]

[Embodiment 3]

In this embodiment, the configuration and the operation of the nonvolatile latch circuit which is an embodiment of the invention disclosed herein will be described with reference to FIGS. 19A and 19B.

[0315]

5 FIG. 19A illustrates a configuration of the nonvolatile latch circuit 400 including the latch portion 411 and the data holding portion 401 for holding data of the latch portion. FIG. 19B illustrates an example of a timing chart of the nonvolatile latch circuit 400.

[0316]

10 FIG. 19A is an example in which the configuration of the latch portion 411 in FIG. 1 is specifically illustrated. FIG. 19A is an example of the configuration of the latch portion 411 in FIG. 1, where an inverter 412 is used for the first element and an inverter 413 is used for the second element. The structure of the transistor 402 can be similar to that in Embodiment 1 or Embodiment 2.

15 [0317]

The latch portion 411 includes the inverter 412 and the inverter 413. The latch portion 411 has a loop structure in which an output of the inverter 412 is electrically connected to an input of the inverter 413, and an output of the inverter 413 is electrically connected to an input of the inverter 412. In addition, the latch portion
20 411 includes a switch 431 and a switch 432, and the output of the inverter 413 is electrically connected to the input of the inverter 412 via the switch 432.

[0318]

The input of the inverter 412 is electrically connected to the wiring 414 supplied with the input signal of the latch circuit via the switch 431. The output of the
25 inverter 412 is electrically connected to the wiring 415 supplied with the output signal of the latch circuit. A node which is connected to the input of the inverter 412 is referred to as a node P. The node P is electrically connected to the wiring 414 supplied with the input signal of the latch circuit. Moreover, the node P is also electrically connected to the output of the inverter 413. Note that the potential of the node P is the
30 same as that of the input of the inverter 412.

[0319]

In the data holding portion 401, the transistor 402 using an oxide

semiconductor as a semiconductor material for forming a channel formation region is used as a switching element. In addition, the data holding portion 401 includes the capacitor 404 which is electrically connected to one of the source electrode and the drain electrode of the transistor 402. One of the electrodes of the capacitor 404 is electrically connected to one of the source electrode and the drain electrode of the transistor 402. The other of the source electrode and the drain electrode of the transistor is electrically connected to the input of the inverter 412 (the node P) in the latch portion.

[0320]

In addition, the other of the source electrode and the drain electrode of the transistor is electrically connected to the wiring 414 supplied with the input signal of the latch circuit via the switch 431. The other electrode of the capacitor 404 is supplied with a potential V_c . A node where the transistor 402 and the capacitor 404 are electrically connected to each other is referred to as a node S.

[0321]

The transistor 402 using an oxide semiconductor has a function of writing data held in the latch portion 411 into the capacitor 404 of the data holding portion 401. In addition, the transistor 402 has a function of holding the data written into the capacitor 404 of the data holding portion 401. Moreover, the transistor 402 has a function of reading the data held in the capacitor 404 of the data holding portion 401 to the latch portion 411.

[0322]

The wiring 414 is supplied with a potential of an input signal IN from a circuit of a previous stage. A circuit of a subsequent stage is supplied with the potential of the wiring 415 as an output signal OUT. The switch 431 is supplied with a potential of a clock signal ϕ_1 . When the clock signal ϕ_1 is supplied with a high-level potential, the switch 431 is turned on. The switch 432 is supplied with a potential of a clock signal ϕ_2 . When the clock signal ϕ_2 is supplied with a high-level potential, the switch 432 is turned on. A gate of the transistor 402 is supplied with a potential of a control signal ϕ_{LS} . When the control signal ϕ_{LS} is supplied with a high-level potential, the transistor 402 is turned on. In a normal operation period, the clock signal ϕ_2 has an inverted

signal of the clock signal $\phi 1$. Here, an example is shown in which the transistors and the switches are turned on when the potentials of the control signals and the clock signals are at high levels.

[0323]

5 Each of the inverter 412 and the inverter 413 of the latch portion 411 is supplied with a high-level power source voltage VDD and a low-level power source voltage VSS.

[0324]

10 Next, FIG. 19B illustrates an example of a timing chart of the potentials of the input signal IN, the output signal OUT, the control signal ϕ_{LS} , the clock signal $\phi 1$, and the clock signal $\phi 2$ in a period in which the nonvolatile latch circuit 400 is in an operation state (an operation period) and in a period in which the nonvolatile latch circuit 400 is in a stop state (a non-operation period). In addition, FIG. 19B illustrates the potentials of the node S of the data holding portion 401, the node P of the latch

15 portion 411, and the power source voltage VDD-L of the inverter 412 and the inverter 413 of the latch portion 411. The node S indicates the potential of one of the electrodes of the capacitor 404. Note that the other electrode of the capacitor 404 is supplied with a predetermined potential V_c , for example, a ground potential.

[0325]

20 In FIG. 19B, a period *a*, a period *b*, a period *d*, and a period *e* are each a period in which the nonvolatile latch circuit 400 is in an operation state (an operation period), and a period *c* is a period in which the nonvolatile latch circuit 400 is in a stop state (a non-operation period). In each of the period *a* and the period *e*, the nonvolatile latch circuit 400 is in a normal operation period, and the clock signal $\phi 1$ and the clock signal

25 $\phi 2$ are each supplied with a high-level potential or a low-level potential alternately. The period *b* is a preparation period prior to the non-operation period. The period *b* is also referred to as a falling period. The period *d* is a preparation period between the non-operation period and restart of a normal operation period after the supply of power. The period *d* is also referred to as a rising period.

30 [0326]

When the clock signal $\phi 1$ is supplied with a high-level potential and the clock

signal $\phi 2$ is supplied with a low-level potential in the normal operation period (period *a*), the switch 432 is turned off and an inverter loop is cut, the switch 431 is turned on, and the potential of the input signal is inputted to the inverter 412. The potential of the input signal is inverted by the inverter 412 and is supplied to a circuit of a subsequent stage as the output signal OUT. If the potential of the input signal is at a high level when the clock signal $\phi 1$ is supplied with a high-level potential, an output signal having a low-level potential can be obtained. If the potential of the input signal is at a low level when the clock signal $\phi 1$ is supplied with a high-level potential, an output signal having a high-level potential can be obtained.

10 [0327]

When the clock signal $\phi 1$ is supplied with a low-level potential and the clock signal $\phi 2$ is supplied with a high-level potential, the switch 431 is turned off, the switch 432 is turned on and an inverter loop is formed, and the potential of the output signal OUT is held (data is latched, i.e., the logical state of the latch circuit is held).

15 [0328]

In the normal operation period, the control signal ϕ_{LS} is supplied with a potential at which the transistor 402 is turned off and not supplied with a potential at which the transistor 402 is turned on. The node S has the potential corresponding to the charge which has been held. Here, the potential of the node S is set to an undefined value.

20

[0329]

Next, when the control signal ϕ_{LS} is supplied with a potential at which the transistor 402 is turned on in the preparation period (period *b*) prior to the non-operation period, the transistor 402 is turned on and the node S is supplied with the potential of the input of the inverter 412 (the node P) in the latch portion (this operation corresponds to writing). When the potential of the input of the inverter 412 (the node P) in the latch portion 412 is set to a high level, the potential of the node S is a high level. The charge corresponding to the potential is accumulated in the node S.

25

[0330]

30

After that, the transistor 402 is turned off by supplying the control signal ϕ_{LS} with the potential at which the transistor 402 is turned off, and the node S becomes a

floating state. As a result, the charge accumulated in the node S is held without any change (holding).

[0331]

Note that it is sufficient that, in the period *b*, the clock signal $\phi 2$ and the clock
5 signal $\phi 1$ hold a potential at the termination of the period *a*. Alternatively, data at the termination of the period *a* may be latched by fixing the potential of the clock signal $\phi 2$ to a high level and the potential of the clock signal $\phi 1$ to a low level.

[0332]

Next, in the non-operation period (period *c*), the supply of power is stopped
10 and the potential of the power source voltage VDD-L is lowered. The potentials of the clock signal $\phi 1$, the clock signal $\phi 2$, the input signal IN, and the output signal OUT can take any values between VDD and VSS. During this time, the potential of the control signal ϕ_{LS} is held at a low level so that the transistor 402 is turned off. For example, the potential is held at a ground potential. In the non-operation period (period *c*), the
15 charge accumulated in the node S is held by turning off the transistor 402 (holding).

[0333]

Next, in the preparation period (period *d*) between the non-operation period and restart of a normal operation period, power is supplied, and the potentials of the clock signal $\phi 2$ and the clock signal $\phi 1$ are each fixed to a low level. Although the potentials
20 of the node P and the output signal OUT depend on the potential of the node P, the potential of the output signal OUT, and the like which are prior to the supply of power, here, it is considered that the node P has a low-level potential and the output signal OUT has a high-level potential.

[0334]

25 Then, when the control signal ϕ_{LS} is supplied with a potential at which the transistor 402 is turned on, the transistor 402 is turned on and the potential held in the node S is supplied to the latch portion 411. Specifically, the charge is distributed to the node S and the input of the inverter 412 (the node P), and the input of the inverter 412 (the node P) is supplied with the potential corresponding to the charge accumulated in
30 the node S. Here, the charge accumulated in the node S is distributed to the latch portion 411, the potential of the input of the inverter 412 (the node P) is raised, and the

potential of the node S is lowered to some extent. As a result, the input of the inverter 412 (the node P) and the node S each substantially have a high-level potential.

[0335]

Then, the potential of the node P in the latch portion is inverted by the inverter 412 and is supplied to a circuit of a subsequent stage as the output signal OUT. Shown here is an example in which the potential held in the node S and the potential supplied to the node P of the latch portion are at high levels, and an output signal having a low-level potential can be obtained. Accordingly, the logical state of the latch circuit can be restored to the logical state prior to the non-operation period.

[0336]

After that, the transistor 402 is turned off by supplying the control signal ϕ_{LS} with the potential at which the transistor 402 is turned off, and the node S becomes a floating state. As a result, the charge accumulated in the node S is held without any change (holding). The charge accumulated in the node S can be rewritten at the timing at which the control signal ϕ_{LS} is supplied next with a potential at which the transistor 402 is turned on. Therefore, the charge accumulated in the node S is held without any change until the timing at which the control signal ϕ_{LS} is supplied next with the potential at which the transistor 402 is turned on.

[0337]

In addition, in the period d , a period in which the clock signal ϕ_2 is set to high level may be provided after the control signal ϕ_{LS} is supplied with a potential at which the transistor 402 is turned on. When the clock signal ϕ_2 is supplied with a high-level potential, the switch 432 is turned on, and an inverter loop is formed. When the inverter loop is formed, the output signal OUT and the node P are each supplied with a high-level potential or a low-level potential, and the potentials are held (data is latched).

[0338]

As described above, the data is read to the latch portion by distributing the charge to the node S and the input of the inverter 412 (the node P). In the case where the charge corresponding to a high-level potential is accumulated in the node S, after the charge is distributed to the node S and the input of the inverter 412 (the node P), the potential of the input of the inverter 412 (the node P) is set higher than the threshold

value of the inverter 412 (an input potential at which the output of the inverter is inverted) without dependence on the potential of the input of the inverter 412 (the node P) before the transistor 402 is turned on.

[0339]

5 In the case where the charge corresponding to a low-level potential is accumulated in the node S, after the charge is distributed to the node S and the input of the inverter 412 (the node P), the potential of the input of the inverter 412 (the node P) is set lower than the threshold value of the inverter 412 (the input potential at which the output of the inverter is inverted) without dependence on the potential of the input of the inverter 412 (the node P) before the transistor 402 is turned on.

[0340]

 In order to achieve the above, it is preferable that a capacitance of the node S be larger than a capacitance of the node P, for example. In other words, it is preferable that the capacitance of the capacitor 404 to which the node S is electrically connected be
15 larger than an input capacitance of the inverter 412, to which the node P is electrically connected, (the input capacitance corresponds to the gate capacitance of a transistor of the inverter). In addition, in the period d , it is effective to provide a period in which the potential V_c is set to a value between VDD and VSS. Accordingly, the reading operation can be performed more stably.

20 [0341]

 In such a manner, without limitation to the case where the node P has a low-level potential and the output signal OUT has a high-level potential, data can be read to the latch portion also in the case where the node P has a high-level potential and the output signal OUT has a low-level potential. In addition, without limitation to the
25 case where the charge corresponding to a high-level potential is accumulated in the node S, data can be read to the latch portion even in the case where the charge corresponding to a low-level potential is accumulated.

[0342]

 Next, the clock signal $\phi 1$ and the clock signal $\phi 2$ are each supplied with a
30 high-level potential or a low-level potential alternately to be a normal operation state (the period e). At the start of the normal operation period (period e), the operations of

the clock signal $\phi 1$ and the clock signal $\phi 2$ may be started from the same potentials (same state) as the termination of the previous normal operation period (period a) or may be started from an inverted potential of the potential at the termination of the period a (this inverted potential is also referred to as a subsequent state of the period a).

5 [0343]

With the use of a transistor using an oxide semiconductor as a semiconductor material for forming a channel formation region, which serves as a switching element of a data holding portion, for the nonvolatile latch circuit according to this embodiment, a nonvolatile latch circuit which has a wide operating temperature range and operates stably even at a high temperature and in which a logical state of storing data is not erased even after power is turned off or a latch circuit provided with a data holding portion where a refresh period is sufficiently long can be realized. Since data writing is performed by switching of the transistor, the number of rewrites is not substantially limited. In addition, the writing voltage is almost equivalent to the threshold voltage of the transistor; thus, the data writing can be performed at a low voltage. Further, a potential is directly supplied to the data holding portion; therefore, the variation in the amount of charge which is stored as data can be suppressed small and data can be read easily.

[0344]

Various logic circuits can be provided by using the nonvolatile latch circuit. For example, the power consumption can be reduced by turning off power of the unused block. In addition, since a logical state is stored even when power is turned off, a system can be started when power is turned on or terminated when power is turned off, at high speed and low power.

25 [0345]

This embodiment can be freely combined with any of the other embodiments.

[0346]

[Embodiment 4]

In this embodiment, another example of the operation of the nonvolatile latch circuit which is an embodiment of the invention disclosed herein will be described with reference to FIG. 20A. The configuration of the nonvolatile latch circuit is the same as that in FIG. 19A, and the timing chart is different from that in FIG. 19B in this example.

[0347]

FIG. 20A illustrates an example of a timing chart of potentials of an input signal IN, an output signal OUT, a control signal ϕ_{LS} , a clock signal ϕ_1 , and a clock signal ϕ_2 in a period in which the nonvolatile latch circuit 400 is in an operation state (an operation period) and in a period in which the nonvolatile latch circuit 400 is in a stop state (a non-operation period). In addition, FIG. 20A illustrates the potentials of the node S of the data holding portion 401, the node P of the latch portion 411, and the power source voltage VDD-L. The node S indicates the potential of one of the electrodes of the capacitor 404. Note that the other electrode of the capacitor 404 is supplied with a potential V_c .

[0348]

In FIG. 20A, a period *a*, a period *b*, a period *d*, and a period *e* are each a period in which the nonvolatile latch circuit 400 is in an operation state (an operation period), and a period *c* is a period in which the nonvolatile latch circuit 400 is in a stop state (a non-operation period). In each of the period *a* and the period *e*, the nonvolatile latch circuit 400 is in a normal operation period, and the clock signal ϕ_1 and the clock signal ϕ_2 are each supplied with a high-level potential or a low-level potential alternately. The period *b* is a preparation period prior to the non-operation period. The period *b* is also referred to as a falling period. The period *d* is a preparation period between the non-operation period and restart of a normal operation period. The period *d* is also referred to as a rising period.

[0349]

In FIG. 20A, the operations of the period *a*, the period *b*, and the period *c* are similar to those in FIG. 19B. Next, in the preparation period (period *d*) between the non-operation period and restart of a normal operation period after the supply of power, the potentials of the clock signal ϕ_2 and the clock signal ϕ_1 are each fixed to a low level. Although the potentials of the node P and the output signal OUT depend on the potential of the node P, the potential of the output signal OUT, and the like which are prior to the supply of power, here, it is considered that the node P has a low-level potential and the output signal OUT has a high-level potential.

[0350]

Then, when the control signal ϕ_{LS} is supplied with a potential at which the transistor 402 is turned on, the transistor 402 is turned on and the potential held in the node S is supplied to the latch portion 411. Specifically, the charge is distributed to the node S and the input of the inverter 412 (the node P), and the input of the inverter 412 (the node P) is supplied with the potential corresponding to the charge accumulated in the node S. Here, the charge accumulated in the node S is distributed to the latch portion 411, the potential of the input of the inverter 412 (the node P) is raised, and the potential of the node S is lowered to some extent.

[0351]

As a result, the input of the inverter 412 (the node P) and the node S each substantially have a high-level potential. Then, the potential of the node P in the latch portion is inverted by the inverter 412 and is supplied to a circuit of a subsequent stage as the output signal OUT. Shown here is an example in which the potential held in the node S and the potential supplied to the node P of the latch portion are at high levels, and an output signal having a low-level potential can be obtained. Accordingly, the logical state of the latch circuit can be restored to the logical state prior to the non-operation period.

[0352]

Next, the clock signal ϕ_2 is supplied with a high-level potential while the control signal ϕ_{LS} is supplied with a potential at which the transistor 402 is turned on. When the clock signal ϕ_2 is supplied with a high-level potential, the switch 432 is turned on, and an inverter loop is formed. When the inverter loop is formed, the output signal OUT and the node P are each supplied with a high-level potential or a low-level potential, and the potentials are held (data is latched).

[0353]

In particular, since the charge is distributed to the node S and the input of the inverter 412 (the node P), even when the potential of the input of the inverter 412 (the node P) is shifted to some extent from a high-level potential or a low-level potential, a high-level potential or a low-level potential is supplied again. Then, the potential of the node P is supplied to the node S. Accordingly, even when the potential of the node S is shifted to some extent from a high-level potential or a low-level potential, a

high-level potential or a low-level potential is supplied again. As a result, the potential of the node S can be restored to the potential before the change (this operation is also referred to as rewriting).

[0354]

5 After that, the transistor 402 is turned off by supplying the control signal ϕ_{LS} with a potential at which the transistor 402 is turned off, and the potential of the node S becomes a floating state. As a result, the charge accumulated in the node S is held without any change (holding). The charge accumulated in the node S can be rewritten at the timing at which the control signal ϕ_{LS} is supplied next with a potential at which
10 the transistor 402 is turned on. Therefore, the charge accumulated in the node S is held without any change until the timing at which the control signal ϕ_{LS} is supplied next with the potential at which the transistor 402 is turned on.

[0355]

 As described above, the data is read to the latch portion by distributing the
15 charge to the node S and the input of the inverter 412 (the node P). In the case where the charge corresponding to a high-level potential is accumulated in the node S, after the charge is distributed to the node S and the input of the inverter 412 (the node P), the potential of the input of the inverter 412 (the node P) is set higher than the threshold value of the inverter 412 (an input potential at which the output of the inverter is
20 inverted) without dependence on the potential of the input of the inverter 412 (the node P) before the transistor 402 is turned on.

[0356]

 In the case where the charge corresponding to a low-level potential is accumulated in the node S, after the charge is distributed to the node S and the input of
25 the inverter 412 (the node P), the potential of the input of the inverter 412 (the node P) is set lower than the threshold value of the inverter 412 (the input potential at which the output of the inverter is inverted) without dependence on the potential of the input of the inverter 412 (the node P) before the transistor 402 is turned on.

[0357]

30 In order to achieve the above, it is preferable that a capacitance of the node S be larger than a capacitance of the node P, for example. In other words, it is preferable

that the capacitance of the capacitor 404 to which the node S is electrically connected be larger than the capacitance of an input capacitance of the inverter 412, to which the node P is electrically connected, (the input capacitance corresponds to the gate capacitance of a transistor of the inverter). In addition, in the period d , it is effective to provide a period in which the potential V_c is set to a value between VDD and VSS. Accordingly, the reading operation can be performed more stably.

[0358]

In such a manner, without limitation to the case where the node P has a low-level potential and the output signal OUT has a high-level potential, data can be read to the latch portion also in the case where the node P has a high-level potential and the output signal OUT has a low-level potential. In addition, without limitation to the case where the charge corresponding to a high-level potential is accumulated in the node S, data can be read to the latch portion even in the case where the charge corresponding to a low-level potential is accumulated.

[0359]

Next, the clock signal $\phi 1$ and the clock signal $\phi 2$ are each supplied with a high-level potential or a low-level potential alternately to be a normal operation state (the period e). At the start of the normal operation period (period e), the operations of the clock signal $\phi 1$ and the clock signal $\phi 2$ may be started from the same potentials (same state) as the termination of the previous normal operation period (period a) or may be started from an inverted potential of the potential at the termination of the period a (this inverted potential is also referred to as a subsequent state of the period a).

[0360]

With the use of a transistor using an oxide semiconductor as a semiconductor material for forming a channel formation region, which serves as a switching element of a data holding portion, for the nonvolatile latch circuit according to this embodiment, a nonvolatile latch circuit which has a wide operating temperature range and operates stably even at a high temperature and in which a logical state of storing data is not erased even after power is turned off or a latch circuit provided with a data holding portion where a refresh period is sufficiently long can be realized.

[0361]

Since data writing is performed by switching of the transistor, the number of rewrites is not substantially limited. In addition, the writing voltage is almost equivalent to the threshold voltage of the transistor; thus, the data writing can be performed at a low voltage. Further, a potential is directly supplied to the data holding
5 portion; therefore, the variation in the amount of charge which is stored as data can be suppressed small and data can be read easily.

[0362]

Various logic circuits can be provided by using the nonvolatile latch circuit. For example, the power consumption can be reduced by turning off power of the unused
10 block. In addition, since a logical state is stored even when power is turned off, a system can be started when power is turned on or terminated when power is turned off, at high speed and low power.

[0363]

This embodiment can be freely combined with any of the other embodiments.

15 [0364]

[Embodiment 5]

In this embodiment, another example of the operation of the nonvolatile latch circuit which is an embodiment of the invention disclosed herein will be described with reference to FIG. 20B. The configuration of the nonvolatile latch circuit is the same as
20 that in FIG. 19A, and the timing chart is different from those in FIG. 19B and FIG. 20A in this example.

[0365]

FIG. 20A illustrates an example of a timing chart of potentials of an input signal IN, an output signal OUT, a control signal ϕ_{LS} , a clock signal ϕ_1 , and a clock
25 signal ϕ_2 in a period in which the nonvolatile latch circuit 400 is in an operation state (an operation period) and in a period in which the nonvolatile latch circuit 400 is in a stop state (a non-operation period). In addition, FIG. 20A illustrates the potentials of the node S of the data holding portion 401, the node P of the latch portion 411, and the power source voltage VDD-L, and a potential V_c of the other electrode of the capacitor
30 404. The node S indicates the potential of one of the electrodes of the capacitor 404.

[0366]

In FIG. 20B, a period *a*, a period *b*, a period *d*, and a period *e* are each a period in which the nonvolatile latch circuit 400 is in an operation state (an operation period), and a period *c* is a period in which the nonvolatile latch circuit 400 is in a stop state (a non-operation period). In each of the period *a* and the period *e*, the nonvolatile latch circuit 400 is in a normal operation period, and the clock signal $\phi 1$ and the clock signal $\phi 2$ are each supplied with a high-level potential or a low-level potential alternately. The period *b* is a preparation period prior to the non-operation period. The period *b* is also referred to as a falling period. The period *d* is a preparation period between the non-operation period and restart of a normal operation period after the supply of power. The period *d* is also referred to as a rising period.

[0367]

In FIG. 20B, the operations of the period *a*, the period *b*, and the period *c* are similar to those in FIG. 19B. Next, in the preparation period (period *d*) between the non-operation period and restart of a normal operation period, power is supplied, and the potentials of the clock signal $\phi 2$ and the clock signal $\phi 1$ are each fixed to a low level. Although the potentials of the node P and the output signal OUT depend on the potential of the node P, the potential of the output signal OUT, and the like which are prior to the supply of power, here, it is considered that the node P has a low-level potential and the output signal OUT has a high-level potential.

[0368]

Then, when the control signal ϕ_{LS} is supplied with a potential at which the transistor 402 is turned on, the transistor 402 is turned on and the potential held in the node S is supplied to the latch portion 411. Specifically, the charge is distributed to the node S and the input of the inverter 412 (the node P). Then, the potential V_c of the other electrode of the capacitor is supplied with a predetermined potential at the timing at which the control signal ϕ_{LS} is supplied with a potential at which the transistor 402 is turned on. The potential V_c is raised from a low-level potential to be a potential between a low-level potential and a high-level potential.

[0369]

Accordingly, the input of the inverter 412 (the node P) is supplied with a potential in which an increase of the potential V_c of the other electrode of the capacitor

is added to the potential which is determined by the distribution of the charge to the input of the inverter 412 (the node P) and the node S. Here, the charge accumulated in the node S is distributed to the latch portion 411 and the potential V_c is supplied with a predetermined potential, the potential of the input of the inverter 412 (the node P) is raised, and the potential of the node S is lowered to some extent. As a result, the input of the inverter 412 (the node P) and the node S each substantially have a high-level potential.

[0370]

Then, the potential of the node P in the latch portion is inverted by the inverter 412 and is supplied to a circuit of a subsequent stage as the output signal OUT. Accordingly, the logical state of the latch circuit can be restored to the logical state prior to the non-operation period. After that, the potential V_c of the other electrode of the capacitor is restored to be a low-level potential.

[0371]

Next, the clock signal ϕ_2 is supplied with a high-level potential while the control signal ϕ_{LS} is supplied with a potential at which the transistor 402 is turned on. When the clock signal ϕ_2 is supplied with a high-level potential, the switch 432 is turned on, and an inverter loop is formed. When the inverter loop is formed, the output signal OUT and the node P are each supplied with a high-level potential or a low-level potential, and the potentials are held (data is latched).

[0372]

In particular, since the charge is distributed to the node S and the input of the inverter 412 (the node P), even when the potential of the input of the inverter 412 (the node P) is shifted to some extent from a high-level potential or a low-level potential, a high-level potential or a low-level potential is supplied again. Then, the potential of the node P is supplied to the node S. Accordingly, even when the potential of the node S is shifted to some extent from a high-level potential or a low-level potential, a high-level potential or a low-level potential is supplied again. As a result, the potential of the node S can be restored to the potential before the change (this operation is also referred to as rewriting).

[0373]

After that, the transistor 402 is turned off by supplying the control signal ϕ_{LS} with a potential at which the transistor 402 is turned off, and the node S becomes a floating state. As a result, the charge accumulated in the node S is held without any change (holding). The charge accumulated in the node S can be rewritten at the timing
5 at which the control signal ϕ_{LS} is supplied next with a potential at which the transistor 402 is turned on. Therefore, the charge accumulated in the node S is held without any change until the timing at which the control signal ϕ_{LS} is supplied next with the potential at which the transistor 402 is turned on.

[0374]

10 As described above, the data is read to the latch portion by distributing the charge to the node S and the input of the inverter 412 (the node P) and controlling the potential V_c . In the case where the charge corresponding to a high-level potential is accumulated in the node S, after the charge is distributed to the node S and the input of the inverter 412 (the node P), the potential of the input of the inverter 412 (the node P)
15 is set higher than the threshold value of the inverter 412 (an input potential at which the output of the inverter is inverted) without dependence on the potential of the input of the inverter 412 (the node P) before the transistor 402 is turned on.

[0375]

In the case where the charge corresponding to a low-level potential is
20 accumulated in the node S, after the charge is distributed to the node S and the input of the inverter 412 (the node P), the potential of the input of the inverter 412 (the node P) is set lower than the threshold value of the inverter 412 (the input potential at which the output of the inverter is inverted) without dependence on the potential of the input of the inverter 412 (the node P) before the transistor 402 is turned on.

25 [0376]

In order to achieve the above, it is preferable that a capacitance of the node S be larger than a capacitance of the node P, for example. In other words, it is preferable that the capacitance of the capacitor 404 to which the node S is electrically connected be larger than an input capacitance of the inverter 412, to which the node P is electrically
30 connected, (the input capacitance corresponds to the gate capacitance of a transistor of the inverter). In addition, in the period d , it is effective to provide a period in which

the potential V_c is set to a value between VDD and VSS. Accordingly, the reading operation can be performed more stably.

[0377]

In such a manner, without limitation to the case where the node P has a low-level potential and the output signal OUT has a high-level potential, data can be read to the latch portion also in the case where the node P has a high-level potential and the output signal OUT has a low-level potential. In addition, without limitation to the case where the charge corresponding to a high-level potential is accumulated in the node S, data can be read to the latch portion even in the case where the charge corresponding to a low-level potential is accumulated.

[0378]

In particular, as described in this embodiment, the potential V_c of the other electrode of the capacitor is supplied with a predetermined potential at the timing at which the control signal ϕ_{LS} is supplied with the potential at which the transistor 402 is turned on, so that the reading operation can be performed more stably.

[0379]

For example, in the case where the capacitance of the capacitor 404 is small or the case where the supply of power is stopped for a long time, it is difficult to maintain the relationship of the potential of the input of the inverter 412 (the node P) after the charge distribution and the threshold value of the inverter 412 (the input potential at which the output of the inverter is inverted); therefore, there is a possibility that the stability of reading might be degraded.

[0380]

Even in such cases, the aforementioned potential relationship can be maintained and a potential difference thereof can be maintained as large as possible by supplying the potential V_c of the other electrode of the capacitor with a predetermined potential. As a result, stable reading can be performed. In other words, the reading operation can be performed even in the case of a capacitor having a smaller capacitance and thus miniaturization is possible. Further, a data holding period can be made longer.

[0381]

Note that the timing at which the potential V_c of the other electrode of the

capacitor is restored to be a low-level potential may be after the clock signal $\phi 2$ is supplied with a high-level potential. The potential of the other electrode of the capacitor may be restored to a low-level potential before the control signal ϕ_{LS} is supplied with the potential at which the transistor 402 is turned off.

5 [0382]

Next, the clock signal $\phi 1$ and the clock signal $\phi 2$ are each supplied with a high-level potential or a low-level potential alternately to be a normal operation state (the period e). At the start of the normal operation period (period e), the operations of the clock signal $\phi 1$ and the clock signal $\phi 2$ may be started from the same potentials (same state) as the termination of the previous normal operation period (period a) or may be started from an inverted potential of the potential at the termination of the period a (this inverted potential is also referred to as a subsequent state of the period a).
10 [0383]

With the use of a transistor using an oxide semiconductor as a semiconductor material for forming a channel formation region, which serves as a switching element of a data holding portion, for the nonvolatile latch circuit according to this embodiment, a nonvolatile latch circuit which has a wide operating temperature range and operates stably even at a high temperature and in which a logical state of storing data is not erased even after power is turned off or a latch circuit provided with a data holding portion where a refresh period is sufficiently long can be realized.
15 [0384]

Since data writing is performed by switching of the transistor, the number of rewrites is not substantially limited. In addition, the writing voltage is almost equivalent to the threshold voltage of the transistor; thus, the data writing can be performed at a low voltage. Further, a potential is directly supplied to the data holding portion; therefore, the variation in the amount of charge which is stored as data can be suppressed small and data can be read easily. In addition, the capacitance of the capacitor of the data holding portion can be reduced and the capacitor can be reduced in size thus miniaturization is possible.
25 [0385]

30 [0385]

Various logic circuits can be provided by using the nonvolatile latch circuit.

For example, the power consumption can be reduced by turning off power of the unused block. In addition, since a logical state is stored even when power is turned off, a system can be started when power is turned on or terminated when power is turned off, at high speed and low power.

5 [0386]

This embodiment can be freely combined with any of the other embodiments.

[0387]

[Embodiment 6]

In this embodiment, a configuration of a logic circuit including a plurality of the nonvolatile latch circuits which are each an embodiment of the invention disclosed
10 herein will be described with reference to FIG. 21.

[0388]

FIG. 21 illustrates a configuration of a logic circuit including two nonvolatile latch circuits 400 each including a latch portion 411 and a data holding portion 401 for
15 holding data of the latch portion. This logic circuit is referred to as D-FF and used as a register in a CPU or various logic circuits.

[0389]

The configuration of the data holding portion 401 is similar to that in FIG. 1. The configuration of the latch portion 411 is an example in which a NAND is used for
20 the first element and a clocked inverter is used for the second element in the configuration of the latch portion 411 of FIG. 1.

[0390]

The latch portion 411 includes a NAND 412 and a clocked inverter 413. The latch portion 411 has a loop structure in which an output of the NAND 412 is
25 electrically connected to an input of the clocked inverter 413, and an output of the clocked inverter 413 is electrically connected to an input of the NAND 412. In addition, the latch portion 411 includes an analog switch 431.

[0391]

One of inputs of the NAND 412 is electrically connected to a wiring 414
30 supplied with an input signal of the latch circuit 400 via the analog switch 431. The output of the NAND 412 is electrically connected to a wiring 415 supplied with an output signal of the latch circuit 400. The other input of the NAND 412 is electrically

connected to a wiring supplied with a signal RSTB. The analog switch 431 is supplied with a clock signal and an inverted signal of the clock signal. The clocked inverter 413 is supplied with a clock signal and an inverted signal of the clock signal.

[0392]

5 The logic circuit in FIG. 21 includes a nonvolatile latch circuit 400a and a nonvolatile latch circuit 400b as the aforementioned nonvolatile latch circuits 400. The nonvolatile latch circuit 400a is electrically connected to the wiring 414 supplied with a potential of an input signal from a circuit of a previous stage. The wiring 415 supplied with a potential of an output signal of the nonvolatile latch circuit 400a is
10 electrically connected to the wiring 414 supplied with a potential of an input signal of the nonvolatile latch circuit 400b. The nonvolatile latch circuit 400b is electrically connected to the wiring 415 which supplies a potential of an output signal of the nonvolatile latch circuit 400b to a circuit of a subsequent stage.

[0393]

15 In the nonvolatile latch circuit 400a, an analog switch 431 is supplied with a clock signal $\phi 1$ and an inverted signal $\phi 1b$ of the clock signal $\phi 1$, and the clocked inverter 413 is supplied with a clock signal $\phi 2$ and an inverted signal $\phi 2b$ of the clock signal $\phi 2$. In the nonvolatile latch circuit 400b, an analog switch 431 is supplied with a clock signal $\phi 2$ and the inverted signal $\phi 2b$ of the clock signal $\phi 2$, and a clocked
20 inverter 413 is supplied with the clock signal $\phi 1$ and the inverted signal $\phi 1b$ of the clock signal $\phi 1$.

[0394]

With the use of a transistor using an oxide semiconductor as a semiconductor material for forming a channel formation region, which serves as a switching element of
25 a data holding portion, for the nonvolatile latch circuit according to this embodiment, a nonvolatile latch circuit which has a wide operating temperature range and operates stably even at a high temperature and in which a logical state of storing data is not erased even after power is turned off or a latch circuit provided with a data holding portion where a refresh period is sufficiently long can be realized.

[0395]

30 Since data writing is performed by switching of the transistor, the number of

rewrites is not substantially limited. In addition, the writing voltage is almost equivalent to the threshold voltage of the transistor; thus, the data writing can be performed at a low voltage. Further, a potential is directly supplied to the data holding portion; therefore, the variation in the amount of charge which is stored as data can be suppressed small and data can be read easily.

[0396]

Various logic circuits can be provided by using the nonvolatile latch circuit. For example, the power consumption can be reduced by turning off power of the unused block. In addition, since a logical state is stored even when power is turned off, a system can be started when power is turned on or terminated when power is turned off, at high speed and low power.

[0397]

This embodiment can be freely combined with any of the other embodiments.

[0398]

[Embodiment 7]

In this embodiment, another example of the configuration of the nonvolatile latch circuit which is an embodiment of the invention disclosed herein will be described with reference to FIG. 22. FIG. 22 is different from FIG. 1 in this example. FIG. 22 illustrates a configuration of a nonvolatile latch circuit 400 including a latch portion 411 and a data holding portion 401 for holding data of the latch portion.

[0399]

The nonvolatile latch circuit 400 in FIG. 22 includes the latch portion 411 having a loop structure and the data holding portion 401 for holding data of the latch portion. In the latch portion 411 having a loop structure, an output of a first element (D1) 412 is electrically connected to an input of a second element (D2) 413, and an output of the second element (D2) 413 is electrically connected to an input of the first element (D1) 412.

[0400]

The input of the first element (D1) 412 is electrically connected to a wiring 414 supplied with an input signal of the latch circuit. The output of the first element (D1) 412 is electrically connected to a wiring 415 supplied with an output signal of the latch circuit.

[0401]

When there is a plurality of inputs of the first element (D1) 412, one of the inputs can be electrically connected to the wiring 414 supplied with the input signal of the latch circuit. When there is a plurality of inputs of the second element (D2) 413,
5 one of the inputs can be electrically connected to the output of the first element (D1) 412.

[0402]

As the first element (D1) 412, an element in which inputted signal is inverted and the resulting signal serves as an output can be used. For example, as the first
10 element (D1) 412, an inverter, a NAND, a NOR, a clocked inverter, or the like can be used. As the second element (D2) 413, an element in which inputted signal is inverted and the resulting signal serves as an output can be used. For example, as the second element (D2) 413, an inverter, a NAND, a NOR, a clocked inverter, or the like can be used.

15 [0403]

In the data holding portion 401, a transistor 402a and a transistor 402b using an oxide semiconductor as a semiconductor material for forming a channel formation region are each used as a switching element. In addition, the data holding portion 401 includes a capacitor 404a which is electrically connected to a source electrode or a drain
20 electrode of the transistor 402a and a capacitor 404b which is electrically connected to a source electrode or a drain electrode of the transistor 402b.

[0404]

One of electrodes of the capacitor 404a is electrically connected to one of the source electrode and the drain electrode of the transistor 402a, and one of electrodes of
25 the capacitor 404b is electrically connected to one of the source electrode and the drain electrode of the transistor 402b. The other of the source electrode and the drain electrode of the transistor 402a is electrically connected to the wiring 414 supplied with the input of the first element (D1) 412 or the input signal of the latch circuit. The other of the source electrode and the drain electrode of the transistor 402b is electrically
30 connected to the output of the first element (D1) 412 or the wiring 415 supplied with the output signal of the latch circuit. The other electrode of the capacitor 404a and the other electrode of the capacitor 404b are each supplied with a potential V_c .

[0405]

The transistor 402a and the transistor 402b using an oxide semiconductor each have a function of writing data held in the latch portion 411 into the capacitor 404a and the capacitor 404b of the data holding portion 401. In addition, the transistor 402a and the transistor 402b each have a function of holding the data written into the capacitor 404a and the capacitor 404b of the data holding portion 401. Moreover, the transistor 402a and the transistor 402b each have a function of reading the data held in the capacitor 404a and the capacitor 404b of the data holding portion 401 to the latch portion 411.

[0406]

A writing operation of the data held in the latch portion 411 into the data holding portion 401, a holding operation of the data, a reading operation of the data from the data holding portion 401 to the latch portion 411, and a rewriting operation of the data will be described. First, the transistor 402a and the transistor 402b are turned on by supplying a gate electrode of each of the transistor 402a and the transistor 402b with a potential at which the transistor 402a and the transistor 402b are turned on.

[0407]

Accordingly, one of the electrodes of the capacitor 404a is supplied with the data held in the latch portion, that is, a potential of the input of the first element (D1) 412, which is held in the latch portion, and one of the electrodes of the capacitor 404b is supplied with a potential of the output of the first element (D1) 412, which is held in the latch portion. As a result, the charge corresponding to the potential of the input of the first element (D1) 412 is accumulated in one of the electrodes of the capacitor 404a, and the charge corresponding to the potential of the output of the first element (D1) 412 is accumulated in one of the electrodes of the capacitor 404b (this operation corresponds to writing).

[0408]

After that, the transistor 402a and the transistor 402b are turned off in such a manner that a potential of the gate electrode of each of the transistor 402a and the transistor 402b is set to a potential at which each of the transistor 402a and the transistor 402b is turned off. Accordingly, the charge accumulated in one of the electrodes of the capacitor 404a and the capacitor 404b is held (holding).

[0409]

In addition, the transistor 402a and the transistor 402b are turned on by supplying the gate electrode of each of the transistor 402a and the transistor 402b with a potential at which each of the transistor 402a and the transistor 402b is turned on.

5 Accordingly, a charge is distributed to one of the electrodes of the capacitor 404a and the input of the first element (D1) 412, and to one of the electrodes of the capacitor 404b and the output of the first element (D1) 412. As a result, the input and the output of the first element (D1) 412 are each supplied with the potential corresponding to the charge accumulated in one of the electrodes of the capacitor 404b. As a result, the data
10 can be read (reading). Rewriting of the data can be performed in a manner similar to that of the writing and holding of the data.

[0410]

With the use of a transistor using an oxide semiconductor as a semiconductor material for forming a channel formation region, which serves as a switching element of
15 a data holding portion, for the nonvolatile latch circuit according to this embodiment, a nonvolatile latch circuit which has a wide operating temperature range and operates stably even at a high temperature and in which a logical state of storing data is not erased even after power is turned off or a latch circuit provided with a data holding portion where a refresh period is sufficiently long can be realized.

20 [0411]

Since data writing is performed by switching of the transistor, the number of rewrites is not substantially limited. In addition, the writing voltage is almost equivalent to the threshold voltage of the transistor; thus, the data writing can be performed at a low voltage. Further, a potential is directly supplied to the data holding
25 portion; therefore, the variation in the amount of charge which is stored as data can be suppressed small and data can be read easily.

[0412]

Various logic circuits can be provided by using the nonvolatile latch circuit. For example, the power consumption can be reduced by turning off power of the unused
30 block. In addition, since a logical state is stored even when power is turned off, a system can be started when power is turned on or terminated when power is turned off, at high speed and low power.

[0413]

This embodiment can be freely combined with any of the other embodiments.

[0414]

[Embodiment 8]

5 In this embodiment, the configuration and the operation of the nonvolatile latch circuit which is an embodiment of the invention disclosed herein will be described with reference to FIG. 23 and FIGS. 24A and 24B.

[0415]

10 FIG. 23 illustrates a configuration of the nonvolatile latch circuit 400 including the latch portion 411 and the data holding portion 401 for holding data of the latch portion. FIGS. 24A and 24B each illustrate an example of a timing chart of the nonvolatile latch circuit 400.

[0416]

15 FIG. 23 is an example in which the configuration of the latch portion 411 in FIG. 22 is specifically illustrated. FIG. 23 is an example of the configuration of the latch portion 411 in FIG. 22, where an inverter 412 is used for the first element and an inverter 413 is used for the second element. Each structure of the transistor 402a and the transistor 402b can be similar to that in Embodiment 1 or Embodiment 2.

[0417]

20 The latch portion 411 includes the inverter 412 and the inverter 413. The latch portion 411 has a loop structure in which an output of the inverter 412 is electrically connected to an input of the inverter 413, and an output of the inverter 413 is electrically connected to an input of the inverter 412. In addition, the latch portion 411 includes a switch 431 and a switch 432, and the output of the inverter 413 is
25 electrically connected to the input of the inverter 412 via the switch 432.

[0418]

The input of the inverter 412 is electrically connected to the wiring 414 supplied with the input signal of the latch circuit via the switch 431. The output of the inverter 412 is electrically connected to the wiring 415 supplied with the output signal
30 of the latch circuit. A node which is connected to the input of the inverter 412 is referred to as a node P. The node P is electrically connected to the wiring 414 supplied with the input signal of the latch circuit. Moreover, the node P is also electrically

connected to the output of the inverter 413. Note that the potential of the node P is the same as that of the input of the inverter 412.

[0419]

5 In the data holding portion 401, the transistor 402a and the transistor 402b using an oxide semiconductor as a semiconductor material for forming a channel formation region are each used as a switching element. In addition, the data holding portion 401 includes the capacitor 404a which is electrically connected to the source electrode or the drain electrode of the transistor 402a and the capacitor 404b which is electrically connected to the source electrode or the drain electrode of the transistor
10 402b.

[0420]

One of the electrodes of the capacitor 404a is electrically connected to one of the source electrode and the drain electrode of the transistor 402a, and one of the electrodes of the capacitor 404b is electrically connected to one of the source electrode
15 and the drain electrode of the transistor 402b. The other of the source electrode and the drain electrode of the transistor 402a is electrically connected to the wiring 414 supplied with the input signal of the latch circuit and the input of the inverter 412 (the node P) in the latch portion.

[0421]

20 The other of the source electrode and the drain electrode of the transistor 402b is electrically connected to the wiring 415 supplied with the output signal of the latch circuit and the output of the inverter 412 in the latch portion. The other electrode of the capacitor 404a and the other electrode of the capacitor 404b are each supplied with a potential V_c . A node where the transistor 402a and the capacitor 404a are electrically
25 connected to each other is referred to as a node S1, and a node where the transistor 402b and the capacitor 404b are electrically connected to each other is referred to as a node S2.

[0422]

30 The transistor 402a and the transistor 402b using an oxide semiconductor each have a function of writing data held in the latch portion 411 into the capacitor 404a and the capacitor 404b of the data holding portion 401. In addition, the transistor 402a and the transistor 402b each have a function of holding the data written into the capacitor

404a and the capacitor 404b of the data holding portion 401. Moreover, the transistor 402a and the transistor 402b each have a function of reading the data held in the capacitor 404a and the capacitor 404b of the data holding portion 401 to the latch portion 411.

5 [0423]

The wiring 414 is supplied with a potential of an input signal IN from a circuit of a previous stage. A circuit of a subsequent stage is supplied with the potential of the wiring 415 as an output signal OUT. The switch 431 is supplied with a potential of a clock signal $\phi 1$. When the clock signal $\phi 1$ is supplied with a high-level potential, the switch 431 is turned on. The switch 432 is supplied with a potential of a clock signal $\phi 2$. When the clock signal $\phi 2$ is supplied with a high-level potential, the switch 432 is turned on. A gate of each of the transistor 402a and the transistor 402b is supplied with a potential of a control signal ϕ_{LS} .

[0424]

15 When the control signal ϕ_{LS} is supplied with a high-level potential, the transistor 402a and the transistor 402b are turned on. In a normal operation period, the clock signal $\phi 2$ is an inverted signal of the clock signal $\phi 1$. Here, an example is shown in which the transistors and the switches are turned on when the potentials of the control signals and the clock signals are at high levels.

20 [0425]

Each of the inverter 412 and the inverter 413 of the latch portion 411 is supplied with a high-level power source voltage VDD and a low-level power source voltage VSS.

[0426]

25 Next, FIGS. 24A and 24B each illustrate an example of a timing chart of the potentials of the input signal IN, the output signal OUT, the control signal ϕ_{LS} , the clock signal $\phi 1$, and the clock signal $\phi 2$ in a period in which the nonvolatile latch circuit 400 is in an operation state (an operation period) and in a period in which the nonvolatile latch circuit 400 is in a stop state (a non-operation period). In addition, FIGS. 24A and 30 24B illustrate the potentials of the node S1 and the node S2 of the data holding portion 401, the node P of the latch portion 411, and the power source voltage VDD-L of the

inverter 412 and the inverter 413 of the latch portion 411.

[0427]

The node S1 indicates the potential of one of the electrodes of the capacitor 404a. The node S2 indicates the potential of one of the electrodes of the capacitor 404b. Note that the other electrode of the capacitor 404a and the other electrode of the capacitor 404b are each supplied with a predetermined potential V_c , for example, a ground potential.

[0428]

First, FIG. 24A is described. In FIG. 24A, a period *a*, a period *b*, a period *d*, and a period *e* are each a period in which the nonvolatile latch circuit 400 is in an operation state (an operation period), and a period *c* is a period in which the nonvolatile latch circuit 400 is in a stop state (a non-operation period). In each of the period *a* and the period *e*, the nonvolatile latch circuit 400 is in a normal operation period, and the clock signal $\phi 1$ and the clock signal $\phi 2$ are each supplied with a high-level potential or a low-level potential alternately. The period *b* is a preparation period prior to the non-operation period. The period *b* is also referred to as a falling period. The period *d* is a preparation period between the non-operation period and restart of a normal operation period. The period *d* is also referred to as a rising period.

[0429]

When the clock signal $\phi 1$ is supplied with a high-level potential and the clock signal $\phi 2$ is supplied with a low-level potential in the normal operation period (period *a*), the switch 432 is turned off and an inverter loop is cut, the switch 431 is turned on, and the potential of the input signal is inputted to the inverter 412. The potential of the input signal is inverted by the inverter 412 and is supplied to a circuit of a subsequent stage as the output signal OUT. If the potential of the input signal is at a high level when the clock signal $\phi 1$ is supplied with a high-level potential, an output signal having a low-level potential can be obtained. If the potential of the input signal is at a low level when the clock signal $\phi 1$ is supplied with a high-level potential, an output signal having a high-level potential can be obtained.

[0430]

When the clock signal $\phi 1$ is supplied with a low-level potential and the clock

signal ϕ_2 is supplied with a high-level potential, the switch 431 is turned off, the switch 432 is turned on and an inverter loop is formed, and the potential of the output signal OUT is held (data is latched, i.e., the logical state of the latch circuit is held).

[0431]

5 In the normal operation period, the control signal ϕ_{LS} is supplied with a potential at which the transistor 402a and the transistor 402b are turned off and not supplied with a potential at which the transistor 402a and the transistor 402b are turned on. The node S1 and the node S2 each have the potential corresponding to a charge which has been held. Here, the potential of the node S1 and the node S2 are each set
10 to an undefined value.

[0432]

Next, when the control signal ϕ_{LS} is supplied with the potential at which the transistor 402a and the transistor 402b are turned on in the preparation period (period *b*) prior to the non-operation period, the transistor 402a and the transistor 402b are turned
15 on. Accordingly, one of the electrodes of the capacitor 404a (the node S1) is supplied with the potential of the input of the inverter 412 (the node P) in the latch portion, and one of the electrodes of the capacitor 404b (the node S2) is supplied with the potential of the output of the inverter 412 (or the wiring 415 supplied with the output signal) in the latch portion.

20 [0433]

As a result, the charge corresponding to the potential of the input of the inverter 412 (the node P) in the latch portion is accumulated in the capacitor 404a, and the charge corresponding to the potential of the output of the inverter 412 (or the wiring 415 supplied with the output signal) in the latch portion is accumulated in the capacitor 404b
25 (this operation corresponds to writing). For example, at the timing at which the control signal ϕ_{LS} is supplied with the potential at which the transistor 402a and the transistor 402b are turned on, the potential of the node S1 is set to a high level if the potential of the input of the inverter 412 (the node P) in the latch portion is at a high level. In addition, the potential of the node S2 is set to a low level if the potential of
30 the output of the inverter 412 (or the wiring 415 supplied with the output signal) in the latch portion is at a low level.

[0434]

After that, the transistor 402a and the transistor 402b are turned off by supplying the control signal ϕ_{LS} with the potential at which the transistor 402a and the transistor 402b are turned off, and the node S1 and the node S2 each become a floating state. As a result, the charges accumulated in the node S1 and the node S2 are held without any change (holding).

[0435]

Note that it is sufficient that, in the period *b*, the clock signal $\phi 2$ and the clock signal $\phi 1$ hold a potential at the termination of the period *a*. Alternatively, data at the termination of the period *a* may be latched by fixing the potential of the clock signal $\phi 2$ to a high level and the potential of the clock signal $\phi 1$ to a low level.

[0436]

Next, in the non-operation period (period *c*), the supply of power is stopped and the potentials of the power source voltage VDD-L of the inverter 412 and the inverter 413 of the latch portion 411 are lowered. The potentials of the clock signal $\phi 1$, the clock signal $\phi 2$, the input signal IN, the output signal OUT and the node P can take any values between VDD and VSS. During this time, the potential of the control signal ϕ_{LS} is held at a low level so that the transistor 402a and the transistor 402b are turned off. For example, the potential is held at a ground potential. In the non-operation period (period *c*), the charges accumulated in the node S1 and the node S2 are held by turning off the transistor 402a and the transistor 402b (holding).

[0437]

Next, the preparation period (period *d*) starts which is between the non-operation period and restart of a normal operation period. FIG. 24A illustrates an example of the case where the potentials of the node P and the output signal OUT are at low levels at the timing at which the control signal ϕ_{LS} is supplied with the potential at which the transistor 402a and the transistor 402b are turned on.

[0438]

In the period *d*, prior to the supply of power to the inverter 412 and the inverter 413 of the latch portion 411, the potential of the clock signal $\phi 2$ is fixed to a high level and the potential of the clock signal $\phi 1$ is fixed to a low level. When the control signal

ϕ_{LS} is supplied in this state with the potential at which the transistor 402a and the transistor 402b are turned on, the transistor 402a and the transistor 402b are turned on, and the latch portion 411 is supplied with the potentials held in the node S1 and the node S2.

5 [0439]

Specifically, the charge is distributed to the node S1 and the input of the inverter 412 (the input corresponds to the node P, and the input of the inverter 412 (the node P) is supplied with the potential corresponding to the charge accumulated in the node S1. Here, the potential of the input of the inverter 412 (the node P) is raised, and

10 the potential of the node S1 is lowered to some extent.

[0440]

In addition, the charge is distributed to the node S2 and the output of the inverter 412 (or the wiring 415 supplied with the output signal), and the output of the inverter 412 (or the wiring 415 supplied with the output signal) is supplied with the

15 potential corresponding to the charge accumulated in the node S2. Here, both the potential of the input of the inverter 412 (the node P) and the potential of the node S2 are still at low levels.

[0441]

When power is supplied in this state to the inverter 412 and the inverter 413, in

20 the latch portion, the potential of the input of the inverter 412 (the node P) is set to a high level and the potential of the output of the inverter 412 (or the wiring 415 supplied with the output signal) is set to a low level by a potential difference between the input and the output of the inverter 412 and a potential difference between the input and the output of the inverter 413.

25 [0442]

Accordingly, the data of the data holding portion is read to the latch portion, and the logical state of the latch circuit can be restored to the logical state prior to the start of the non-operation period. When the potential difference between the input and the output of the inverter 412 and the potential difference between the input and the

30 output of the inverter 413 are generated in such a manner before power is supplied, the latch circuit can be used as a differential amplifier. As a result, more stable reading can be performed as compared to FIG. 19B.

[0443]

When power is supplied and the inverter loop is formed, the node P and the output signal OUT are each supplied with a high-level potential or a low-level potential, and the potentials are held (data is latched). Then, the potential of the node P and the potential of the output signal OUT are supplied to the node S1 and the node S2, respectively. Accordingly, the node S1 and the node S2 are supplied again with a high-level potential or a low-level potential. As a result, the potentials of the node S1 and the node S2 can be restored to the potentials before the change (this operation is also referred to as rewriting).

[0444]

After that, the transistor 402a and the transistor 402b are turned off by supplying the control signal ϕ_{LS} with the potential at which the transistor 402a and the transistor 402b are turned off, and the node S1 and the node S2 each become a floating state. As a result, the charges accumulated in the node S1 and the node S2 are held without any change (holding). The charges accumulated in the node S1 and the node S2 can be rewritten at the timing at which the control signal ϕ_{LS} is supplied next with the potential at which the transistor 402a and the transistor 402b are turned on. Therefore, the charges accumulated in the node S1 and the node S2 are held without any change until the timing at which the control signal ϕ_{LS} is supplied next with the potential at which the transistor 402a and the transistor 402b are turned on.

[0445]

As described above, the data is read to the latch portion in such a manner that the charge is distributed to the node S1 and the input of the inverter 412 (the node P) and the charge is distributed to the node S2 and the output of the inverter 412 (or the wiring 415 supplied with the output signal). In the case where the charge corresponding to a high-level potential is accumulated in the node S1 and the charge corresponding to a low-level potential is accumulated in the node S2, after the charge distribution, the potential of the input of the inverter 412 (the node P) is set higher than the potential of the output of the inverter 412 (or the wiring 415 supplied with the output signal) without dependence on the potentials of the input of the inverter 412 (the node P) and the output of the inverter 412 (or the wiring 415 supplied with the output

signal) before the transistor 402a and the transistor 402b are turned on.

[0446]

In the case where the charge corresponding to a low-level potential is accumulated in the node S1 and the charge corresponding to a high-level potential is accumulated in the node S2, after the charge distribution, the potential of the input of the inverter 412 (the node P) is set lower than the potential of the output of the inverter 412 (or the wiring 415 supplied with the output signal) without dependence on the potentials of the input of the inverter 412 (the node P) and the output of the inverter 412 (or the wiring 415 supplied with the output signal) before the transistor 402a and the transistor 402b are turned on. In addition, after the charge distribution, the potential of the input of the inverter 412 (the node P) and the potential of the output of the inverter 412 (or the wiring 415 supplied with the output signal) are both set not to be lowered too much. For example, both of the potentials are set so as not to be lower than the threshold voltage of the transistor included in the inverter.

15 [0447]

In order to achieve the above, it is preferable that a capacitance of the node S1 be larger than a capacitance of the node P, for example. In other words, it is preferable that the capacitance of the capacitor 404a to which the node S1 is electrically connected be larger than an input capacitance of the inverter 412, to which the node P is electrically connected, (the input capacitance corresponds to the gate capacitance of a transistor of the inverter). In addition, in the period d , it is effective to provide a period in which the potential V_c is set to a value between VDD and VSS. Accordingly, the reading operation can be performed more stably.

[0448]

25 In such a manner, without limitation to the case where the node P has a low-level potential and the output signal OUT has a high-level potential, data can be read to the latch portion also in the case where the node P has a high-level potential and the output signal OUT has a low-level potential. In addition, without limitation to the case where the charge corresponding to a high-level potential is accumulated in the node S1, data can be read to the latch portion even in the case where the charge corresponding to a low-level potential is accumulated.

30 [0449]

Next, the clock signal $\phi 1$ and the clock signal $\phi 2$ are each supplied with a high-level potential or a low-level potential alternately to be a normal operation state (the period e). At the start of the normal operation period (period e), the operations of the clock signal $\phi 1$ and the clock signal $\phi 2$ may be started from the same potentials (same state) as the termination of the previous normal operation period (period a) or may be started from an inverted potential of the potential at the termination of the period a (this inverted potential is also referred to as a subsequent state of the period a).
[0450]

Next, FIG. 24B is described. In FIG. 24B, the operations of the period a , the period b , and the period c are similar to those in FIG. 24A.
[0451]

Next, the preparation period (period d) starts which is between the non-operation period and restart of a normal operation period. FIG. 24B illustrates an example of the case where the potentials of the node P and the output signal OUT are at high levels at the timing at which the control signal ϕ_{LS} is supplied with the potential at which the transistor 402a and the transistor 402b are turned on.
[0452]

In the period d , prior to the supply of power to the inverter 412 and the inverter 413 of the latch portion 411, the potential of the clock signal $\phi 2$ is fixed to a high level and the potential of the clock signal $\phi 1$ is fixed to a low level. When the control signal ϕ_{LS} is supplied in this state with the potential at which the transistor 402a and the transistor 402b are turned on, the transistor 402a and the transistor 402b are turned on, and the latch portion 411 is supplied with the potentials held in the node S1 and the node S2.

[0453]

Specifically, the charge is distributed to the node S1 and the input of the inverter 412 (the node P), and the input of the inverter 412 (the node P) is supplied with the potential corresponding to the charge accumulated in the node S1. Here, both the potential of the input of the inverter 412 (the node P) and the potential of the node S1 are still at high levels.

[0454]

In addition, the charge is distributed to the node S2 and the output of the inverter 412 (or the wiring 415 supplied with the output signal, and the output of the inverter 412 (or the wiring 415 supplied with the output signal) is supplied with the potential corresponding to the charge accumulated in the node S2. Here, the potential of the output (the output signal OUT) of the inverter 412 is lowered, and the potential of the node S2 is raised to some extent.

[0455]

When power is supplied in this state to the inverter 412 and the inverter 413, in the latch portion, the potential of the input of the inverter 412 (the node P) is set to a high level and the potential of the output of the inverter 412 (or the wiring 415 supplied with the output signal) is set to a low level by a potential difference between the input and the output of the inverter 412 and a potential difference between the input and the output of the inverter 413.

[0456]

Accordingly, the data of the data holding portion is read to the latch portion, and the logical state of the latch circuit can be restored to the logical state prior to the start of the non-operation period. When the potential difference between the input and the output of the inverter 412 and the potential difference between the input and the output of the inverter 413 are generated in such a manner before power is supplied, the latch circuit can be used as a differential amplifier. As a result, more stable reading can be performed as compared to FIG. 19B.

[0457]

When power is supplied and the inverter loop is formed, the node P and the output signal OUT are each supplied with a high-level potential or a low-level potential, and the potentials are held (data is latched). Then, the potential of the node P and the potential of the output signal OUT are supplied to the node S1 and the node S2, respectively. Accordingly, the node S1 and the node S2 are supplied again with a high-level potential or a low-level potential. As a result, the potentials of the node S1 and the node S2 can be restored to the potentials before the change (this operation is also referred to as rewriting).

[0458]

After that, the transistor 402a and the transistor 402b are turned off by

supplying the control signal ϕ_{LS} with the potential at which the transistor 402a and the transistor 402b are turned off, and the node S1 and the node S2 each become a floating state. As a result, the charges accumulated in the node S1 and the node S2 are held without any change (holding). The charges accumulated in the node S1 and the node S2 can be rewritten at the timing at which the control signal ϕ_{LS} is supplied next with the potential at which the transistor 402a and the transistor 402b are turned on. Therefore, the charges accumulated in the node S1 and the node S2 are held without any change until the timing at which the control signal ϕ_{LS} is supplied next with the potential at which the transistor 402a and the transistor 402b are turned on.

10 [0459]

Next, the clock signal $\phi 1$ and the clock signal $\phi 2$ are each supplied with a high-level potential or a low-level potential alternately to be a normal operation state (the period e). At the start of the normal operation period (period e), the operations of the clock signal $\phi 1$ and the clock signal $\phi 2$ may be started from the same potentials (same state) as the termination of the previous normal operation period (period a) or may be started from an inverted potential of the potential at the termination of the period a (this inverted potential is also referred to as a subsequent state of the period a).

15 [0460]

Note that although the example in which the potential difference between the input and the output of the inverter 412 and the potential difference between the input and the output of the inverter 413 are generated before power is supplied is shown here, the nonvolatile latch circuit with the configuration described in this embodiment (FIG. 23) can also be operated with the use of a timing chart similar to that in FIG. 19B.

[0461]

25 Note that in the period d , it is effective to provide a period in which the potential V_c is set to a value between VDD and VSS. Accordingly, the reading operation can be performed more stably.

[0462]

30 With the use of a transistor using an oxide semiconductor as a semiconductor material for forming a channel formation region, which serves as a switching element of a data holding portion, for the nonvolatile latch circuit according to this embodiment, a

nonvolatile latch circuit which has a wide operating temperature range and operates stably even at a high temperature and in which a logical state of storing data is not erased even after power is turned off or a latch circuit provided with a data holding portion where a refresh period is sufficiently long can be realized.

5 [0463]

Since data writing is performed by switching of the transistor, the number of rewrites is not substantially limited. In addition, the writing voltage is almost equivalent to the threshold voltage of the transistor; thus, the data writing can be performed at a low voltage. Further, a potential is directly supplied to the data holding
10 portion; therefore, the variation in the amount of charge which is stored as data can be suppressed small and data can be read easily. Moreover, more stable reading can be performed in the case where the latch circuit is used as a differential amplifier.

[0464]

Various logic circuits can be provided by using the nonvolatile latch circuit.
15 For example, the power consumption can be reduced by turning off power of the unused block. In addition, since a logical state is stored even when power is turned off, a system can be started when power is turned on or terminated when power is turned off, at high speed and low power.

[0465]

20 This embodiment can be freely combined with any of the other embodiments.

[0466]

[Embodiment 9]

In this embodiment, another example of the operation of the nonvolatile latch circuit which is an embodiment of the invention disclosed herein will be described with
25 reference to FIG. 25. The configuration of the nonvolatile latch circuit is the same as that in FIG. 23, and the timing chart is different from those in FIGS. 24A and 24B in this example.

[0467]

Next, FIG. 25 illustrates an example of a timing chart of the potentials of an
30 input signal IN, an output signal OUT, a control signal ϕ_{LS} , a clock signal ϕ_1 , and a clock signal ϕ_2 in a period in which the nonvolatile latch circuit 400 is in an operation

state (an operation period) and in a period in which the nonvolatile latch circuit 400 is in a stop state (a non-operation period). In addition, FIG. 25 illustrates the potentials of the node S1 and the node S2 of the data holding portion 401, the node P of the latch portion 411, and the power source voltage VDD-L of the inverter 412 and the inverter 413 of the latch portion 411, and a potential V_c of the other electrode of each of the capacitor 404a and the capacitor 404b. The node S1 indicates the potential of one of the electrodes of the capacitor 404a. The node S2 indicates the potential of one of the electrodes of the capacitor 404b.

[0468]

In FIG. 25, a period *a*, a period *b*, a period *d*, and a period *e* are each a period in which the nonvolatile latch circuit 400 is in an operation state (an operation period), and a period *c* is a period in which the nonvolatile latch circuit 400 is in a stop state (a non-operation period). In each of the period *a* and the period *e*, the nonvolatile latch circuit 400 is in a normal operation period, and the clock signal $\phi 1$ and the clock signal $\phi 2$ are each supplied with a high-level potential or a low-level potential alternately. The period *b* is a preparation period prior to the non-operation period. The period *b* is also referred to as a falling period. The period *d* is a preparation period between the non-operation period and restart of a normal operation period. The period *d* is also referred to as a rising period.

[0469]

In FIG. 25, the operations of the period *a* and the period *b* are similar to those in FIGS. 24A and 24B. Next, in the non-operation period (period *c*), the supply of power is stopped and the potentials of the power source voltage VDD-L of the inverter 412 and the inverter 413 of the latch portion 411 are lowered. The potentials of the clock signal $\phi 1$, the clock signal $\phi 2$, and the input signal IN can take any values between VDD and VSS. During this time, the potential of the control signal ϕ_{LS} is held at a low level so that the transistor 402a and the transistor 402b are turned off. For example, the potential is held at a ground potential.

[0470]

In the non-operation period (period *c*), the charges accumulated in the node S1 and the node S2 are held by turning off the transistor 402a and the transistor 402b

(holding). In addition, the potential of the output signal OUT is held at a low level. Moreover, the potential of the node P is lowered gradually.

[0471]

Next, the preparation period (period d) starts which is between the non-operation period and restart of a normal operation period. In the period d , prior to the supply of power to the inverter 412 and the inverter 413 of the latch portion 411, the potential of the clock signal ϕ_2 is fixed to a high level and the potential of the clock signal ϕ_1 is fixed to a low level. When the control signal ϕ_{LS} is supplied in this state with the potential at which the transistor 402a and the transistor 402b are turned on, the transistor 402a and the transistor 402b are turned on, and the latch portion 411 is supplied with the potentials held in the node S1 and the node S2.

[0472]

Specifically, the charge is distributed to the node S1 and the input of the inverter 412 (the node P). Then, the potential V_c of the other electrode of the capacitor 404a is supplied with a predetermined potential at the timing at which the control signal ϕ_{LS} is supplied with a potential at which the transistor 402a is turned on. The potential V_c is raised from a low-level potential to be a potential between a low-level potential and a high-level potential. Accordingly, the input of the inverter 412 (the node P) is supplied with a potential in which an increase of the potential V_c of the other electrode of the capacitor 404a is added to the potential which is determined by the distribution of the charge to the input of the inverter 412 (the node P) and the node S1. Here, the potential of the input of the inverter 412 (the node P) is raised, and the potential of the node S1 is lowered to some extent.

[0473]

In addition, the charge is distributed to the node S2 and the output of the inverter 412 (or the wiring 415 supplied with the output signal). Then, the potential V_c of the other electrode of the capacitor 404b is supplied with a predetermined potential at the timing at which the control signal ϕ_{LS} is supplied with a potential at which the transistor 402b is turned on. The potential V_c is raised from a low-level potential to be a potential between a low-level potential and a high-level potential.

[0474]

Accordingly, the output of the inverter 412 (or the wiring 415 supplied with the output signal) is supplied with a potential in which an increase of the potential V_c of the other electrode of the capacitor 404b is added to the potential which is determined by the distribution of the charge to the output of the inverter 412 (or the wiring 415 supplied with the output signal) and the node S2. Here, the potential of the output of the inverter 412 (or the wiring 415 supplied with the output signal) and the potential of the node S2 are raised to some extent by the increase of the potential V_c of the other electrode of the capacitor 404b.

[0475]

When power is supplied in this state to the inverter 412 and the inverter 413, in the latch portion, the potential of the input of the inverter 412 (the node P) is set to a high level and the potential of the output of the inverter 412 (or the wiring 415 supplied with the output signal) is set to a low level by a potential difference between the input and the output of the inverter 412 and a potential difference between the input and the output of the inverter 413.

[0476]

Accordingly, the data of the data holding portion is read to the latch portion, and the logical state of the latch circuit can be restored to the logical state prior to the start of the non-operation period. When the potential difference between the input and the output of the inverter 412 and the potential difference between the input and the output of the inverter 413 are generated in such a manner before power is supplied, the latch circuit can be used as a differential amplifier. As a result, more stable reading can be performed as compared to FIG. 19B.

[0477]

When power is supplied and the inverter loop is formed, the node P and the output signal OUT are each supplied with a high-level potential or a low-level potential, and the potentials are held (data is latched). Then, the potential of the node P and the potential of the output signal OUT are supplied to the node S1 and the node S2, respectively. Accordingly, the node S1 and the node S2 are supplied again with a high-level potential or a low-level potential. As a result, the potentials of the node S1 and the node S2 can be restored to the potentials before the change (this operation is also referred to as rewriting).

[0478]

After that, the potential V_c of the other electrodes of the capacitors is restored to be a low-level potential.

[0479]

5 After that, the transistor 402a and the transistor 402b are turned off by supplying the control signal ϕ_{LS} with the potential at which the transistor 402a and the transistor 402b are turned off, and the node S1 and the node S2 each become a floating state. As a result, the charges accumulated in the node S1 and the node S2 are held without any change (holding).

10 [0480]

The charges accumulated in the node S1 and the node S2 can be rewritten at the timing at which the control signal ϕ_{LS} is supplied next with the potential at which the transistor 402a and the transistor 402b are turned on. Therefore, the charges accumulated in the node S1 and the node S2 are held without any change until the
15 timing at which the control signal ϕ_{LS} is supplied next with the potential at which the transistor 402a and the transistor 402b are turned on.

[0481]

As described above, the data is read to the latch portion in such a manner that the charge is distributed to the node S1 and the input of the inverter 412 (the node P)
20 and the charge is distributed to the node S2 and the output of the inverter 412 (or the wiring 415 supplied with the output signal). In the case where the charge corresponding to a high-level potential is accumulated in the node S1 and the charge corresponding to a low-level potential is accumulated in the node S2, after the charge distribution, the potential of the input of the inverter 412 (the node P) is set higher than
25 the potential of the output of the inverter 412 (or the wiring 415 supplied with the output signal) without dependence on the potentials of the input of the inverter 412 (the node P) and the output of the inverter 412 (or the wiring 415 supplied with the output signal) before the transistor 402a and the transistor 402b are turned on.

[0482]

30 In the case where the charge corresponding to a low-level potential is accumulated in the node S1 and the charge corresponding to a high-level potential is

accumulated in the node S2, after the charge distribution, the potential of the input of the inverter 412 (the node P) is set lower than the potential of the output of the inverter 412 (or the wiring 415 supplied with the output signal) without dependence on the potentials of the input of the inverter 412 (the node P) and the output of the inverter 412 (or the wiring 415 supplied with the output signal) before the transistor 402a and the transistor 402b are turned on. In addition, after the charge distribution, the potential of the input of the inverter 412 (the node P) and the potential of the output of the inverter 412 (or the wiring 415 supplied with the output signal) are both set not to be lowered too much. For example, both of the potentials are set so as not to be lower than the threshold voltage of the transistor included in the inverter.

[0483]

In order to achieve the above, it is preferable that a capacitance of the node S1 be larger than a capacitance of the node P, for example. In other words, it is preferable that the capacitance of the capacitor 404a to which the node S1 is electrically connected be larger than an input capacitance of the inverter 412, to which the node P is electrically connected, (the input capacitance corresponds to the gate capacitance of a transistor of the inverter). In addition, in the period d , it is effective to provide a period in which the potential V_c is set to a value between VDD and VSS. Accordingly, the reading operation can be performed more stably.

[0484]

In such a manner, without limitation to the case where the node P has a low-level potential and the output signal OUT has a high-level potential, data can be read to the latch portion also in the case where the node P has a high-level potential and the output signal OUT has a low-level potential. In addition, without limitation to the case where the charge corresponding to a high-level potential is accumulated in the node S1, data can be read to the latch portion even in the case where the charge corresponding to a low-level potential is accumulated.

[0485]

In particular, as described in this embodiment, the potential V_c of the other electrode of the capacitor is supplied with a predetermined potential at the timing at which the control signal ϕ_{LS} is supplied with the potential at which the transistor 402a and the transistor 402b are turned on, so that the reading operation can be performed

more stably.

[0486]

For example, in the case where the capacitance of the capacitor 404a and the capacitor 404b are small or the case where the supply of power is stopped for a long
5 time, it is difficult to keep a potential difference between the input of the inverter 412 (the node P) and the output of the inverter 412 (or the wiring 415 supplied with the output signal) after the charge distribution, and the potentials of the input of the inverter 412 (the node P) and the output of the inverter 412 (or the wiring 415 supplied with the output signal) after the charge distribution become low. Therefore, there is a
10 possibility that the stability of reading might be degraded.

[0487]

Even in such cases, after the charge distribution, the potentials of the input of the inverter 412 (the node P) and the output of the inverter 412 (or the wiring 415 supplied with the output signal) can be controlled to appropriate potentials by supplying
15 the potential V_c of the other electrode of each of the capacitor 404a and the capacitor 404b with a predetermined potential. As a result, stable reading can be performed. In other words, the reading operation can be performed even in the case of a capacitor having a smaller capacitance and thus miniaturization is possible. Further, a data holding period can be made longer.

20 [0488]

Next, the clock signal $\phi 1$ and the clock signal $\phi 2$ are each supplied with a high-level potential or a low-level potential alternately to be a normal operation state (the period e). At the start of the normal operation period (period e), the operations of the clock signal $\phi 1$ and the clock signal $\phi 2$ may be started from the same potentials
25 (same state) as the termination of the previous normal operation period (period a) or may be started from an inverted potential of the potential at the termination of the period a (this inverted potential is also referred to as a subsequent state of the period a).

[0489]

Note that although the example in which the potential difference between the
30 input and the output of the inverter 412 and the potential difference between the input and the output of the inverter 413 are generated before power is supplied is shown here,

the nonvolatile latch circuit with the configuration described in this embodiment (FIG. 23) can also be operated with the use of a timing chart similar to that in FIG. 19B.

[0490]

With the use of a transistor using an oxide semiconductor as a semiconductor material for forming a channel formation region, which serves as a switching element of a data holding portion, for the nonvolatile latch circuit according to this embodiment, a nonvolatile latch circuit which has a wide operating temperature range and operates stably even at a high temperature and in which a logical state of storing data is not erased even after power is turned off or a latch circuit provided with a data holding portion where a refresh period is sufficiently long can be realized.

[0491]

Since data writing is performed by switching of the transistor, the number of rewrites is not substantially limited. In addition, the writing voltage is almost equivalent to the threshold voltage of the transistor; thus, the data writing can be performed at a low voltage. Further, a potential is directly supplied to the data holding portion; therefore, the variation in the amount of charge which is stored as data can be suppressed small and data can be read easily. In addition, the capacitance of the capacitor of the data holding portion can be reduced and the capacitor can be reduced in size thus miniaturization is possible.

[0492]

Various logic circuits can be provided by using the nonvolatile latch circuit. For example, the power consumption can be reduced by turning off power of the unused block. In addition, since a logical state is stored even when power is turned off, a system can be started when power is turned on or terminated when power is turned off, at high speed and low power.

[0493]

This embodiment can be freely combined with any of the other embodiments.

[0494]

[Embodiment 10]

In this embodiment, another example of a configuration of a logic circuit including a plurality of the nonvolatile latch circuits which are each an embodiment of the invention disclosed herein will be described with reference to FIG. 26.

[0495]

FIG. 26 illustrates a configuration of a logic circuit including two nonvolatile latch circuits 400 each including a latch portion 411 and a data holding portion 401 for holding data of the latch portion. This logic circuit is referred to as D-FF and used as a register in a CPU or various logic circuits.

[0496]

The configuration of the data holding portion 401 is similar to that in FIG. 22. The configuration of the latch portion 411 is an example in which a NAND is used for the first element and a clocked inverter is used for the second element in the configuration of the latch portion 411 of FIG. 22.

[0497]

The latch portion 411 includes a NAND 412 and a clocked inverter 413. The latch portion 411 has a loop structure in which an output of the NAND 412 is electrically connected to an input of the clocked inverter 413, and an output of the clocked inverter 413 is electrically connected to an input of the NAND 412. In addition, the latch portion 411 includes an analog switch 431.

[0498]

One of inputs of the NAND 412 is electrically connected to a wiring 414 supplied with an input signal of the latch circuit 400 via the analog switch 431. The output of the NAND 412 is electrically connected to a wiring 415 supplied with an output signal of the latch circuit 400. The other input of the NAND 412 is electrically connected to a wiring supplied with a signal RSTB. The analog switch 431 is supplied with a clock signal and an inverted signal of the clock signal. The clocked inverter 413 is supplied with a clock signal and an inverted signal of the clock signal.

[0499]

The logic circuit in FIG. 26 includes a nonvolatile latch circuit 400a and a nonvolatile latch circuit 400b as the aforementioned nonvolatile latch circuits 400. The nonvolatile latch circuit 400a is electrically connected to the wiring 414 supplied with a potential of an input signal from a circuit of a previous stage. The wiring 415 supplied with a potential of an output signal of the nonvolatile latch circuit 400a is electrically connected to the wiring 414 supplied with a potential of an input signal of the nonvolatile latch circuit 400b. The nonvolatile latch circuit 400b is electrically

connected to the wiring 415 which supplies a potential of an output signal of the nonvolatile latch circuit 400b to a circuit of a subsequent stage.

[0500]

In the nonvolatile latch circuit 400a, an analog switch 431 is supplied with a clock signal $\phi 1$ and an inverted signal $\phi 1b$ of the clock signal $\phi 1$, and the clocked inverter 413 is supplied with a clock signal $\phi 2$ and an inverter signal $\phi 2b$ of the clock signal $\phi 2$. In the nonvolatile latch circuit 400b, an analog switch 431 is supplied with the clock signal $\phi 2$ and the inverted signal $\phi 2b$ of the clock signal $\phi 2$, and a clocked inverter 413 is supplied with the clock signal $\phi 1$ and the inverter signal $\phi 1b$ of the clock signal $\phi 1$.

[0501]

With the use of a transistor using an oxide semiconductor as a semiconductor material for forming a channel formation region, which serves as a switching element of a data holding portion, for the nonvolatile latch circuit according to this embodiment, a nonvolatile latch circuit which has a wide operating temperature range and operates stably even at a high temperature and in which a logical state of storing data is not erased even after power is turned off or a latch circuit provided with a data holding portion where a refresh period is sufficiently long can be realized.

[0502]

Since data writing is performed by switching of the transistor, the number of rewrites is not substantially limited. In addition, the writing voltage is almost equivalent to the threshold voltage of the transistor; thus, the data writing can be performed at a low voltage. Further, a potential is directly supplied to the data holding portion; therefore, the variation in the amount of charge which is stored as data can be suppressed small and data can be read easily.

[0503]

Various logic circuits can be provided by using the nonvolatile latch circuit. For example, the power consumption can be reduced by turning off power of the unused block. In addition, since a logical state is stored even when power is turned off, a system can be started when power is turned on or terminated when power is turned off, at high speed and low power.

[0504]

This embodiment can be freely combined with any of the other embodiments.

[0505]

[Embodiment 11]

5 Next, another example of the manufacturing method of a transistor using an oxide semiconductor which can be used as the transistor 402 in the above embodiments (such as Embodiment 1 or Embodiment 2) will be described with reference to FIGS. 27A to 27E. In this embodiment, description is made in detail on the case where an oxide semiconductor (particularly with an amorphous structure) which is highly purified
10 is used. Although a top-gate transistor is used as an example in the following description, the structure of the transistor is not limited thereto.

[0506]

First, an insulating layer 202 is formed over a bottom substrate 200. Then, an oxide semiconductor layer 206 is formed over the insulating layer 202 (see FIG. 27A).

15 [0507]

Here, the bottom substrate 200 corresponds to the substrate including the transistor 160 in the lower portion and the like, which is shown in the above embodiments. The above embodiments can be referred to for details of the bottom substrate 200. Note that a surface of the bottom substrate 200 is preferably as flat as possible. In order to achieve this, the surface may be subjected to a chemical
20 mechanical polishing (CMP) method or the like so as to have a peak-to-valley height of 5 nm or less, preferably 1 nm or less, or a root-mean-square roughness (RMS) of 2 nm or less, preferably 0.4 nm or less.

[0508]

25 The insulating layer 202 serves as a base and can be formed in a manner similar to that of the insulating layer 168, the protective insulating layer 144, or the like shown in the above embodiments. The above embodiments can be referred to for details of the insulating layer 202. Note that it is preferable to form the insulating layer 202 so as to contain hydrogen or water as little as possible.

30 [0509]

As the oxide semiconductor layer 206, any of the following oxide semiconductors can be used: an In-Sn-Ga-Zn-O-based oxide semiconductor which is a

four-component metal oxide; an In-Ga-Zn-O-based oxide semiconductor, an In-Sn-Zn-O-based oxide semiconductor, an In-Al-Zn-O-based oxide semiconductor, a Sn-Ga-Zn-O-based oxide semiconductor, an Al-Ga-Zn-O-based oxide semiconductor, or a Sn-Al-Zn-O-based oxide semiconductor which are three-component metal oxides; an
5 In-Zn-O-based oxide semiconductor, a Sn-Zn-O-based oxide semiconductor, an Al-Zn-O-based oxide semiconductor, a Zn-Mg-O-based oxide semiconductor, a Sn-Mg-O-based oxide semiconductor, or an In-Mg-O-based oxide semiconductor which are two-component metal oxides; or an In-O-based oxide semiconductor, a Sn-O-based oxide semiconductor, or a Zn-O-based oxide semiconductor which are one-component
10 metal oxides.

[0510]

In particular, an In-Ga-Zn-O-based oxide semiconductor material has sufficiently high resistance when there is no electric field and thus a sufficiently low off-state current can be obtained. In addition, having a high field-effect mobility, the
15 In-Ga-Zn-O-based oxide semiconductor material is suitable for a semiconductor device.

[0511]

A typical example of the In-Ga-Zn-O-based oxide semiconductor material is represented by $\text{InGaO}_3(\text{ZnO})_m$ ($m > 0$). Another example of the oxide semiconductor material is represented by $\text{InMO}_3(\text{ZnO})_m$ ($m > 0$) where M is used instead of Ga. Here,
20 M denotes one or more of metal elements selected from gallium (Ga), aluminum (Al), iron (Fe), nickel (Ni), manganese (Mn), cobalt (Co), and the like. For example, M can be Ga, Ga and Al, Ga and Fe, Ga and Ni, Ga and Mn, Ga and Co, or the like. Note that the aforementioned composition is only an example obtained from a crystalline structure.

25 [0512]

In this embodiment, the oxide semiconductor layer 206 with an amorphous structure is formed with a sputtering method using an In-Ga-Zn-O-based metal oxide target.

[0513]

30 As the metal oxide target used for forming the oxide semiconductor layer 206 with a sputtering method, for example, a metal oxide target having a composition ratio of $\text{In}_2\text{O}_3:\text{Ga}_2\text{O}_3:\text{ZnO} = 1:1:1$ [molar ratio] can be used. Furthermore, a metal oxide

target having a composition ratio of $\text{In}_2\text{O}_3:\text{Ga}_2\text{O}_3:\text{ZnO} = 1:1:2$ [molar ratio] or a metal oxide target having a composition ratio of $\text{In}_2\text{O}_3:\text{Ga}_2\text{O}_3:\text{ZnO} = 1:1:4$ [molar ratio] can also be used.

[0514]

5 The relative density of the oxide semiconductor in the metal oxide target is greater than or equal to 80 %, preferably greater than or equal to 95 %, and more preferably greater than or equal to 99.9 %. The use of the metal oxide target with high relative density makes it possible to form the oxide semiconductor layer 206 having a dense structure.

10 [0515]

The atmosphere in which the oxide semiconductor layer 206 is formed is preferably a rare gas (typically argon) atmosphere, an oxygen atmosphere, or a mixed atmosphere containing a rare gas (typically argon) and oxygen. Specifically, it is preferable to use, for example, a high-purity gas atmosphere from which an impurity
15 such as hydrogen, water, a hydroxyl group, or hydride is removed to a concentration of a few ppm or less (preferably, a few ppb or less).

[0516]

At the time of forming the oxide semiconductor layer 206, for example, the substrate is held in a treatment chamber kept under reduced pressure and the substrate is
20 heated to a temperature of 100 °C to 550 °C inclusive, preferably 200 °C to 400 °C inclusive. Then, a sputtering gas from which hydrogen, water, or the like is removed is introduced into the treatment chamber while moisture in the treatment chamber is removed, whereby the oxide semiconductor layer 206 is formed using the aforementioned target. The oxide semiconductor layer 206 is formed while the
25 substrate is heated, so that the concentration of an impurity contained in the oxide semiconductor layer 206 can be reduced. Moreover, damage due to sputtering can be reduced. An entrapment vacuum pump is preferably used in order to remove moisture in the treatment chamber. For example, a cryopump, an ion pump, or a titanium sublimation pump can be used. Alternatively, a turbo pump provided with a cold trap
30 may also be used. Since hydrogen, water, or the like is removed from the treatment chamber evacuated with a cryopump, the concentration of an impurity in the oxide

semiconductor layer 206 can be reduced.

[0517]

The oxide semiconductor layer 206 can be formed under the following conditions, for example: the distance between the substrate and the target is 170 mm; the pressure is 0.4 Pa; the direct-current (DC) power is 0.5 kW; and the atmosphere is oxygen (the proportion of the oxygen flow is 100 %), argon (the proportion of the argon flow is 100 %), or a mixed atmosphere containing oxygen and argon. Note that it is preferable to use a pulsed direct-current (DC) power source because dust (such as powder substances formed at the time of film formation) can be reduced and the thickness distribution can be reduced. The thickness of the oxide semiconductor layer 206 is 2 nm to 200 nm inclusive, preferably 5 nm to 30 nm inclusive. Note that the appropriate thickness of the oxide semiconductor layer differs depending on the oxide semiconductor material to be used, the intended purpose of a semiconductor device, or the like; therefore, the thickness may be determined in accordance with the material, the intended purpose, or the like.

[0518]

Note that before the oxide semiconductor layer 206 is formed with a sputtering method, reverse sputtering is preferably performed in which plasma is generated with an argon gas introduced, so that dust on the surface of the insulating layer 202 is removed. Here, the reverse sputtering is a method in which ions collide with a surface to be processed so that the surface is modified, in contrast to normal sputtering in which ions collide with a sputtering target. An example of a method for making ions collide with a surface to be processed is a method in which a high-frequency voltage is applied to the surface to be processed under an argon atmosphere so that plasma is generated near a substrate. Note that an atmosphere of nitrogen, helium, oxygen, or the like may be used instead of an argon atmosphere.

[0519]

Next, the oxide semiconductor layer 206 is processed with a method such as etching using a mask, whereby an island-shaped oxide semiconductor layer 206a is formed.

[0520]

As an etching method for the oxide semiconductor layer 206, either dry etching or wet etching may be employed. It is needless to say that dry etching and wet etching can be used in combination. The etching conditions (e.g., an etching gas or an etchant, etching time, and temperature) are set as appropriate depending on the material so that the oxide semiconductor layer can be etched into a desired shape. The above embodiments can be referred to for details of the etching conditions. The oxide semiconductor layer 206 can be etched in a manner similar to that of the oxide semiconductor layer shown in the above embodiments. The above embodiments can be referred to for details of the etching.

10 [0521]

After that, heat treatment (first heat treatment) is preferably performed on the oxide semiconductor layer 206a. Through the first heat treatment, excess hydrogen (including water and hydroxyl groups) in the oxide semiconductor layer 206a can be removed, the structure of the oxide semiconductor layer 206a can be aligned, and defects in the oxide semiconductor layer 206a can be reduced. The first heat treatment is performed at a temperature of, for example, 300 °C to 550 °C inclusive, or 400 °C to 550 °C inclusive.

[0522]

The heat treatment can be performed in such a manner that, for example, the bottom substrate 200 is introduced into an electric furnace using a resistance heating element or the like, and then heated under a nitrogen atmosphere at 450 °C for 1 hour. The oxide semiconductor layer 206a is not exposed to the air during the heat treatment so that the entry of water or hydrogen can be prevented.

[0523]

The heat treatment apparatus is not limited to the electric furnace and can be an apparatus for heating an object to be processed by thermal conduction or thermal radiation from a medium such as a heated gas. For example, a rapid thermal annealing (RTA) apparatus such as a gas rapid thermal annealing (GRTA) apparatus or a lamp rapid thermal annealing (LRTA) apparatus can be used. An LRTA apparatus is an apparatus for heating an object to be processed by radiation of light (an electromagnetic wave) emitted from a lamp such as a halogen lamp, a metal halide lamp, a xenon arc

lamp, a carbon arc lamp, a high-pressure sodium lamp, or a high-pressure mercury lamp. A GRTA apparatus is an apparatus for performing heat treatment using a high-temperature gas. As the gas, an inert gas which does not react with an object to be processed by heat treatment, for example, nitrogen or a rare gas such as argon is used.

[0524]

For example, as the first heat treatment, a GRTA process may be performed as follows. The substrate is put in an inert gas atmosphere, heated for several minutes, and taken out of the inert gas atmosphere. The GRTA process enables high-temperature heat treatment for a short time. Moreover, the GRTA process can be employed even when the temperature exceeds the upper temperature limit of the substrate because it is heat treatment for a short time.

[0525]

Note that the inert gas atmosphere may be changed during the process to an atmosphere including oxygen. This is because defects caused by oxygen deficiency can be reduced by performing the first heat treatment under an atmosphere containing oxygen.

[0526]

For example, in the case where an electrical furnace is used in the first heat treatment, an atmosphere can be changed when a heat treatment temperature falls. For example, the heat treatment can be performed (at a constant temperature) under an atmosphere of an inert gas such as a rare gas (e.g., helium, neon, or argon) or nitrogen, and the atmosphere can be switched to an atmosphere containing oxygen when the heat treatment temperature falls. As the atmosphere containing oxygen, an oxygen gas or a mixed gas of an oxygen gas and a nitrogen gas can be used.

[0527]

Note that as the inert gas atmosphere, it is preferable to employ an atmosphere that contains nitrogen or a rare gas (e.g., helium, neon, or argon) as its main component and that does not contain water, hydrogen, or the like. For example, the purity of nitrogen or a rare gas such as helium, neon, or argon introduced into the heat treatment apparatus is greater than or equal to 6 N (99.9999 %), preferably greater than or equal to

7 N (99.99999 %) (i.e., the impurity concentration is less than or equal to 1 ppm, preferably less than or equal to 0.1 ppm).

[0528]

5 In any case, when the impurity are reduced through the first heat treatment to form the i-type or substantially i-type oxide semiconductor layer 206a, a transistor with excellent properties can be realized.

[0529]

10 Note that the first heat treatment can also be performed on the oxide semiconductor layer 206 that has not yet been processed into the island-shaped oxide semiconductor layer 206a. In that case, after the first heat treatment, the bottom substrate 200 is taken out of the heating apparatus and a photolithography step is performed.

[0530]

15 The first heat treatment, which has an effect of removing hydrogen or water, can also be referred to as dehydration treatment, dehydrogenation treatment, or the like. The dehydration treatment or dehydrogenation treatment can be performed, for example, after the oxide semiconductor layer is formed, or after a source or drain electrode is stacked over the oxide semiconductor layer 206a. Such dehydration treatment or dehydrogenation treatment may be performed once or plural times.

20 [0531]

Next, a conductive layer is formed to be in contact with the oxide semiconductor layer 206a. Then, a source or drain electrode 208a and a source or drain electrode 208b are formed by selectively etching the conductive layer (see FIG. 27B). This step is similar to the step for forming the source or drain electrode 142a and the like described in the above embodiments. The above embodiments can be referred to for details of the step.

[0532]

30 Next, a gate insulating layer 212 in contact with part of the oxide semiconductor layer 206a is formed (see FIG. 27C). The description of the gate insulating layer in the above embodiments can be referred to for details of the gate insulating layer 212.

[0533]

After the gate insulating layer 212 is formed, second heat treatment is preferably performed under an inert gas atmosphere or an oxygen atmosphere. The heat treatment is performed at a temperature of 200 °C to 450 °C inclusive, preferably 250 °C to 350 °C inclusive. For example, the heat treatment may be performed at
5 250 °C for 1 hour under a nitrogen atmosphere. The second heat treatment can reduce variation in electric characteristics of the transistor. In the case where the gate insulating layer 212 contains oxygen, by supplying oxygen to the oxide semiconductor layer 206a to reduce oxygen deficiency of the oxide semiconductor layer 206a, an i-type (intrinsic) or substantially i-type oxide semiconductor layer can also be formed.

10 [0534]

Note that although the second heat treatment is performed in this embodiment immediately after the gate insulating layer 212 is formed, the timing of the second heat treatment is not limited thereto.

[0535]

15 Next, a gate electrode 214 is formed over the gate insulating layer 212 in a region overlapping with the oxide semiconductor layer 206a (see FIG. 27D). The gate electrode 214 can be formed by forming a conductive layer over the gate insulating layer 212 and then selectively patterning the conductive layer. The description of the gate electrode in the above embodiments can be referred to for details of the gate
20 electrode 214.

[0536]

Next, an interlayer insulating layer 216 and an interlayer insulating layer 218 are formed over the gate insulating layer 212 and the gate electrode 214 (see FIG. 27E). The interlayer insulating layer 216 and the interlayer insulating layer 218 can be formed
25 with a PVD method, a CVD method, or the like. The interlayer insulating layer 216 and the interlayer insulating layer 218 can be formed using a material including an inorganic insulating material such as silicon oxide, silicon oxynitride, silicon nitride, hafnium oxide, aluminum oxide, or tantalum oxide. Note that although a stacked structure of the interlayer insulating layer 216 and the interlayer insulating layer 218 is
30 used in this embodiment, an embodiment of the invention disclosed herein is not limited thereto. A single-layer structure or a stacked structure including three or more layers

can also be used.

[0537]

Note that the interlayer insulating layer 218 is preferably formed so as to have a planarized surface. This is because an electrode, a wiring, or the like can be favorably formed over the interlayer insulating layer 218 when the interlayer insulating layer 218 is formed so as to have a planarized surface.

[0538]

Through the above steps, a transistor 250 including the highly-purified oxide semiconductor layer 206a is completed.

10 [0539]

The transistor 250 illustrated in FIG. 27E includes the following: the oxide semiconductor layer 206a provided over the bottom substrate 200 with the insulating layer 202 interposed therebetween; the source or drain electrode 208a and the source or drain electrode 208b electrically connected to the oxide semiconductor layer 206a; the gate insulating layer 212 covering the oxide semiconductor layer 206a, the source or drain electrode 208a, and the source or drain electrode 208b; the gate electrode 214 over the gate insulating layer 212; the interlayer insulating layer 216 over the gate insulating layer 212 and the gate electrode 214; and the interlayer insulating layer 218 over the interlayer insulating layer 216.

20 [0540]

In the transistor 250 shown in this embodiment, the oxide semiconductor layer 206a is highly purified. Therefore, the concentration of hydrogen in the oxide semiconductor layer 206a is less than or equal to $5 \times 10^{19} / \text{cm}^3$, preferably less than or equal to $5 \times 10^{18} / \text{cm}^3$, more preferably less than or equal to $5 \times 10^{17} / \text{cm}^3$, and still more preferably less than or equal to $1 \times 10^{16} / \text{cm}^3$. In addition, the carrier density of the oxide semiconductor layer 206a is sufficiently low (e.g., less than $1 \times 10^{12} / \text{cm}^3$, preferably less than $1 \times 10^{11} / \text{cm}^3$) as compared to that of a typical silicon wafer (approximately $1 \times 10^{14} / \text{cm}^3$). As a result of this, a sufficiently low off-state current can be obtained. For example, when a drain voltage V_D is +1 V or +10 V and a gate voltage V_G ranges from -5 V to -20 V, the off-state current is less than or equal to 1×10^{-13} A at room temperature. Moreover, the aforementioned transistor has

characteristics of a normally-off transistor. Therefore, leakage current, that is, an off-state current at a state where a voltage between a gate electrode and a source electrode is approximately 0 V is much smaller than that of a transistor using silicon. For example, the leakage current per unit channel width is less than or equal to 10 aA/ μm at room temperature.

[0541]

In this manner, by using the highly-purified and intrinsic oxide semiconductor layer 206a, the off-state current of the transistor can be sufficiently reduced.

[0542]

Note that although, in this embodiment, the transistor 250 is used as the transistor 402 shown in the above embodiments, the invention disclosed herein does not need to be construed as being limited to that case. For example, when the electric characteristics of an oxide semiconductor are sufficiently increased, the oxide semiconductor can be used for all the transistors including transistors included in an integrated circuit. In such a case, it is not necessary to employ a stacked structure as shown in the above embodiments, and a semiconductor device can be formed using, for example, a substrate such as a glass substrate.

[0543]

The structures, methods, and the like described in this embodiment can be combined as appropriate with any of the structures, methods, and the like described in the other embodiments.

[0544]

[Embodiment 12]

Next, another example of the manufacturing method of a transistor using an oxide semiconductor which can be used as the transistor 402 in the above embodiments (such as Embodiment 1 or Embodiment 2) will be described with reference to FIGS. 28A to 28E. In this embodiment, description is made in detail on the case where, as an oxide semiconductor layer, a first oxide semiconductor layer having a crystallized region and a second oxide semiconductor layer that is obtained by crystal growth from the crystallized region of the first oxide semiconductor layer are used. Although a top-gate transistor is used as an example in the following description, the structure of the transistor is not limited thereto.

[0545]

First, an insulating layer 302 is formed over a bottom substrate 300. Next, a first oxide semiconductor layer is formed over the insulating layer 302, and then subjected to first heat treatment so that a region including at least a surface of the first
5 oxide semiconductor layer is crystallized, whereby a first oxide semiconductor layer 304 is formed (see FIG. 28A).

[0546]

Here, the bottom substrate 300 corresponds to the substrate including the transistor 160 in the lower portion and the like, which is shown in the above
10 embodiments. The above embodiments can be referred to for details of the bottom substrate 300. Note that the planarity of the surface of the bottom substrate 300 is particularly important in this embodiment, because it is indispensable for uniform crystal growth. In order to obtain an oxide semiconductor layer with preferable crystallinity, the surface of the bottom substrate 300 may have a peak-to-valley height of
15 1 nm or less, preferably 0.2 nm or less, or a root-mean-square roughness (RMS) of 0.5 nm or less, preferably 0.1 nm or less.

[0547]

The insulating layer 302 serves as a base and can be formed in a manner similar to that of the insulating layer 168, the protective insulating layer 144, or the like
20 shown in the above embodiments. The above embodiments can be referred to for details of the insulating layer 302. Note that it is preferable to form the insulating layer 302 so as to contain hydrogen or water as little as possible.

[0548]

The first oxide semiconductor layer 304 can be formed in a manner similar to
25 that of the oxide semiconductor layer 206 shown in the above embodiment. The above embodiment can be referred to for details of the first oxide semiconductor layer 304 and a manufacturing method thereof. Note that in this embodiment, the first oxide semiconductor layer 304 is intentionally crystallized through the first heat treatment; therefore, the first oxide semiconductor layer 304 is preferably formed using a metal
30 oxide target which causes crystallization easily. For example, ZnO can be used. Further, it is also preferable to use an In-Ga-Zn-O-based oxide in which the proportion of Zn in metal elements (In, Ga, Zn) is greater than or equal to 60 %, because an

In-Ga-Zn-O-based oxide containing Zn at high concentration is easily crystallized. The thickness of the first oxide semiconductor layer 304 is preferably 3 nm to 15 nm inclusive, and in this embodiment, 5 nm for example. Note that the appropriate thickness of the oxide semiconductor layer 304 differs depending on the oxide semiconductor material to be used, the intended purpose of a semiconductor device, or the like; therefore, the thickness may be determined in accordance with the material, the intended purpose, or the like.

[0549]

The first heat treatment is performed at a temperature of 450 °C to 850 °C inclusive, preferably 550 °C to 750 °C inclusive. The time for the first heat treatment is preferably 1 minute to 24 hours inclusive. The temperature and time differ depending on the kind or composition ratio of the oxide semiconductor. In addition, the first heat treatment is preferably performed under an atmosphere that does not contain hydrogen or water, such as an atmosphere of nitrogen, oxygen, or a rare gas (e.g., helium, neon, or argon), from which water is sufficiently removed.

[0550]

The heat treatment apparatus is not limited to the electric furnace can be an apparatus for heating an object to be processed by thermal conduction or thermal radiation from a medium such as a heated gas. For example, a rapid thermal annealing (RTA) apparatus such as a gas rapid thermal annealing (GRTA) apparatus or a lamp rapid thermal annealing (LRTA) apparatus can be used. An LRTA apparatus is an apparatus for heating an object to be processed by radiation of light (an electromagnetic wave) emitted from a lamp such as a halogen lamp, a metal halide lamp, a xenon arc lamp, a carbon arc lamp, a high-pressure sodium lamp, or a high-pressure mercury lamp. A GRTA apparatus is an apparatus for performing heat treatment using a high-temperature gas. As the gas, an inert gas which does not react with an object to be processed by heat treatment, for example, nitrogen or a rare gas such as argon is used.

[0551]

Through the aforementioned first heat treatment, a region including at least the surface of the first oxide semiconductor layer is crystallized. The crystallized region is

formed in such a manner that crystal growth proceeds from the surface of the first oxide semiconductor layer toward the inside of the first oxide semiconductor layer. Note that in some cases, the crystallized region includes a plate-like crystal with an average thickness of 2 nm to 10 nm inclusive. In some cases, the crystallized region also includes a crystal which has an a-b surface substantially parallel to the surface of the oxide semiconductor layer and in which a c-axis is oriented in a direction substantially perpendicular to the surface of the oxide semiconductor layer. Here, a "direction substantially parallel" means a direction within $\pm 10^\circ$ of the parallel direction, and a "direction substantially perpendicular" means a direction within $\pm 10^\circ$ of the perpendicular direction.

[0552]

Through the first heat treatment during which the crystallized region is formed, hydrogen (including water or hydroxyl groups) in the first oxide semiconductor layer is preferably removed. In order to remove hydrogen or the like, the first heat treatment may be performed under an atmosphere of nitrogen, oxygen, or a rare gas (e.g., helium, neon, or argon), which has a purity of 6 N (99.9999 %) or more (i.e., the impurity concentration is less than or equal to 1 ppm), more preferably a purity of 7 N (99.99999 %) or more (i.e., the impurity concentration is less than or equal to 0.1 ppm). Alternatively, the first heat treatment may be performed in ultra-dry air containing H₂O with 20 ppm or less, preferably 1 ppm or less.

[0553]

Furthermore, through the first heat treatment during which the crystallized region is formed, oxygen is preferably supplied to the first oxide semiconductor layer. Oxygen can be supplied to the first oxide semiconductor layer by, for example, changing the atmosphere for the heat treatment to an oxygen atmosphere.

[0554]

The first heat treatment in this embodiment is as follows: hydrogen or the like is removed from the oxide semiconductor layer through heat treatment under a nitrogen atmosphere at 700 °C for 1 hour, and then the atmosphere is changed to an oxygen atmosphere so that oxygen is supplied to the inside of the first oxide semiconductor layer. Note that the main purpose of the first heat treatment is to form the crystallized

region; accordingly, heat treatment for removing hydrogen or the like and treatment for supplying oxygen may be performed separately. For example, heat treatment for crystallization can be performed after heat treatment for removing hydrogen or the like and treatment for supplying oxygen.

5 [0555]

Through such first heat treatment, the crystallized region is formed, hydrogen (including water and hydroxyl groups) or the like is removed, and the first oxide semiconductor layer 304 supplied with oxygen can be obtained.

[0556]

10 Next, a second oxide semiconductor layer 305 is formed over the first oxide semiconductor layer 304 including the crystallized region at least on its surface (see FIG. 28B).

[0557]

The second oxide semiconductor layer 305 can be formed in a manner similar
15 to that of the oxide semiconductor layer 206 shown in the above embodiments. The above embodiments can be referred to for details of the second oxide semiconductor layer 305 and a manufacturing method thereof. Note that the second oxide semiconductor layer 305 is preferably formed to be thicker than the first oxide semiconductor layer 304. Further, the second oxide semiconductor layer 305 is
20 preferably formed so that the total thickness of the first oxide semiconductor layer 304 and the second oxide semiconductor layer 305 are 3 nm to 50 nm inclusive. Note that the appropriate thickness of the oxide semiconductor layer differs depending on the oxide semiconductor material to be used, the intended purpose of a semiconductor device, or the like; therefore, the thickness may be determined in accordance with the
25 material, the intended purpose, or the like.

[0558]

The second oxide semiconductor layer 305 and the first oxide semiconductor layer 304 are preferably formed using materials which have the same main component and further have close lattice constants after crystallization (lattice mismatch is less than
30 or equal to 1 %). This is because in the crystallization of the second oxide semiconductor layer 305, crystal growth easily proceeds from the crystallized region of the first oxide semiconductor layer 304 in the case where materials having the same

main component are used. In addition, the use of materials having the same main component realizes favorable interface physical properties or electric characteristics.

[0559]

5 Note that in the case where a desired film quality is obtained through crystallization, the second oxide semiconductor layer 305 may be formed using a material which has a main component different from that of the material of the first oxide semiconductor layer 304.

[0560]

10 Next, second heat treatment is performed on the second oxide semiconductor layer 305, whereby crystal growth proceeds from the crystallized region of the first oxide semiconductor layer 304, and a second oxide semiconductor layer 306 is formed (see FIG. 28C).

[0561]

15 The second heat treatment is performed at a temperature of 450 °C to 850 °C inclusive, preferably 600 °C to 700 °C inclusive. The time for the second heat treatment is 1 minute to 100 hours inclusive, preferably 5 hours to 20 hours inclusive, and typically 10 hours. Note that also the second heat treatment is preferably performed under an atmosphere that does not contain hydrogen or water.

[0562]

20 Details of the atmosphere and the effect of the second heat treatment are similar to those of the first heat treatment. The heat treatment apparatus that can be used is also similar to that of the first heat treatment. For example, in the second heat treatment, a furnace is filled with a nitrogen atmosphere when a temperature rises, and the furnace is filled with an oxygen atmosphere when the temperature falls, whereby
25 hydrogen or the like can be removed under the nitrogen atmosphere and oxygen can be supplied under the oxygen atmosphere.

[0563]

30 Through the aforementioned second heat treatment, crystal growth can proceed from the crystallized region of the first oxide semiconductor layer 304 to the whole of the second oxide semiconductor layer 305, so that the second oxide semiconductor layer 306 can be formed. In addition, it is possible to form the second oxide semiconductor

layer 306 from which hydrogen (including water and hydroxyl groups) or the like is removed and to which oxygen is supplied. Furthermore, the orientation of the crystallized region of the first oxide semiconductor layer 304 can be improved through the second heat treatment.

5 [0564]

For example, in the case where an In-Ga-Zn-O-based oxide semiconductor material is used for the second oxide semiconductor layer 306, the second oxide semiconductor layer 306 can include a crystal represented by $\text{InGaO}_3(\text{ZnO})_m$ ($m > 0$, and m is not a natural number), a crystal represented by $\text{In}_2\text{Ga}_2\text{ZnO}_7$ (In:Ga:Zn:O =
10 2:2:1:7), or the like. Such crystals are oriented through the second heat treatment so that a c-axis is in a direction substantially perpendicular to the surface of the second oxide semiconductor layer 306.

[0565]

Here, the aforementioned crystals include any of In, Ga, and Zn, and can be
15 considered to have a stacked structure of a plurality of layers parallel to an a-axis and a b-axis. Specifically, the aforementioned crystals have a structure in which a layer containing In and a layer not containing In (a layer containing Ga or Zn) are stacked in the c-axis direction.

[0566]

20 In an In-Ga-Zn-O-based oxide semiconductor crystal, a layer containing In, that is, a layer in a direction parallel to the a-axis and the b-axis has favorable conductivity. This is because electrical conduction in the In-Ga-Zn-O-based oxide semiconductor crystal is mainly controlled by In, and the 5s orbital of an In atom overlaps with the 5s orbital of an adjacent In atom, so that a carrier path is formed.

25 [0567]

Further, in the case where the first oxide semiconductor layer 304 includes an amorphous region at the interface with the insulating layer 302, through the second heat treatment, crystal growth proceeds in some cases from the crystallized region formed on the surface of the first oxide semiconductor layer 304 toward the bottom of the first
30 oxide semiconductor layer to crystallize the amorphous region. Note that in some cases, the amorphous region remains depending on the material of the insulating layer 302, the heat treatment conditions, and the like.

[0568]

In the case where the first oxide semiconductor layer 304 and the second oxide semiconductor layer 305 are formed using oxide semiconductor materials having the same main component, in some cases, the first oxide semiconductor layer 304 and the second oxide semiconductor layer 306 have the same crystal structure, as illustrated in FIG. 28C. Therefore, although indicated by a dotted line in FIG. 28C, the boundary between the first oxide semiconductor layer 304 and the second oxide semiconductor layer 306 cannot be distinguished in some cases so that the first oxide semiconductor layer 304 and the second oxide semiconductor layer 306 can be considered as the same layer.

[0569]

Next, the first oxide semiconductor layer 304 and the second oxide semiconductor layer 306 are processed with a method such as etching using a mask, whereby an island-shaped first oxide semiconductor layer 304a and an island-shaped second oxide semiconductor layer 306a are formed (see FIG. 28D).

[0570]

As an etching method for the first oxide semiconductor layer 304 and the second oxide semiconductor layer 306, either dry etching or wet etching may be employed. It is needless to say that dry etching and wet etching can be used in combination. The etching conditions (e.g., an etching gas or an etchant, etching time, and temperature) are set as appropriate depending on the material so that the oxide semiconductor layer can be etched into a desired shape. The first oxide semiconductor layer 304 and the second oxide semiconductor layer 306 can be etched in a manner similar to that of the oxide semiconductor layer shown in the above embodiments. The above embodiments can be referred to for details of the etching.

[0571]

A region of the oxide semiconductor layers, which becomes a channel formation region, preferably has a planarized surface. For example, the surface of the second oxide semiconductor layer preferably has a peak-to-valley height of 1 nm or less (more preferably 0.2 nm or less) in a region overlapping with a gate electrode (the channel formation region).

[0572]

Next, a conductive layer is formed to be in contact with the second oxide semiconductor layer 306a. Then, a source or drain electrode 308a and a source or drain electrode 308b are formed by selectively etching the conductive layer (see FIG. 28D). The source or drain electrode 308a and the source or drain electrode 308b can be formed in a manner similar to that of the source or drain electrode 142a and the source or drain electrode 142b shown in the above embodiments. The above embodiments can be referred to for details of the source or drain electrode 308a and the source or drain electrode 308b.

[0573]

In the step illustrated in FIG. 28D, crystal layers on the side surfaces of the first oxide semiconductor layer 304a and the second oxide semiconductor layer 306a, which are in contact with the source or drain electrode 308a and the source or drain electrode 308b, are brought into an amorphous state in some cases.

[0574]

Next, a gate insulating layer 312 in contact with part of the second oxide semiconductor layer 306a is formed. The gate insulating layer 312 can be formed with a CVD method or a sputtering method. Then, a gate electrode 314 is formed over the gate insulating layer 312 in a region overlapping with the first oxide semiconductor layer 304a and the second oxide semiconductor layer 306a. After that, an interlayer insulating layer 316 and an interlayer insulating layer 318 are formed over the gate insulating layer 312 and the gate electrode 314 (see FIG. 28E). The gate insulating layer 312, the gate electrode 314, the interlayer insulating layer 316, and the interlayer insulating layer 318 can be formed in a manner similar to that of the gate insulating layer and the like shown in the above embodiments. The above embodiments can be referred to for details of the gate insulating layer 312, the gate electrode 314, the interlayer insulating layer 316, and the interlayer insulating layer 318.

[0575]

After the gate insulating layer 312 is formed, third heat treatment is preferably performed under an inert gas atmosphere or an oxygen atmosphere. The third heat treatment is performed at a temperature of 200 °C to 450 °C inclusive, preferably 250 °C to 350 °C inclusive. For example, the heat treatment may be performed at

250 °C for 1 hour under an atmosphere containing oxygen. The third heat treatment can reduce variation in electric characteristics of the transistor. In the case where the gate insulating layer 312 contains oxygen, by supplying oxygen to the second oxide semiconductor layer 306a to reduce oxygen deficiency of the second oxide semiconductor layer 306a, an i-type (intrinsic) or substantially i-type oxide semiconductor layer can also be formed.

[0576]

Note that although the third heat treatment is performed in this embodiment after the gate insulating layer 312 is formed, the timing of the third heat treatment is not limited thereto. Further, the third heat treatment may be omitted in the case where oxygen is supplied to the second oxide semiconductor layer through other treatment such as the second heat treatment.

[0577]

The gate electrode 314 can be formed by forming a conductive layer over the gate insulating layer 312 and then selectively patterning the conductive layer. The description of the gate electrode in the above embodiments can be referred to for details of the gate electrode 314.

[0578]

The interlayer insulating layer 316 and the interlayer insulating layer 318 can be formed with a PVD method, a CVD method, or the like. The interlayer insulating layer 316 and the interlayer insulating layer 318 can be formed using a material including an inorganic insulating material such as silicon oxide, silicon oxynitride, silicon nitride, hafnium oxide, aluminum oxide, or tantalum oxide. Note that although a stacked structure of the interlayer insulating layer 316 and the interlayer insulating layer 318 is used in this embodiment, an embodiment of the invention disclosed herein is not limited thereto. A single-layer structure or a stacked structure including three or more layers can also be used.

[0579]

Note that the interlayer insulating layer 318 is preferably formed so as to have a planarized surface. This is because an electrode, a wiring, or the like can be favorably formed over the interlayer insulating layer 318 when the interlayer insulating layer 318 is formed so as to have a planarized surface.

[0580]

Through the above steps, a transistor 350 is completed. The transistor 350 uses the first oxide semiconductor layer 304a and the second oxide semiconductor layer 306a which is obtained by crystal growth from the crystallized region of the first oxide semiconductor layer 304a.

[0581]

The transistor 350 illustrated in FIG. 28E includes the following: the first oxide semiconductor layer 304a provided over the bottom substrate 300 with the insulating layer 302 interposed therebetween; the second oxide semiconductor layer 306a provided over the first oxide semiconductor layer 304a; the source or drain electrode 308a and the source or drain electrode 308b electrically connected to the second oxide semiconductor layer 306a; the gate insulating layer 312 covering the second oxide semiconductor layer 306a, the source or drain electrode 308a, and the source or drain electrode 308b; the gate electrode 314 over the gate insulating layer 312; the interlayer insulating layer 316 over the gate insulating layer 312 and the gate electrode 314; and the interlayer insulating layer 318 over the interlayer insulating layer 316.

[0582]

In the transistor 350 shown in this embodiment, the first oxide semiconductor layer 304a and the second oxide semiconductor layer 306a are highly purified. Therefore, the concentration of hydrogen in the first oxide semiconductor layer 304a and the second oxide semiconductor layer 306a is less than or equal to $5 \times 10^{19} / \text{cm}^3$, preferably less than or equal to $5 \times 10^{18} / \text{cm}^3$, more preferably less than or equal to $5 \times 10^{17} / \text{cm}^3$, and still more preferably less than or equal to $1 \times 10^{16} / \text{cm}^3$. In addition, the carrier density of the first semiconductor layer 304a and the second oxide semiconductor layer 306a is sufficiently low (e.g., less than $1 \times 10^{12} / \text{cm}^3$, preferably less than $1 \times 10^{11} / \text{cm}^3$) as compared to that of a typical silicon wafer (approximately $1 \times 10^{14} / \text{cm}^3$). As a result of this, a sufficiently low off-state current can be obtained. For example, when a drain voltage V_D is +1 V or +10 V and a gate voltage V_G ranges from -5 V to -20 V, the off-state current is less than or equal to 1×10^{-13} A at room temperature. Moreover, the aforementioned transistor has characteristics of a normally-off transistor. Therefore, leakage current, that is, an off-state current at a

state where a voltage between a gate electrode and a source electrode is approximately 0 V is much smaller than that of a transistor using silicon. For example, the leakage current per unit channel width is less than or equal to 10 aA/ μm at room temperature.

[0583]

5 In this manner, by using the highly-purified and intrinsic first oxide semiconductor layer 304a and second oxide semiconductor layer 306a, the off-state current of the transistor can be sufficiently reduced.

[0584]

10 Furthermore, in this embodiment, the first oxide semiconductor layer 304a having a crystallized region and the second oxide semiconductor layer 306a which is obtained by crystal growth from the crystallized region of the first oxide semiconductor layer 304a are used as the oxide semiconductor layer. Thus, the field-effect mobility can be increased and a transistor with favorable electric characteristics can be realized.

[0585]

15 Note that although, in this embodiment, the transistor 350 is used as the transistor 402 shown in the above embodiments, the invention disclosed herein does not need to be construed as being limited to that case. For example, the transistor 350 shown in this embodiment uses the first oxide semiconductor layer 304a having a crystallized region and the second oxide semiconductor layer 306a which is obtained by
20 crystal growth from the crystallized region of the first oxide semiconductor layer 304a, and thus has a high field-effect mobility. Accordingly, the oxide semiconductor can be used for all the transistors including transistors included in an integrated circuit. In such a case, it is not necessary to employ a stacked structure as shown in the above embodiments, and a semiconductor device can be formed using, for example, a
25 substrate such as a glass substrate.

[0586]

The structures, methods, and the like described in this embodiment can be combined as appropriate with any of the structures, methods, and the like described in the other embodiments.

30 [0587]

[Embodiment 13]

In this embodiment, examples of the electronic device on which a

semiconductor device using the nonvolatile latch circuit according to any of the above embodiments is mounted will be described with reference to FIGS. 29A to 29F. The electronic device on which the semiconductor device using the nonvolatile latch circuit according to any of the above embodiments is mounted has excellent characteristics which cannot be seen in the conventional technology. Therefore, it is possible to provide an electronic device having a novel structure with the semiconductor device using the nonvolatile latch circuit. Note that the semiconductor device using the nonvolatile latch circuit according to any of the above embodiments is integrated and mounted on a circuit board or the like to be mounted on an electronic device.

10 [0588]

FIG. 29A illustrates a laptop personal computer including the semiconductor device using the nonvolatile latch circuit according to any of the above embodiments. The laptop personal computer includes a main body 1301, a housing 1302, a display portion 1303, a keyboard 1304, and the like. A laptop personal computer with excellent performance can be provided by applying the semiconductor device according to the invention disclosed herein to a laptop personal computer.

[0589]

FIG. 29B illustrates a portable digital assistant (PDA) including the semiconductor device using the nonvolatile latch circuit according to any of the above embodiments. A main body 1311 includes a display portion 1313, an external interface 1315, operation keys 1314, and the like. Further, a stylus 1312 is provided as an accessory for operation. A portable digital assistant (PDA) with excellent performance can be provided by applying the semiconductor device according to the invention disclosed herein to a portable digital assistant (PDA).

25 [0590]

FIG. 29C illustrates an e-book reader 1320 as an example of electronic paper including the semiconductor device using the nonvolatile latch circuit according to any of the above embodiments. The e-book reader 1320 includes two housings: a housing 1321 and a housing 1323. The housing 1321 is combined with the housing 1323 by a hinge 1337, so that the e-book reader 1320 can be opened and closed with the hinge 1337 used as an axis. Such a structure allows the e-book reader 1320 to be used as paper books.

[0591]

The housing 1321 includes a display portion 1325, and the housing 1323 includes a display portion 1327. The display portion 1325 and the display portion 1327 can display a continuous image or different images. The structure for displaying
5 different images allows text to be displayed on the right display portion (the display portion 1325 in FIG. 29C) and images to be displayed on the left display portion (the display portion 1327 in FIG. 29C).

[0592]

FIG. 29C illustrates an example of the case where the housing 1321 includes an
10 operating portion and the like. For example, the housing 1321 includes a power button 1331, operation keys 1333, a speaker 1335, and the like. The operation keys 1333 allow pages to be turned. Note that a keyboard, a pointing device, or the like may also be provided on the surface of the housing, on which the display portion is provided. Furthermore, an external connection terminal (an earphone terminal, a USB terminal, a
15 terminal that can be connected to various cables such as an AC adapter and a USB cable, or the like), a recording medium insertion portion, and the like may be provided on the back surface or the side surface of the housing. The e-book reader 1320 can also serve as an electronic dictionary.

[0593]

20 In addition, the e-book reader 1320 may have a structure capable of transmitting and receiving data wirelessly. Through wireless communication, desired book data or the like can be purchased and downloaded from an electronic book server.

[0594]

Note that electronic paper can be used in any field as long as data is displayed.
25 For example, electronic paper can be applied to posters, advertisement in vehicles such as trains, and a variety of cards such as credit cards, as well as e-book readers. Electronic paper with excellent performance can be provided by applying the semiconductor device according to the invention disclosed herein to electronic paper.

[0595]

30 FIG. 29D illustrates a cellular phone including the semiconductor device using the nonvolatile latch circuit according to any of the above embodiments. The cellular phone includes two housings: a housing 1340 and a housing 1341. The housing 1341

includes a display panel 1342, a speaker 1343, a microphone 1344, a pointing device 1346, a camera lens 1347, an external connection terminal 1348, and the like. The housing 1340 includes a solar cell 1349 for charging the cellular phone, an external memory slot 1350, and the like. An antenna is built in the housing 1341.

5 [0596]

The display panel 1342 includes a touch panel. A plurality of operation keys 1345 which are displayed as an image are shown by dashed lines in FIG. 29D. Note that the cellular phone includes a booster circuit for increasing a voltage outputted from the solar cell 1349 to a voltage needed for each circuit. In addition to the above

10 structure, a noncontact IC chip, a small recording device, or the like may be built in the cellular phone.

[0597]

The display orientation of the display panel 1342 changes as appropriate in accordance with the application mode. Further, the camera lens 1347 is provided on

15 the same surface as the display panel 1342, so that the cellular phone can be used as a video phone. The speaker 1343 and the microphone 1344 can be used for videophone calls, recording, and playing sound, etc. as well as voice calls. Moreover, the housing 1340 and the housing 1341 which are unfolded as in FIG. 29D can overlap with each other by sliding. Thus, the cellular phone can be in a suitable size for portable use.

20 [0598]

The external connection terminal 1348 is connectable to an AC adaptor and a variety of cables such as a USB cable, which enables charging of the cellular phone and data communication. Moreover, a larger amount of data can be saved and moved by inserting a recording medium to the external memory slot 1350. In addition to the

25 above functions, an infrared communication function, a television reception function, or the like may be provided. A cellular phone with excellent performance can be provided by applying the semiconductor device according to the invention disclosed herein to a cellular phone.

[0599]

30 FIG. 29E illustrates a digital camera including the semiconductor device using the nonvolatile latch circuit according to any of the above embodiments. The digital camera includes a main body 1361, a display portion A 1367, an eyepiece portion 1363,

an operation switch 1364, a display portion B 1365, a battery 1366, and the like. A digital camera with excellent performance can be provided by applying the semiconductor device according to the invention disclosed herein to a digital camera.

[0600]

5 FIG. 29F illustrates a television set including the semiconductor device using the nonvolatile latch circuit according to any of the above embodiments. A television set 1370 includes a housing 1371 provided with a display portion 1373. Images can be displayed on the display portion 1373. Here, the housing 1371 is supported by a stand 1375.

10 [0601]

The television set 1370 can operate by an operation switch included in the housing 1371 or by a remote controller 1380 separately provided. Channels and volume can be controlled by operation keys 1379 included in the remote controller 1380, and images displayed on the display portion 1373 can thus be controlled. Further, the remote controller 1380 can be provided with a display portion 1377 for displaying data outputted from the remote controller 1380.

[0602]

Note that the television set 1370 preferably includes a receiver, a modem, and the like. With the receiver, a general television broadcast can be received.

20 Furthermore, when the television set 1370 is connected to a communication network by wired or wireless connection via the modem, one-way (from a transmitter to a receiver) or two-way (between a transmitter and a receiver, between receivers, or the like) data communication can be performed. A television set with excellent performance can be provided by applying the semiconductor device according to the invention disclosed herein to a television set.

[0603]

The structures, methods, and the like described in this embodiment can be combined as appropriate with any of the structures, methods, and the like described in the other embodiments.

30

The present application is based on Japanese Patent Application serial No. 2009-282139 filed with the Japan Patent Office on December 11, 2009, the entire

contents of which are hereby incorporated by reference.

CLAIMS

1. A nonvolatile latch circuit comprising:
a latch portion; and
5 a data holding portion configured to hold data of the latch portion,
the data holding portion comprising:
a transistor; and
a capacitor,
wherein a channel formation region of the transistor includes an oxide
10 semiconductor layer, and
wherein one of a source electrode and a drain electrode of the
transistor is electrically connected to one of electrodes of the capacitor, and
wherein the other of the source electrode and the drain electrode of the
transistor is electrically connected to a wiring supplied with an input signal.
15
2. The nonvolatile latch circuit according to claim 1, wherein the oxide
semiconductor layer contains indium, gallium, and zinc.
3. The nonvolatile latch circuit according to claim 1, wherein the transistor
20 controls writing of data into the capacitor.
4. The nonvolatile latch circuit according to claim 1, wherein the transistor has
a function of holding data in the capacitor.
- 25 5. The nonvolatile latch circuit according to claim 1, wherein the transistor has
a function of reading data held in the capacitor to the latch portion.
6. A nonvolatile logic circuit, wherein a plurality of nonvolatile latch circuits
according to claim 1 is included.
- 30 7. A semiconductor device, wherein a logic circuit including the nonvolatile
latch circuit according to claim 1 is used.

8. A nonvolatile latch circuit comprising:

a latch portion; and

a data holding portion configured to hold data of the latch portion,

5 the latch portion comprising:

a first element; and

a second element,

wherein an output of the first element is electrically connected to an input of the second element, and an output of the second element is electrically
10 connected to an input of the first element, and

wherein the input of the first element is electrically connected to a wiring supplied with a input signal, and the output of the first element is electrically connected to a wiring supplied with an output signal,

the data holding portion comprising:

15 a transistor; and

a capacitor,

wherein a channel formation region of the transistor includes an oxide semiconductor layer,

wherein one of a source electrode and a drain electrode of the transistor is electrically connected to one of electrodes of the capacitor, and
20

wherein the other of the source electrode and the drain electrode of the transistor is electrically connected to the input of the first element.

9. The nonvolatile latch circuit according to claim 8, wherein the first element
25 is an inverter and the second element is an inverter.

10. The nonvolatile latch circuit according to claim 8, wherein the latch portion comprises:

a first switch; and

30 a second switch,

wherein the output of the second element is electrically connected to the input of the first element through the second switch, and

wherein the input of the first element is electrically connected to a wiring supplied with an input signal through the first switch.

11. The nonvolatile latch circuit according to claim 8, wherein the first element
5 is a NAND and the second element is a crooked inverter.

12. The nonvolatile latch circuit according to claim 8, wherein the oxide semiconductor layer contains indium, gallium, and zinc.

10 13. A nonvolatile logic circuit, wherein a plurality of nonvolatile latch circuits according to claim 8 is included.

14. A semiconductor device, wherein a logic circuit including the nonvolatile latch circuit according to claim 8 is used.

15

15. A nonvolatile latch circuit comprising:

a latch portion; and

a data holding portion configured to hold data of the latch portion,

the data holding portion comprising:

20

a first transistor;

a second transistor;

a first capacitor; and

a second capacitor,

wherein channel formation regions of the first transistor and the

25

second transistor each include an oxide semiconductor layer,

wherein one of a source electrode and a drain electrode of the first transistor is electrically connected to one of electrodes of the first capacitor,

wherein the other of the source electrode and the drain electrode of the first transistor is electrically connected to a wiring supplied with an input signal,

30

wherein one of a source electrode and a drain electrode of the second transistor is electrically connected to one of electrodes of the second capacitor, and

wherein the other of the source electrode and the drain electrode of the

second transistor is electrically connected to a wiring supplied with an output signal

16. The nonvolatile latch circuit according to claim 15, wherein the oxide semiconductor layer contains indium, gallium, and zinc.

5

17. The nonvolatile latch circuit according to claim 15, wherein the first transistor controls writing of data into the first capacitor, and wherein the second transistor controls writing of data into the second capacitor.

10

18. The nonvolatile latch circuit according to claim 15, wherein the first transistor has a function of holding data in the first capacitor, and wherein the second transistor has a function of holding data in the second capacitor.

15

19. The nonvolatile latch circuit according to claim 15, wherein the first transistor has a function of reading data held in the first capacitor to the latch portion, and wherein the second transistor has a function of reading data held in the second capacitor to the latch portion.

20

20. A nonvolatile logic circuit, wherein a plurality of nonvolatile latch circuits according to claim 15 is included.

21. A semiconductor device, wherein a logic circuit including the nonvolatile latch circuit according to claim 15 is used.

25

22. A nonvolatile latch circuit comprising:

a latch portion; and

a data holding portion configured to hold data of the latch portion,

the latch portion comprising:

a first element; and

30

a second element,

wherein an output of the first element is electrically connected to an input of the second element, and an output of the second element is electrically

connected to an input of the first element, and

wherein the input of the first element is electrically connected to a wiring supplied with an input signal, and the output of the first element is electrically connected to a wiring supplied with an output signal,

5 the data holding portion comprising:

a first transistor;

a second transistor;

a first capacitor; and

a second capacitor,

10 wherein channel formation regions of the first transistor and the second transistor each include an oxide semiconductor layer,

wherein one of a source electrode and a drain electrode of the first transistor is electrically connected to one of electrodes of the first capacitor,

15 wherein the other of the source electrode and the drain electrode of the first transistor is electrically connected to the input of the first element,

wherein one of a source electrode and a drain electrode of the second transistor is electrically connected to one of electrodes of the second capacitor, and

wherein the other of the source electrode and the drain electrode of the second transistor is electrically connected to the output of the first element.

20

23. A nonvolatile latch circuit according to claim 22, wherein the latch portion comprises:

a first switch; and

a second switch,

25 wherein the output of the second element is electrically connected to an input of the first element through the second switch, and

wherein the input of the first element is electrically connected to the wiring supplied with the input signal through the first switch.

30

24. A nonvolatile latch circuit according to claim 22, wherein the first element is a NAND and the second element is a crooked inverter.

25. The nonvolatile latch circuit according to claim 22, wherein the oxide semiconductor layer contains indium, gallium, and zinc.

26. A nonvolatile logic circuit, wherein a plurality of nonvolatile latch circuits
5 according to claim 22 is included.

27. A semiconductor device, wherein a logic circuit including the nonvolatile latch circuit according to claim 22 is used.

10 28. A CPU comprising:
a logic circuit,
the logic circuit comprising:
a latch circuit,
the latch circuit comprising:
15 a latch portion; and
a data holding portion configured to hold data of the latch
portion,
the data holding portion comprising:
a transistor; and
20 a capacitor,
wherein a channel formation region of the transistor
includes an oxide semiconductor layer,
wherein one of a source electrode and a drain
electrode of the transistor is electrically connected to one of electrodes of the capacitor,
25 and
wherein the other of the source electrode and the
drain electrode of the transistor is electrically connected to a wiring supplied with an
input signal.

30 29. A CPU comprising:
a logic circuit,
the logic circuit comprising:

a latch circuit,

the latch circuit comprising:

a latch portion; and

a data holding portion configured to hold data of the latch

5 portion,

the data holding portion comprising:

a first transistor;

a second transistor;

a first capacitor; and

10 a second capacitor,

wherein channel formation regions of the first transistor and the second transistor each include an oxide semiconductor layer,

wherein one of a source electrode and a drain electrode of the first transistor is electrically connected to one of electrodes of the first capacitor,

15

wherein the other of the source electrode and the drain electrode of the first transistor is electrically connected to a wiring supplied with an input signal,

wherein one of a source electrode and a drain electrode of the second transistor is electrically connected to one of electrodes of the second capacitor, and

20

wherein the other of the source electrode and the drain electrode of the second transistor is electrically connected to a wiring supplied with an output signal.

25

FIG. 1

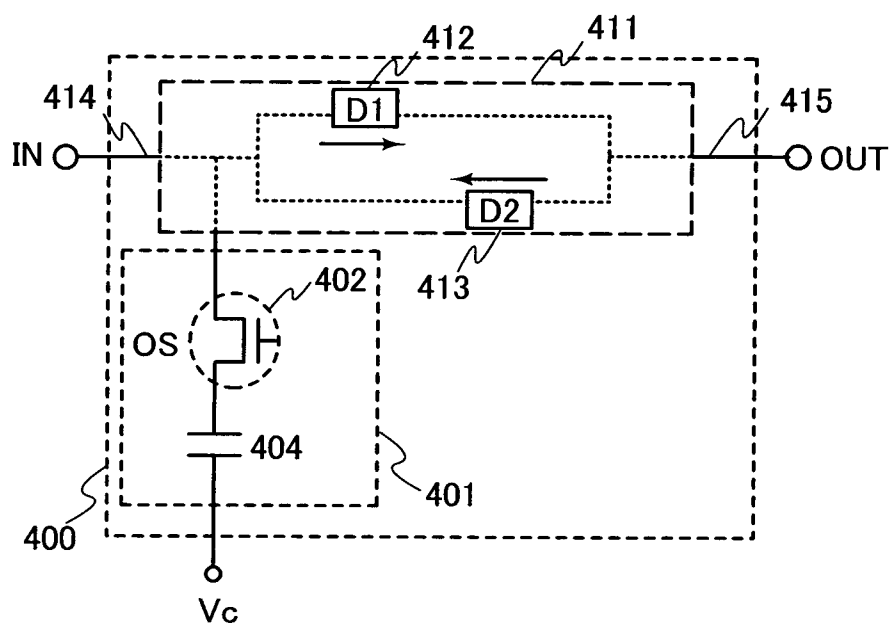


FIG. 2A

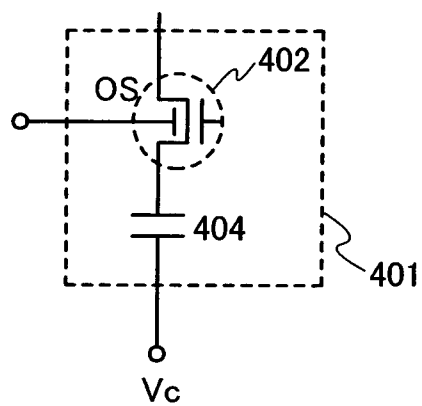


FIG. 2B

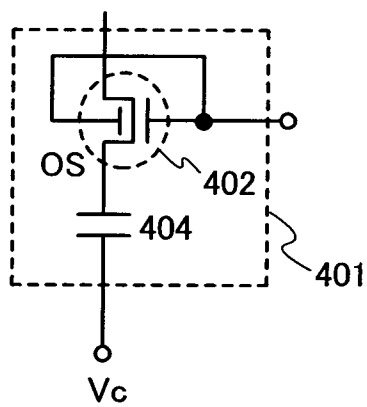


FIG. 3A

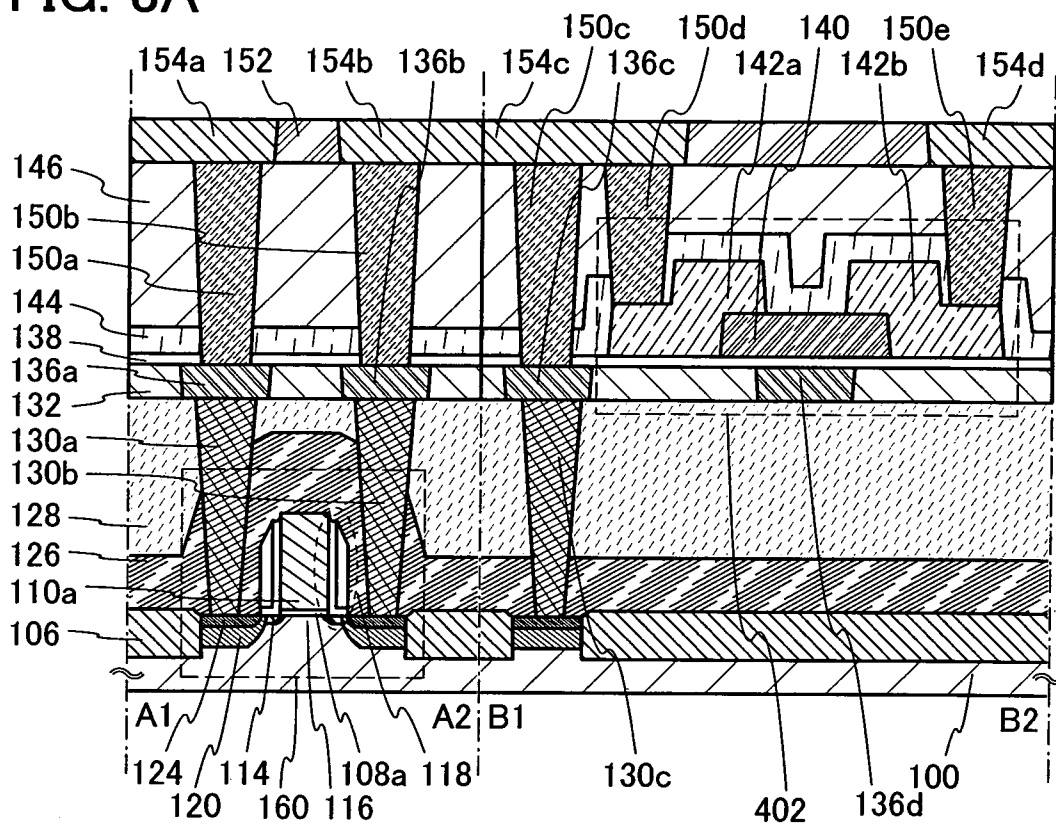


FIG. 3B

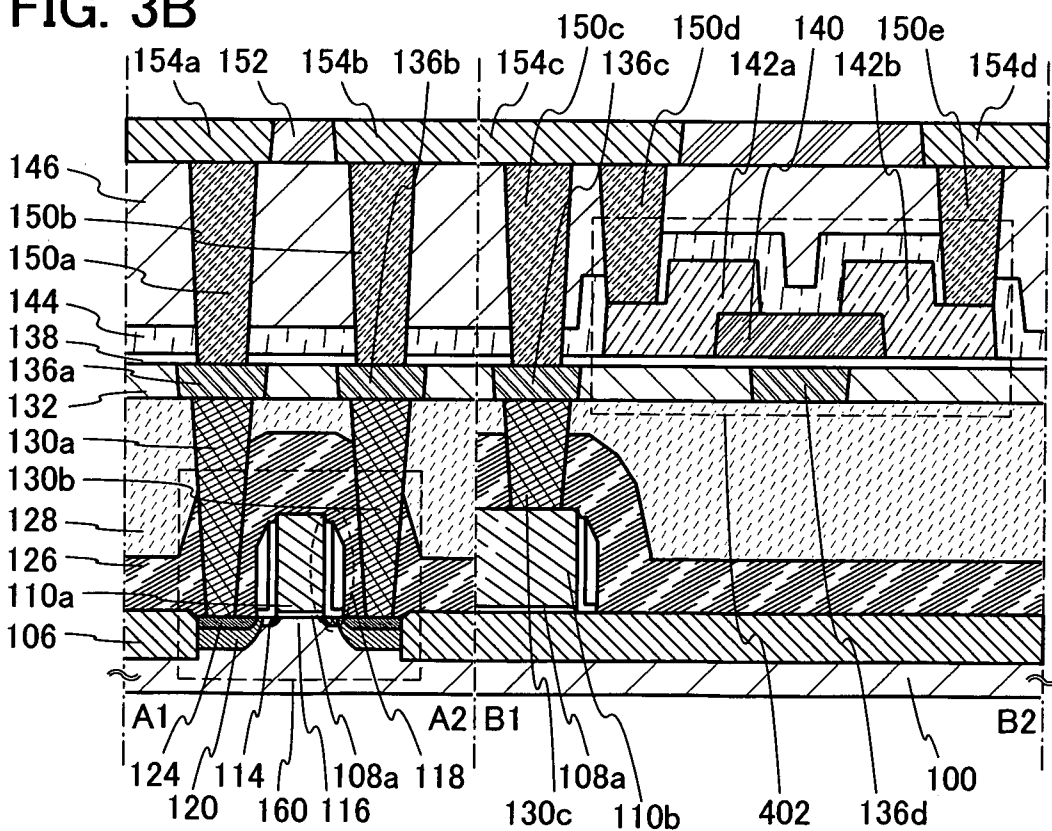


FIG. 4A

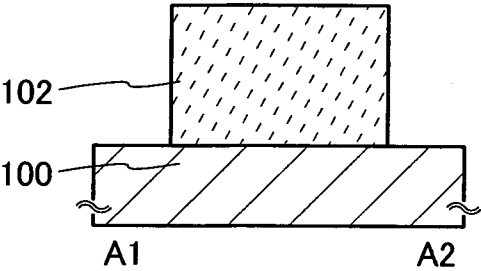


FIG. 4E

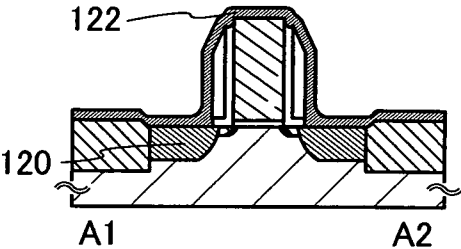


FIG. 4B

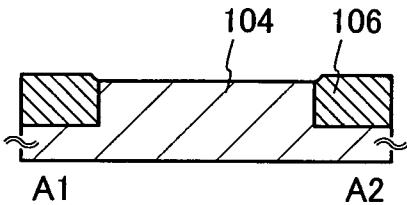


FIG. 4F

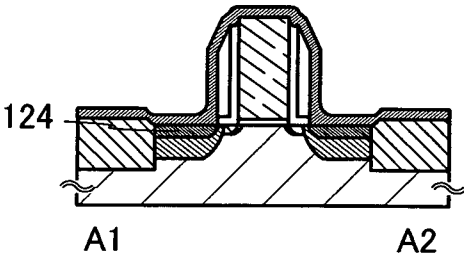


FIG. 4C

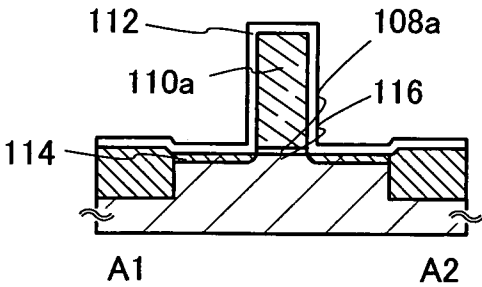


FIG. 4G

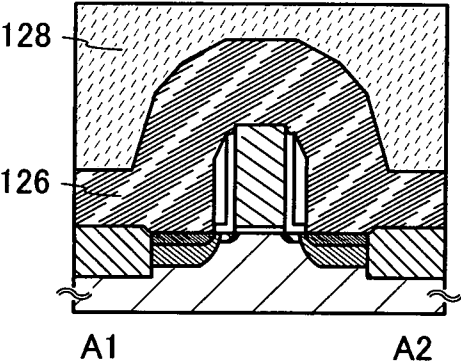


FIG. 4D

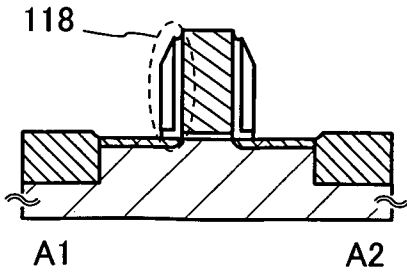
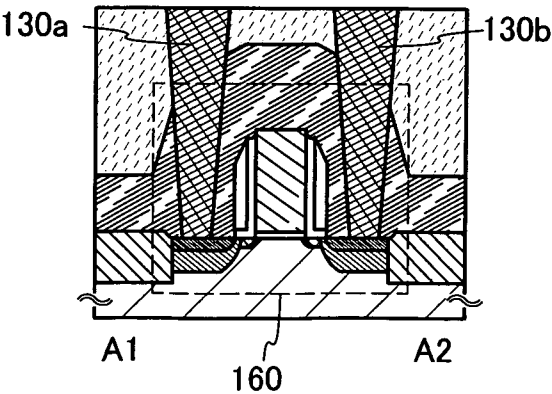


FIG. 4H



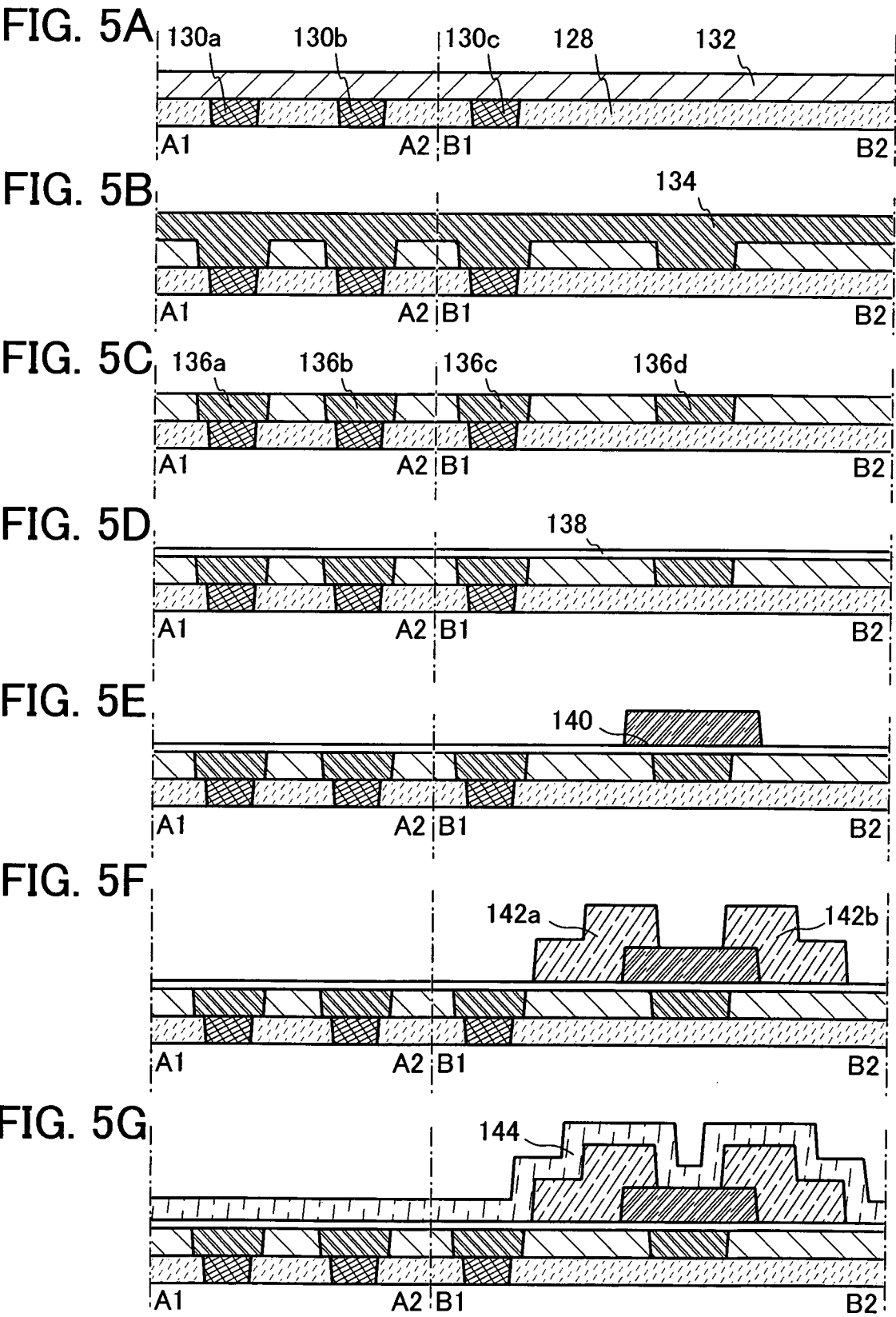


FIG. 6A

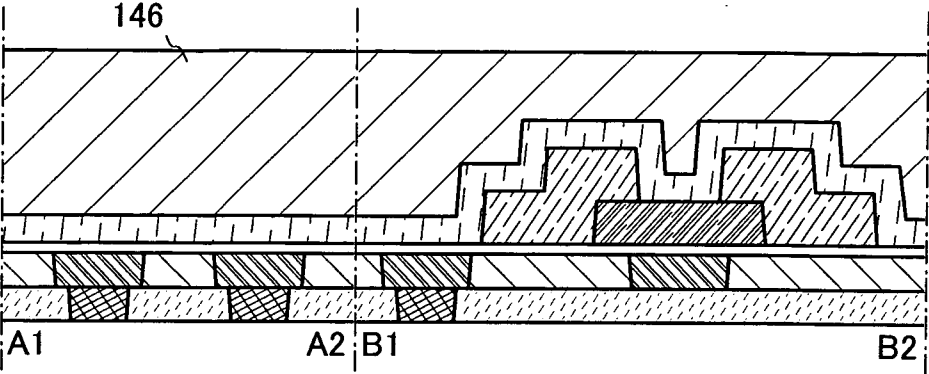


FIG. 6B

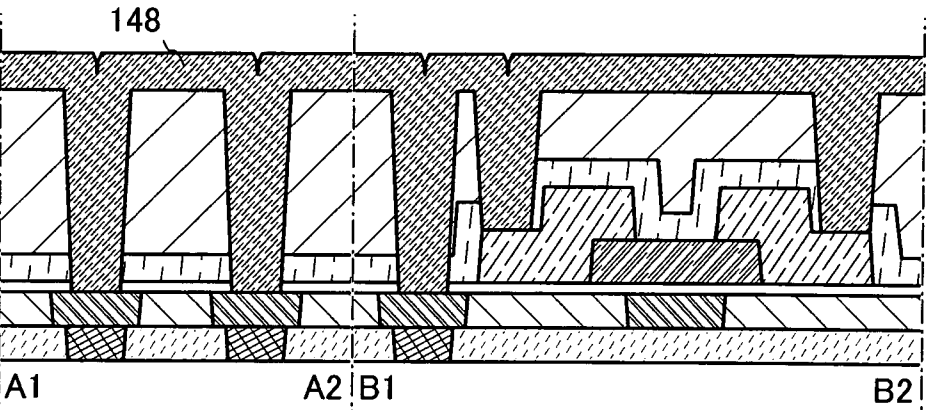


FIG. 6C

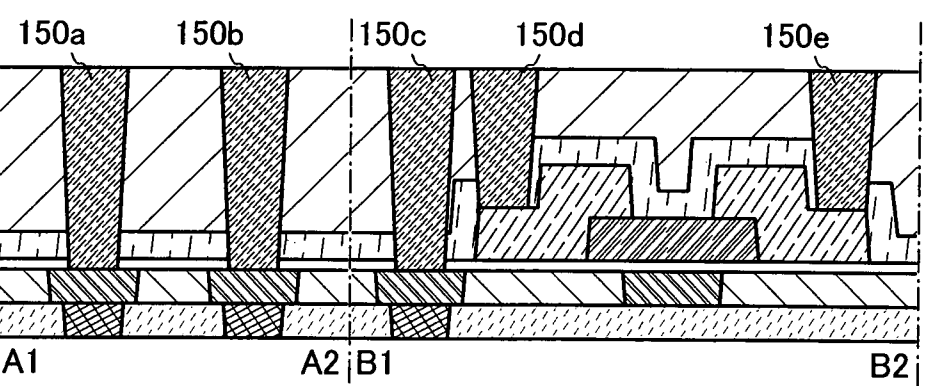


FIG. 6D

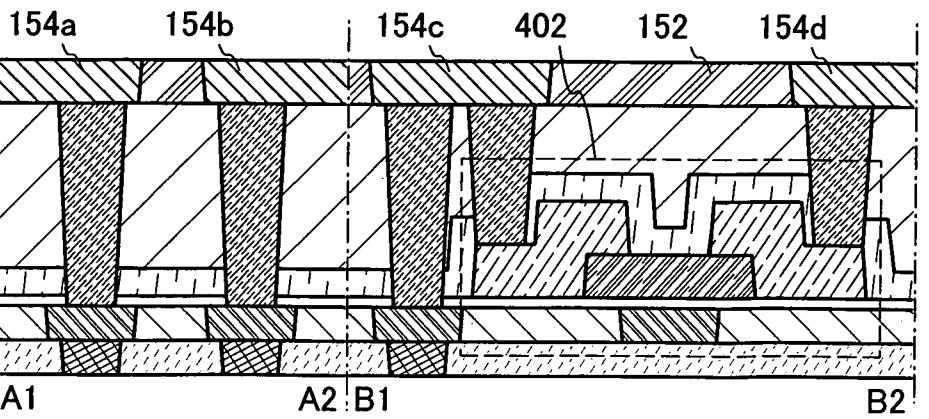
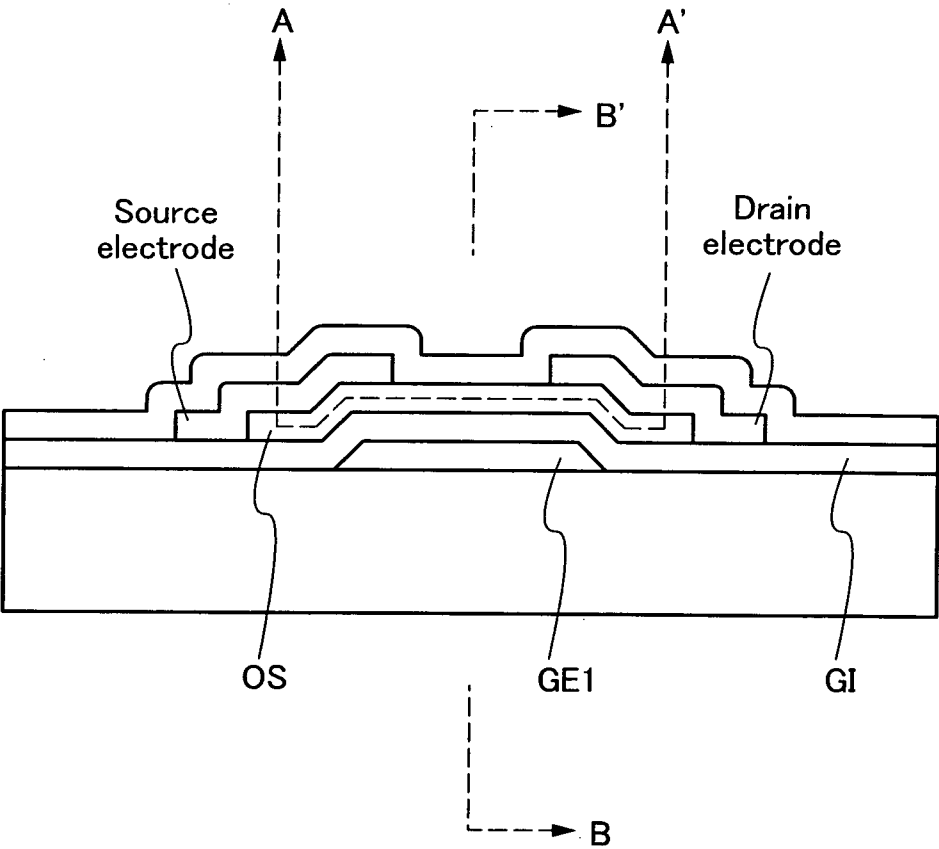
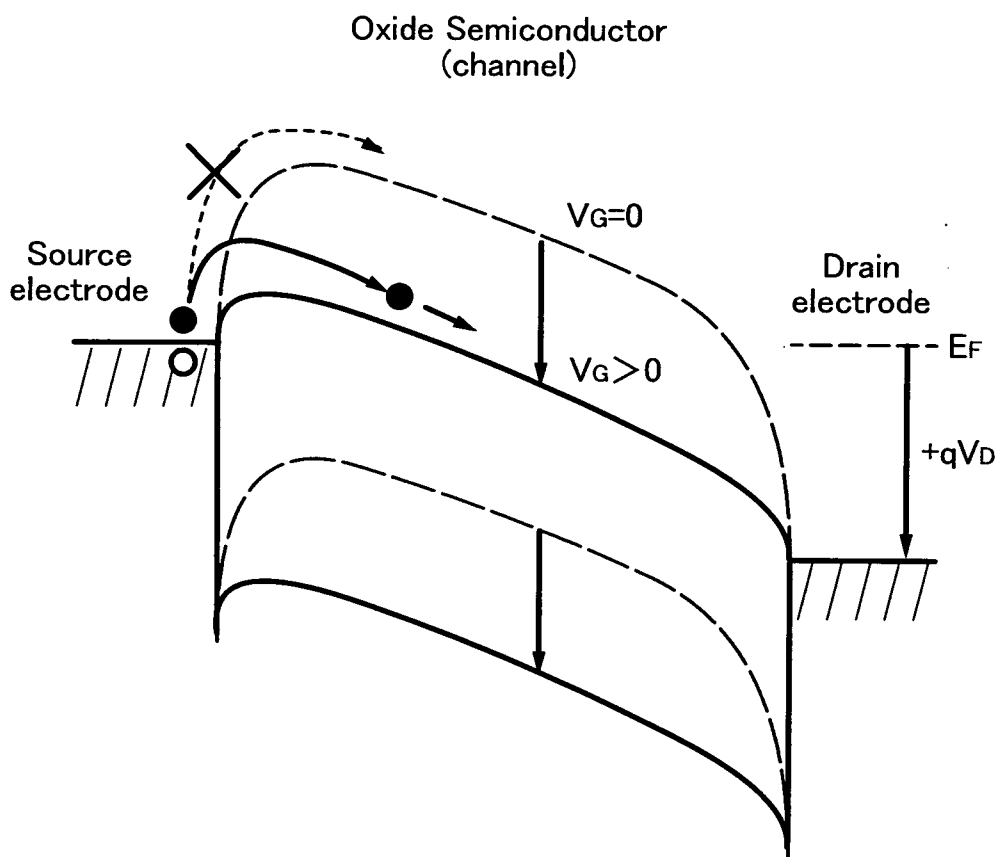


FIG.7



8/31

FIG. 8



9/31

FIG. 9A

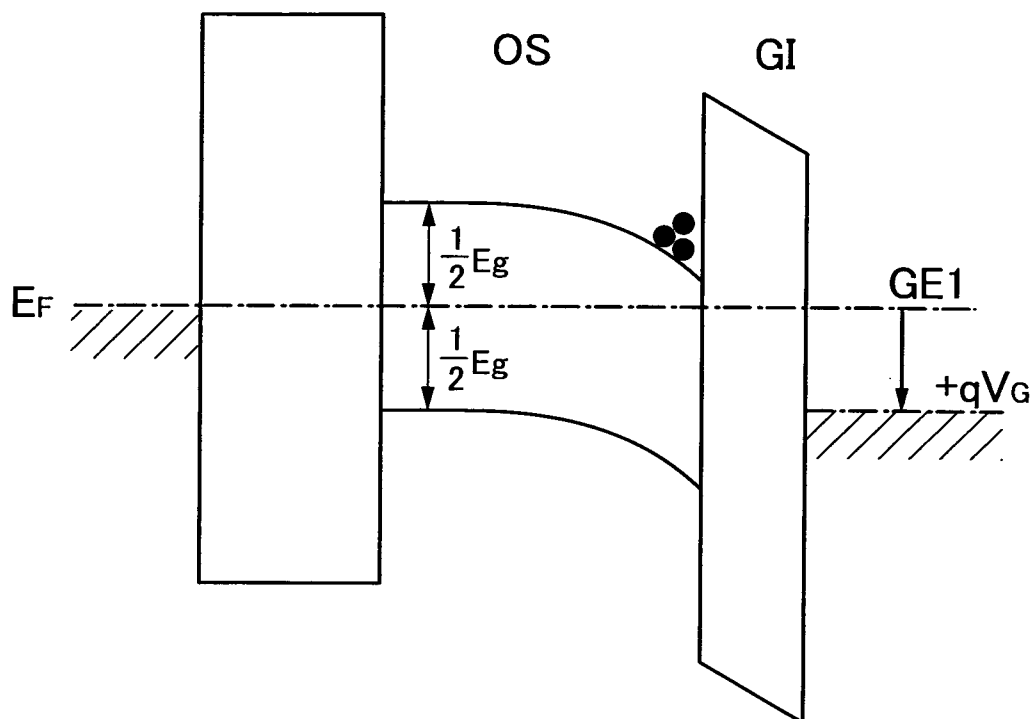
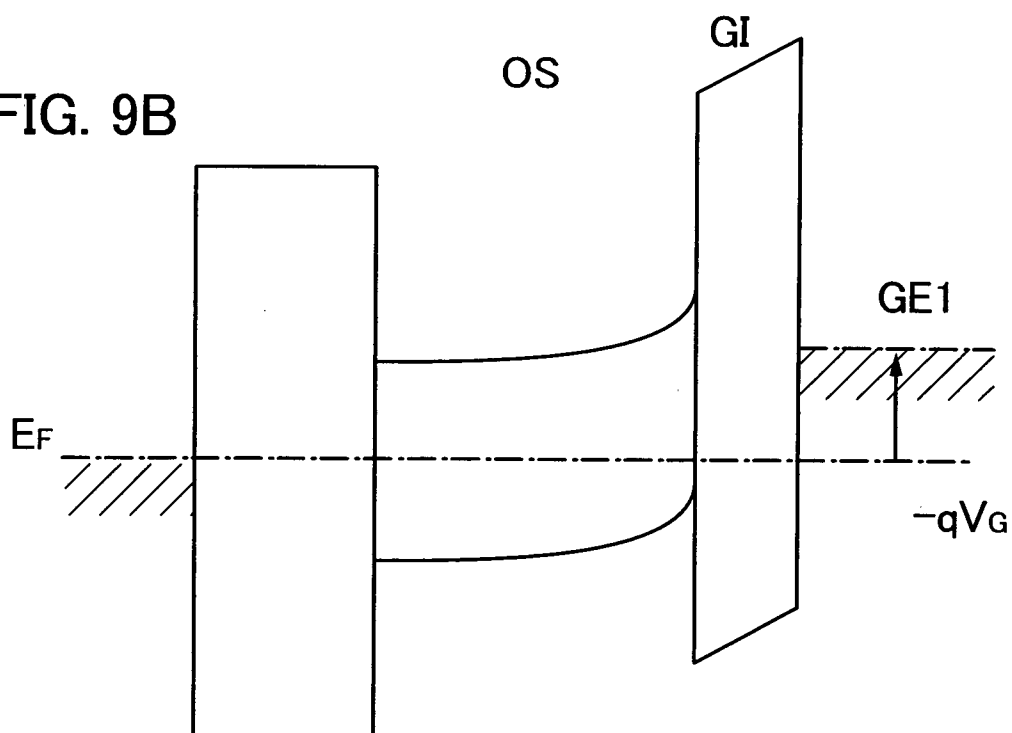
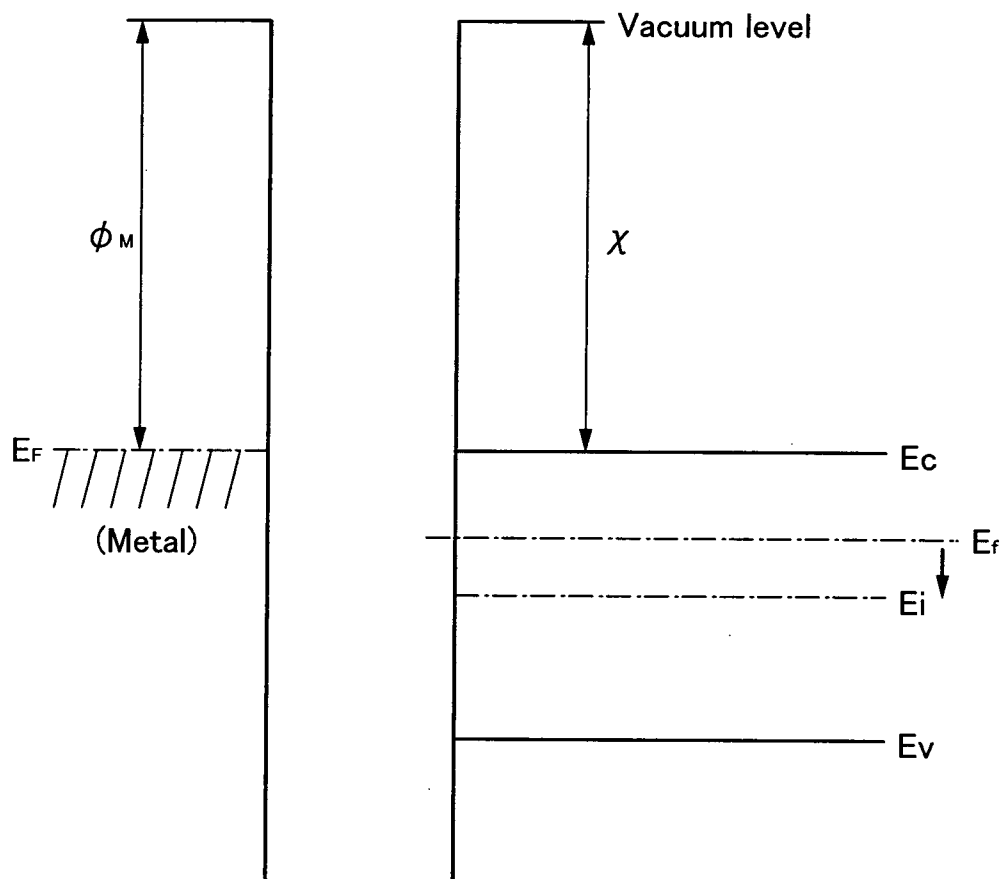
Energy band diagram of B-B' section ($V_G > 0$)

FIG. 9B

Energy band diagram of B-B' section ($V_G < 0$)

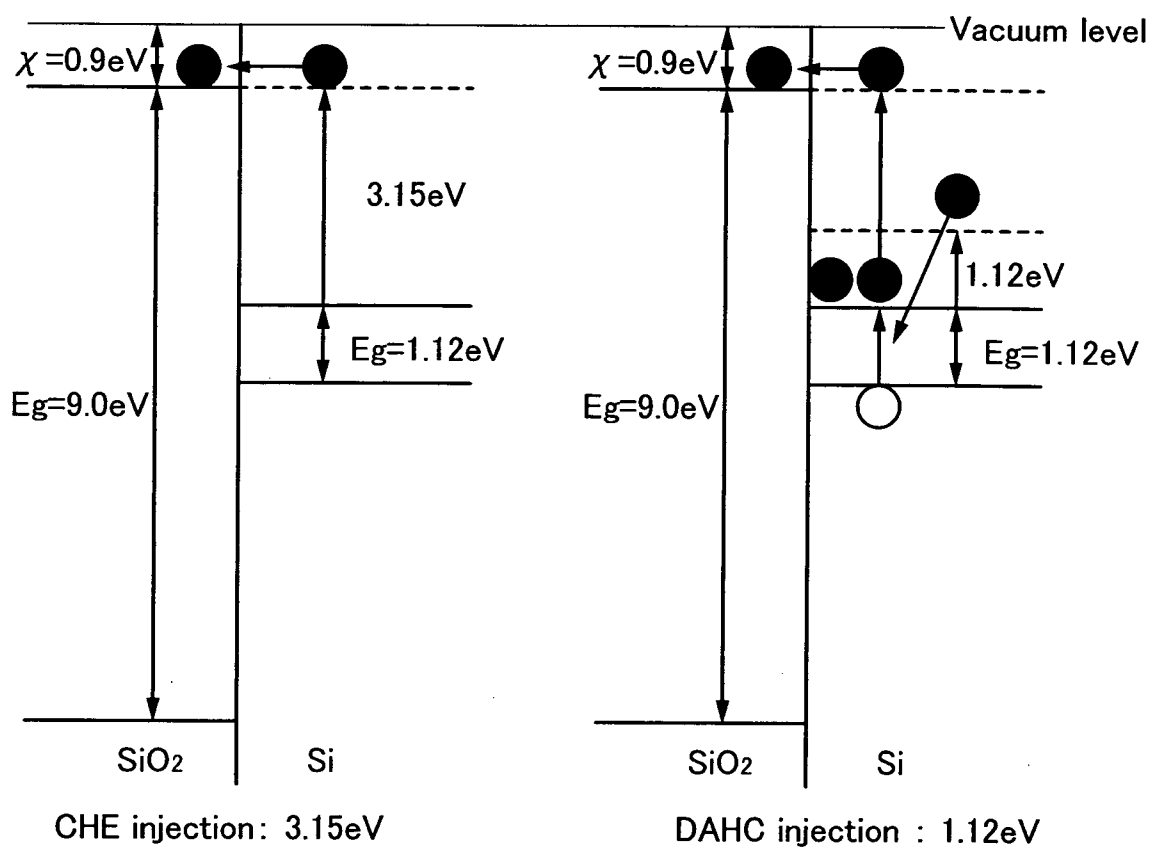
10/31

FIG. 10



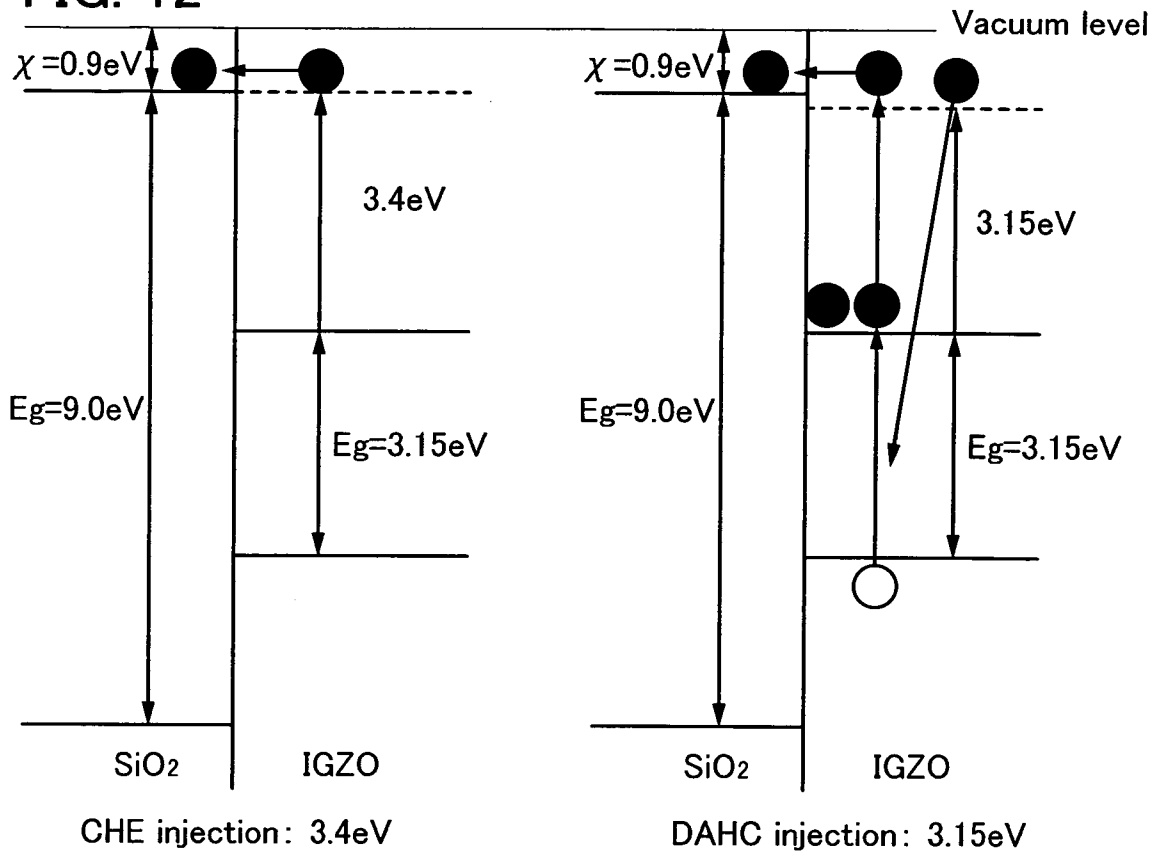
11/31

FIG. 11



12/31

FIG. 12



13/31

FIG. 13

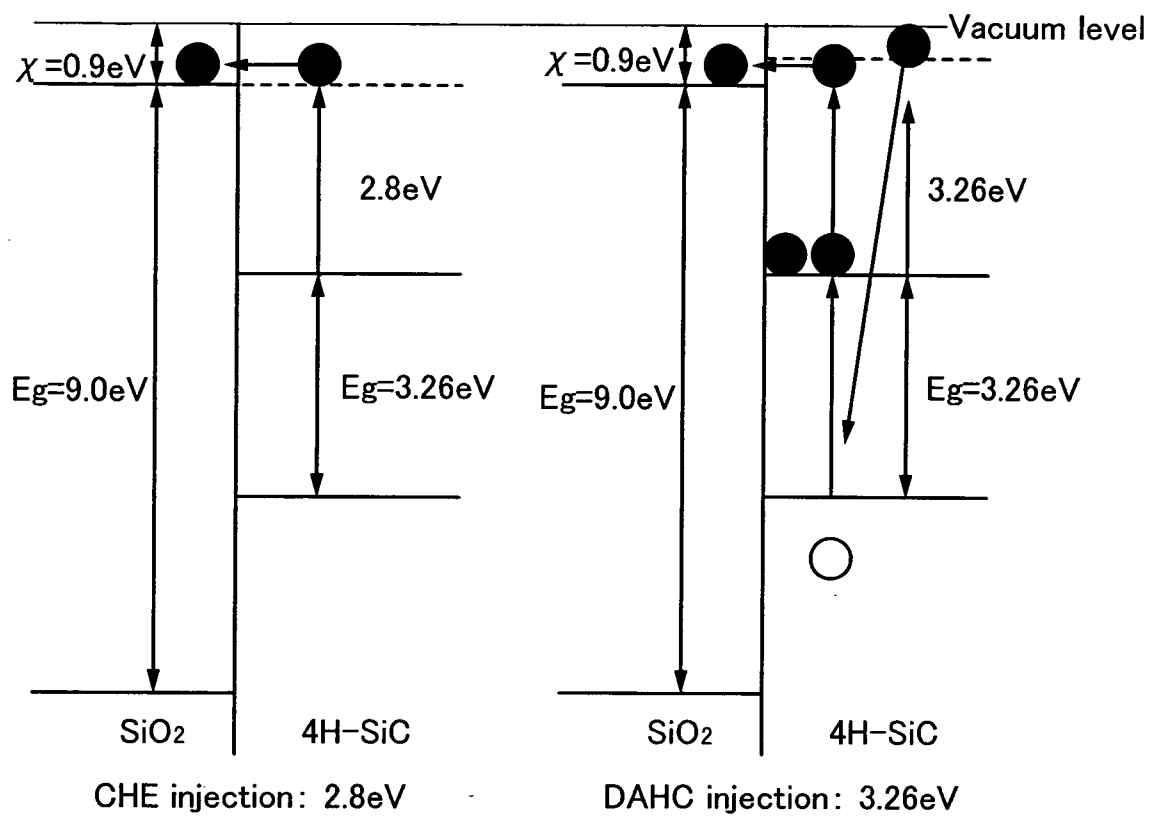


FIG. 14

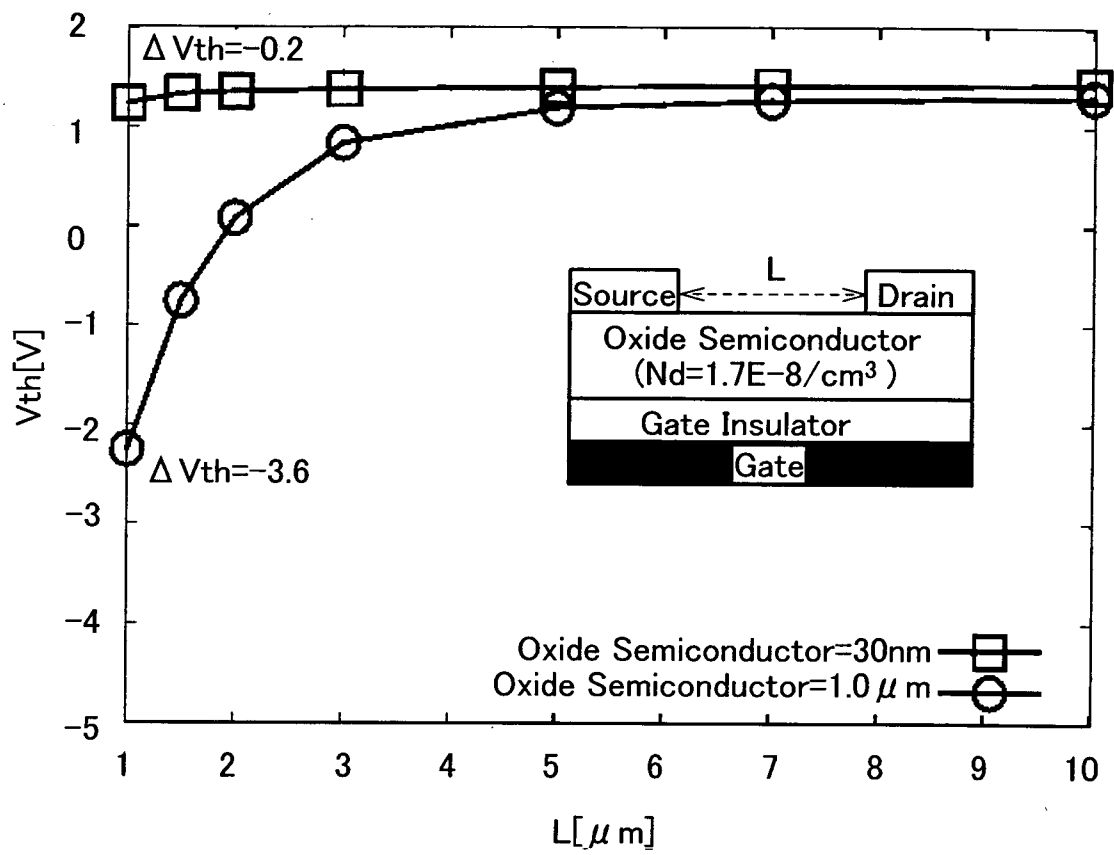


FIG. 15

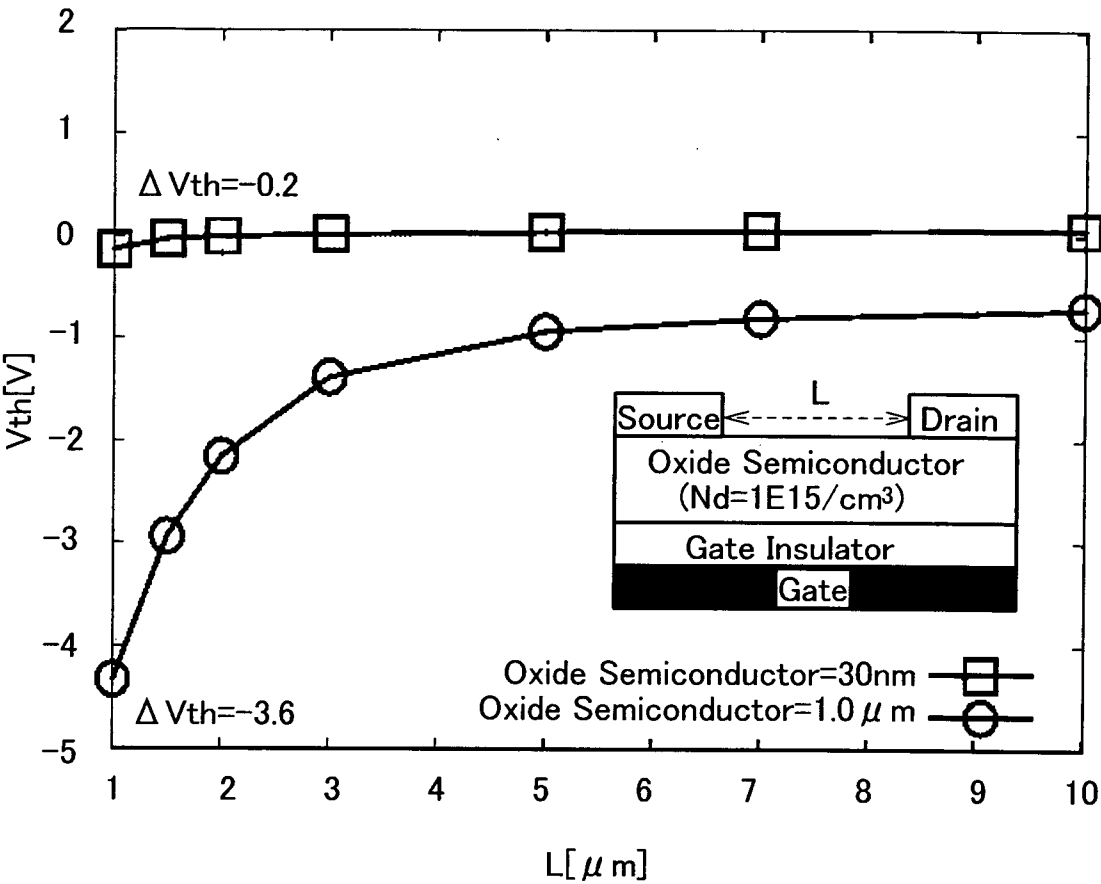


FIG. 16

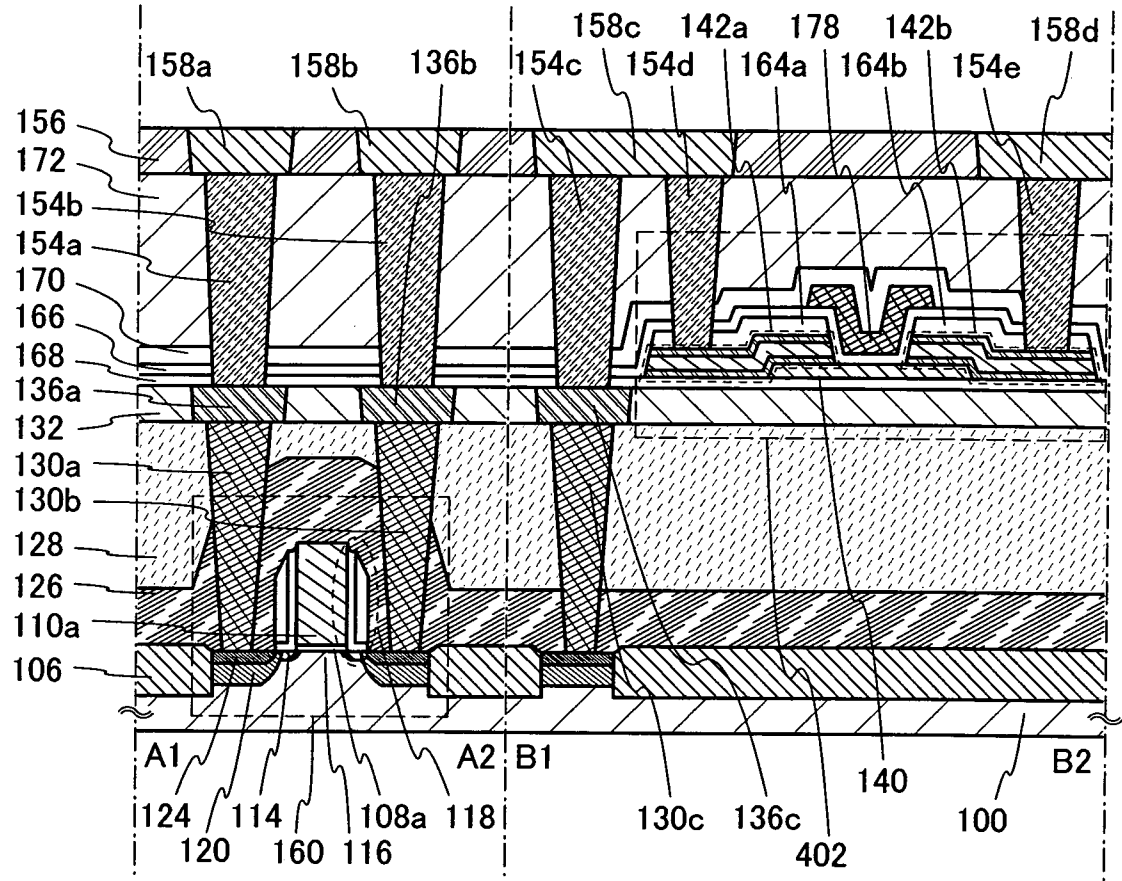


FIG. 17A

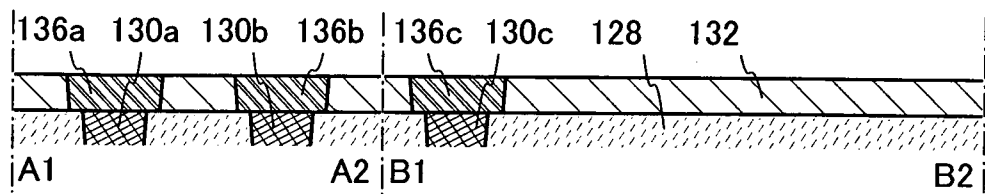


FIG. 17B

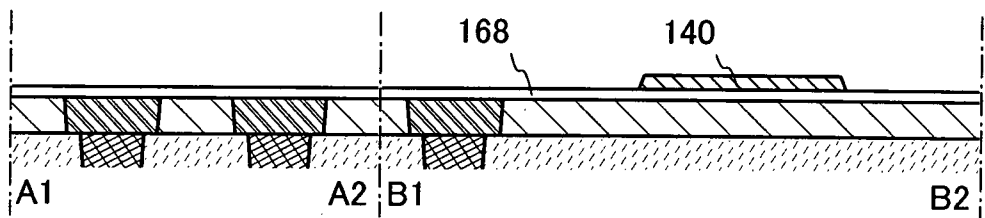


FIG. 17C

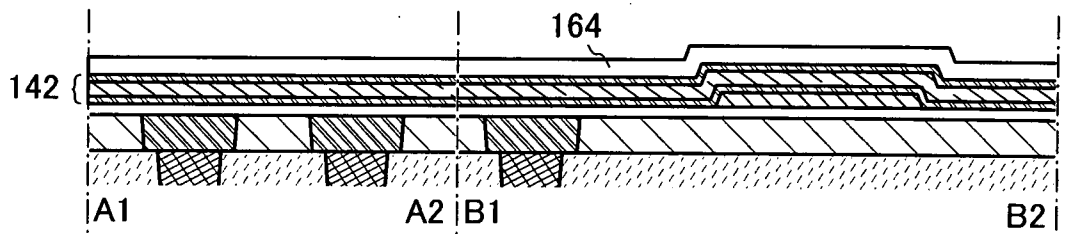


FIG. 17D

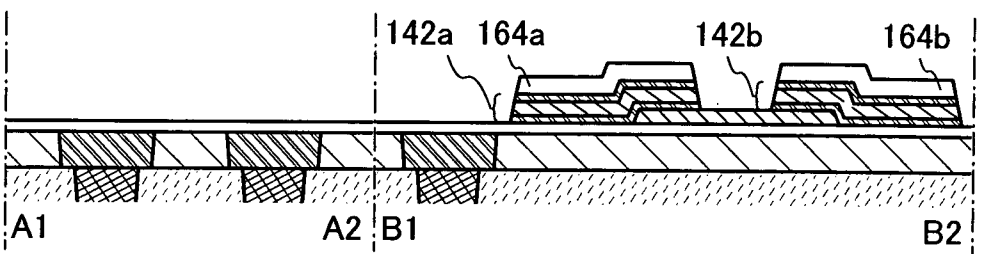


FIG. 17E

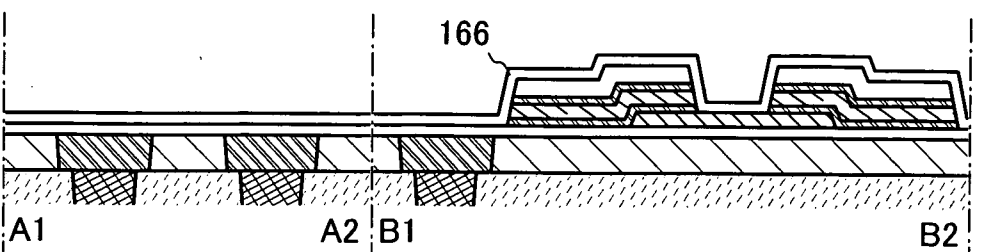


FIG. 18A

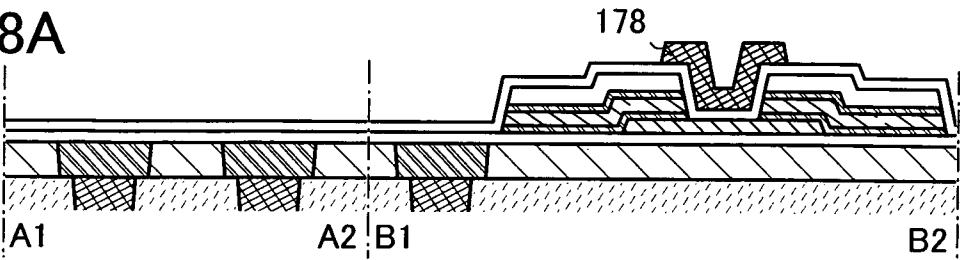


FIG. 18B

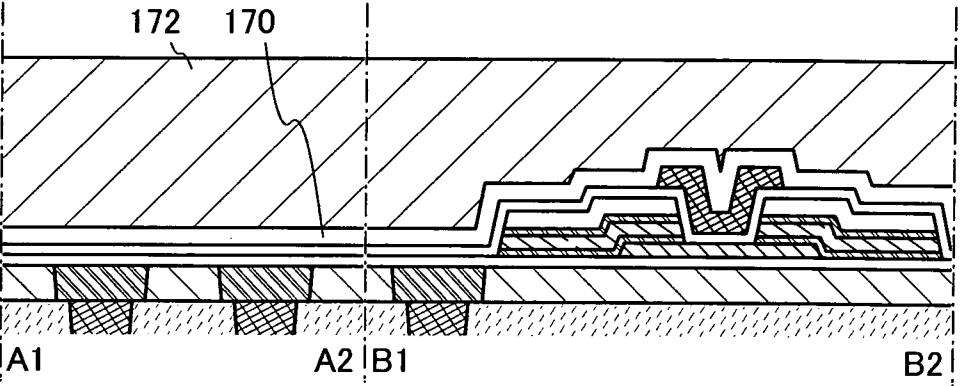


FIG. 18C

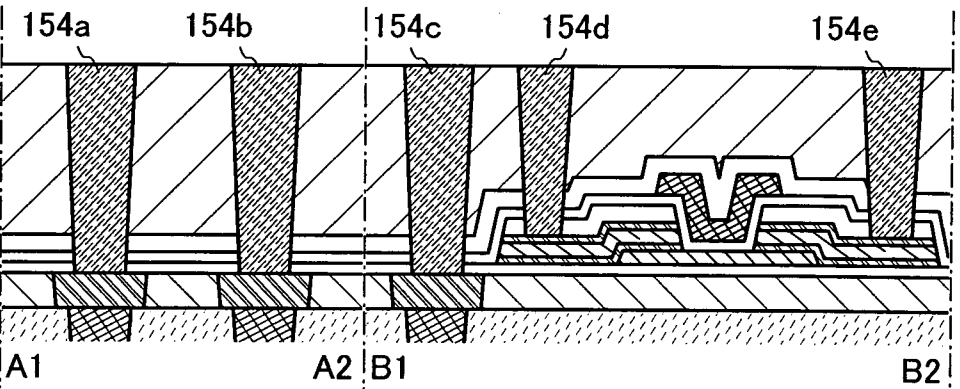
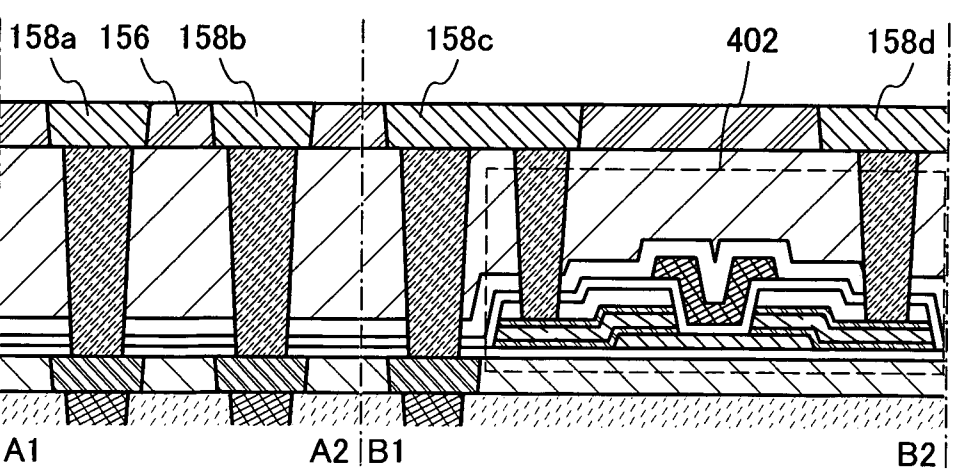


FIG. 18D



19/31

FIG. 19A

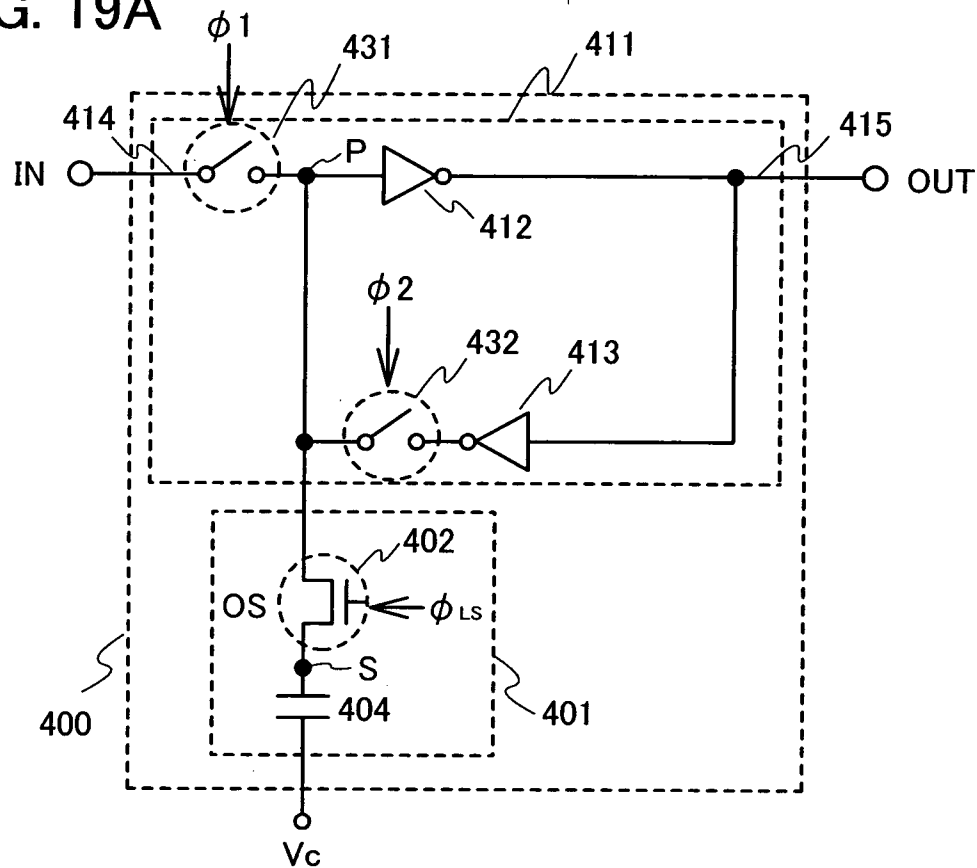


FIG. 19B

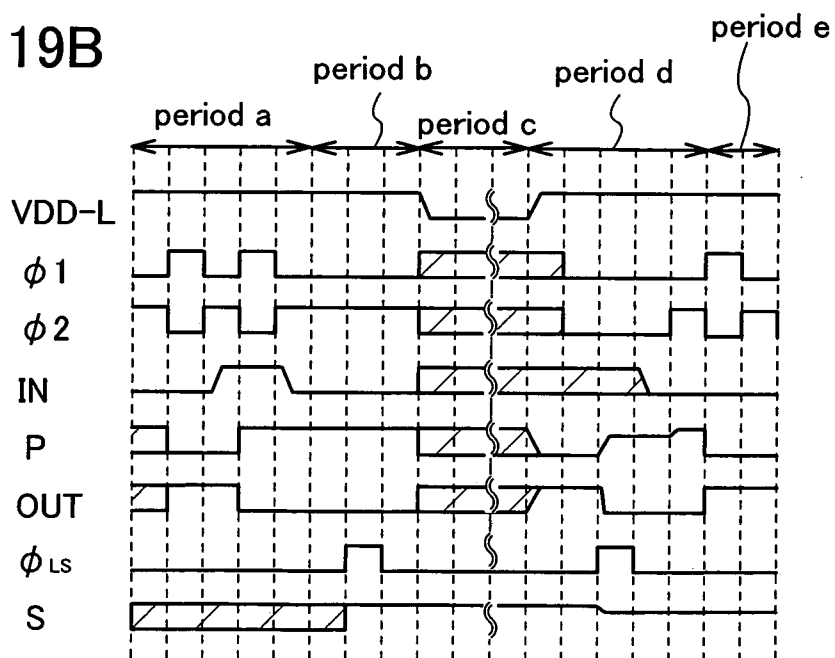


FIG. 20A

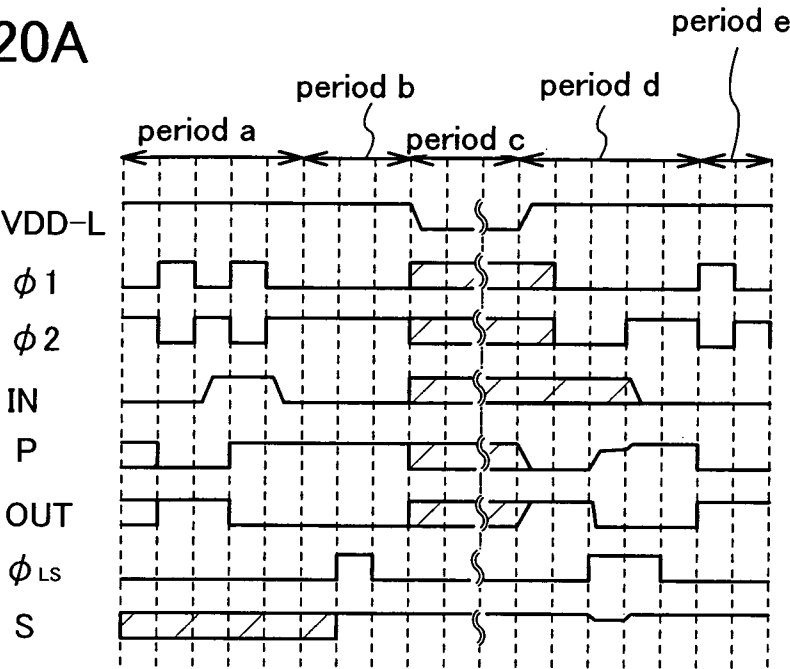


FIG. 20B

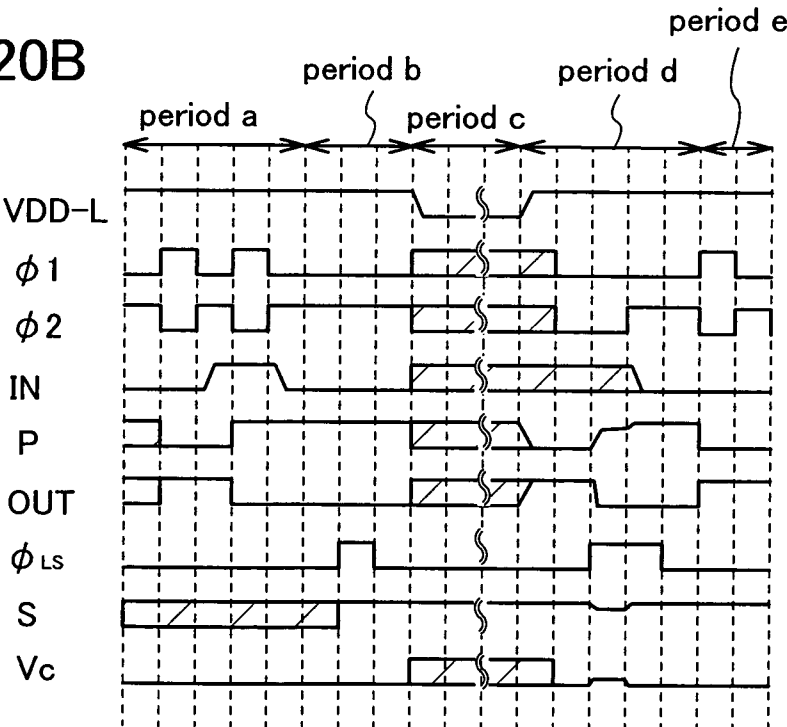


FIG. 21

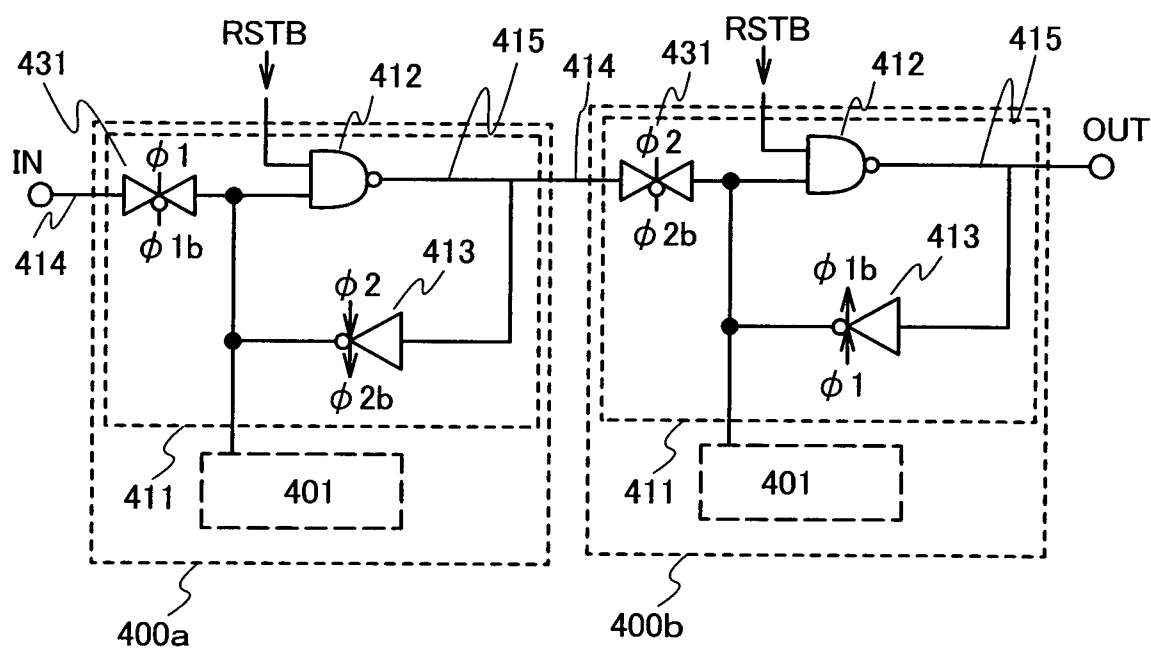
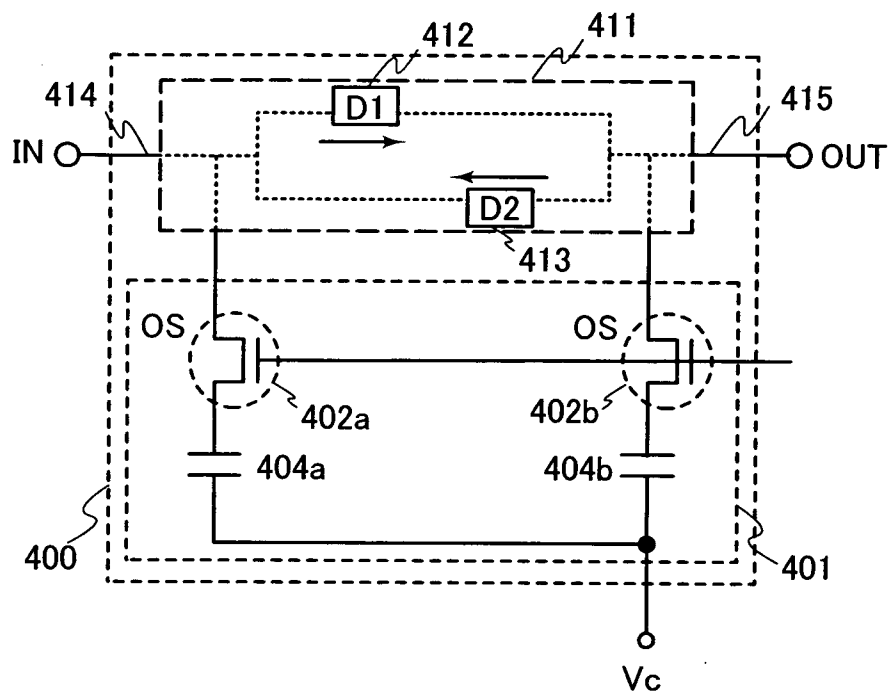


FIG. 22



23/31

FIG. 23

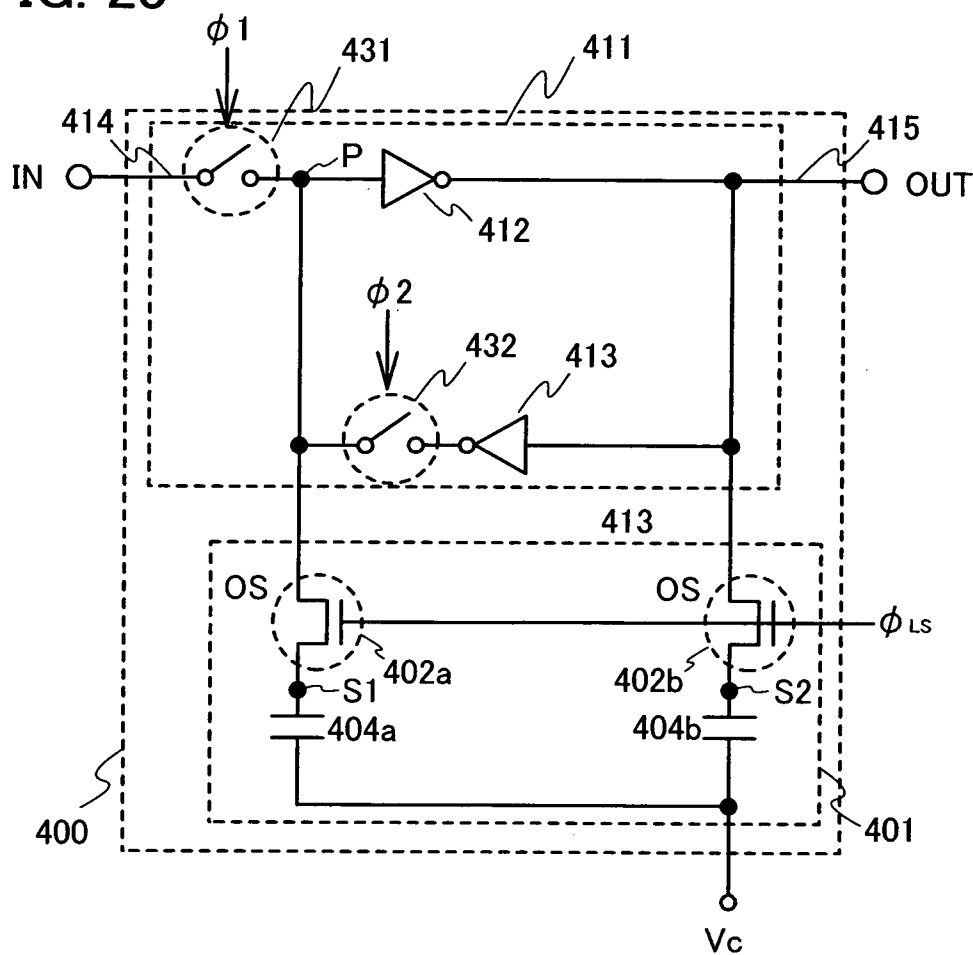


FIG. 24A

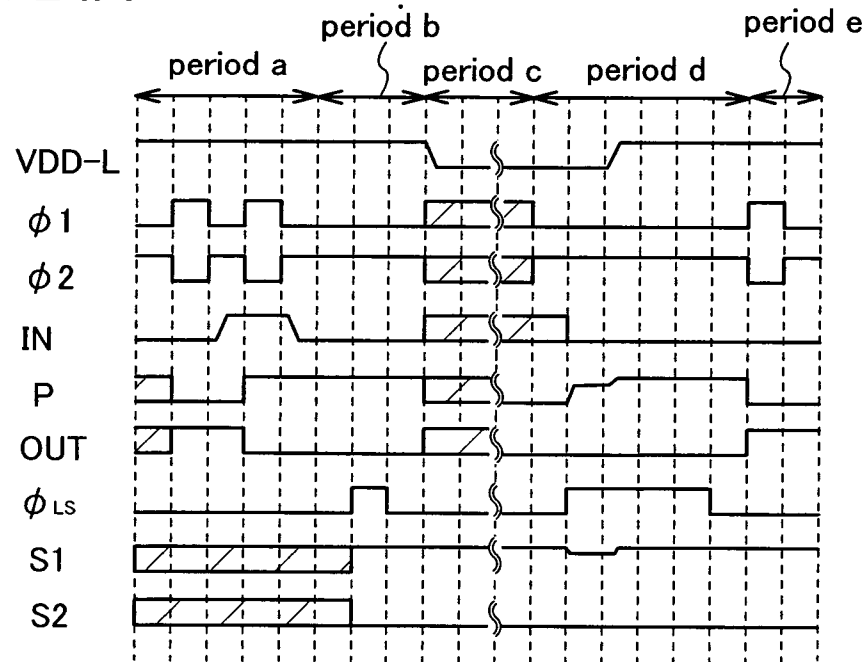


FIG. 24B

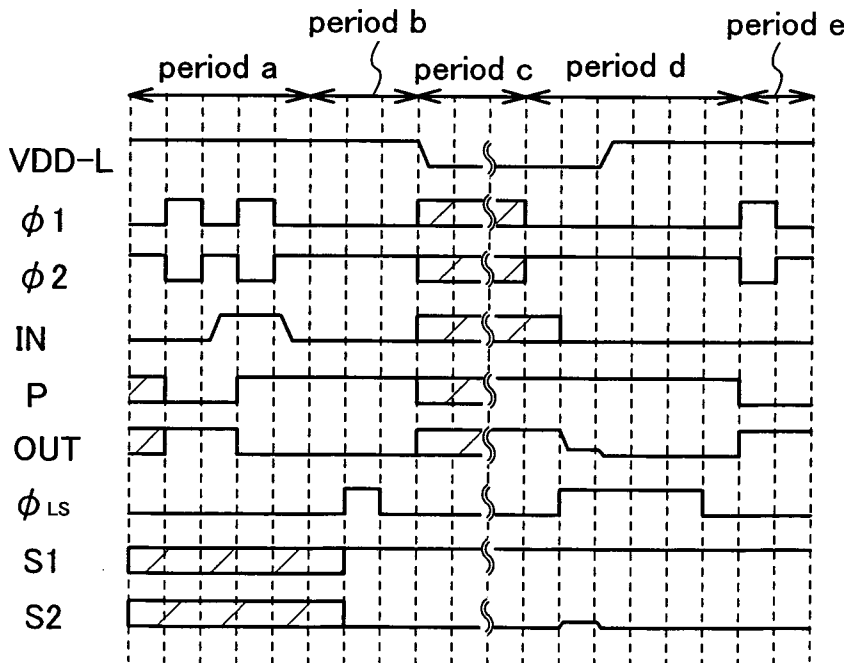


FIG. 25

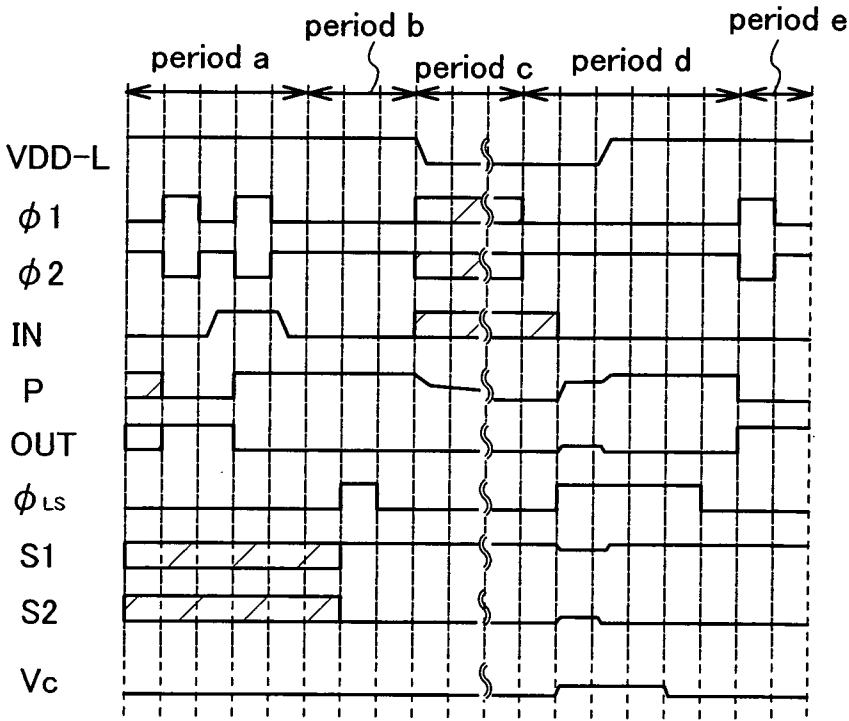
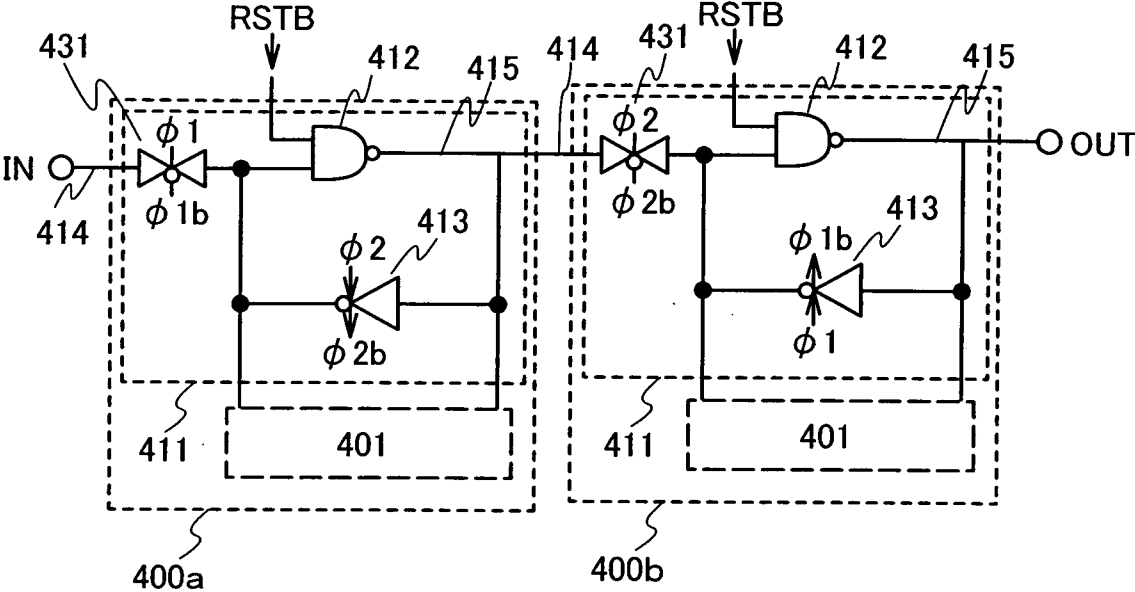


FIG. 26



27/31

FIG. 27A

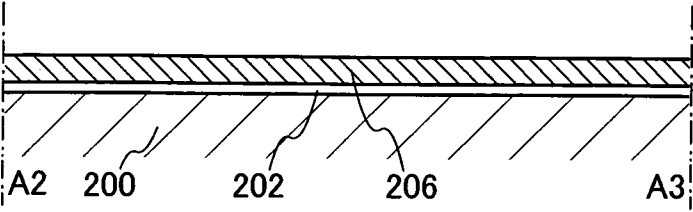


FIG. 27B

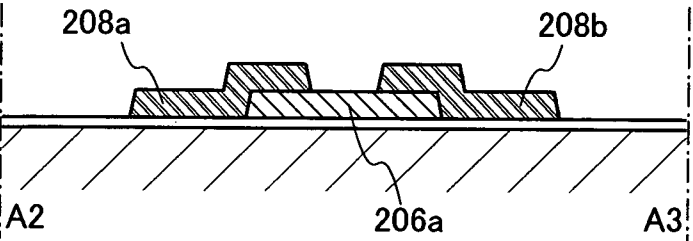


FIG. 27C

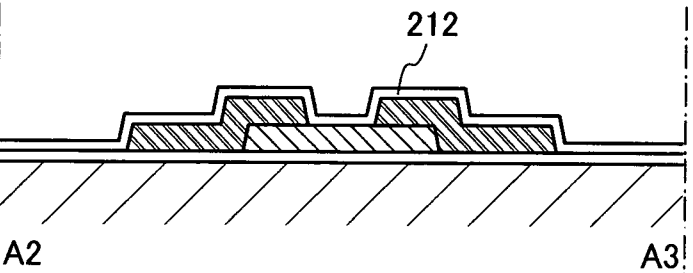


FIG. 27D

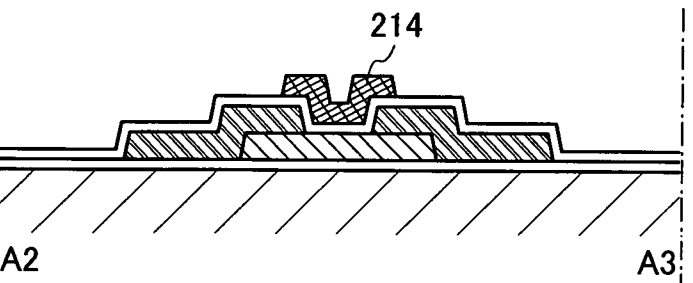
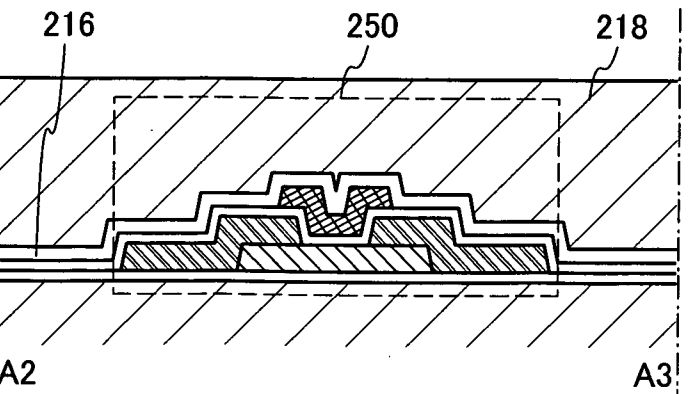


FIG. 27E



28/31

FIG. 28A

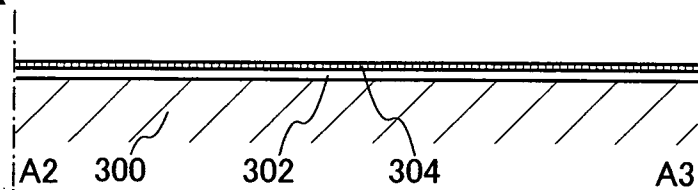


FIG. 28B

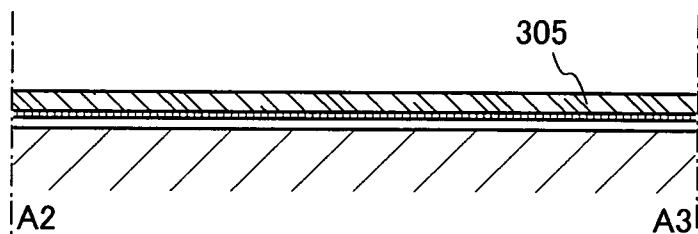


FIG. 28C

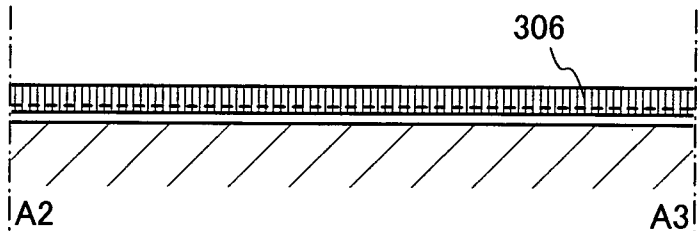


FIG. 28D

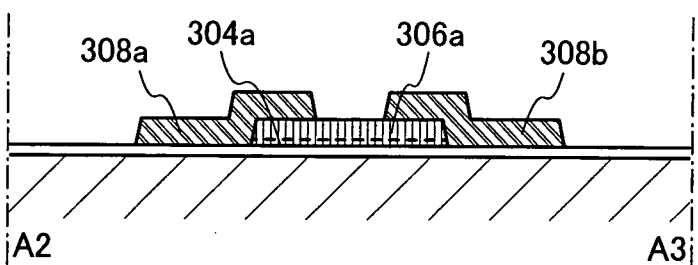
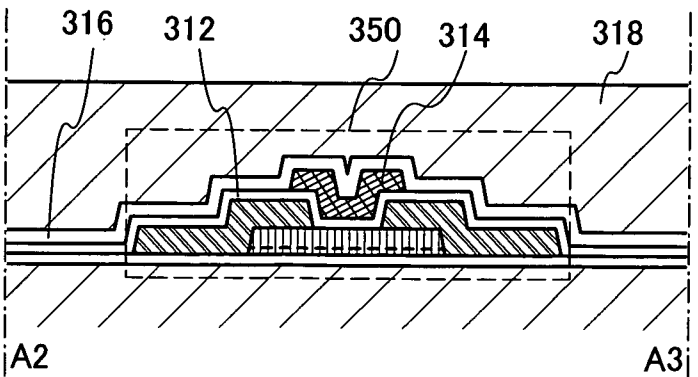


FIG. 28E



29/31

FIG. 29A

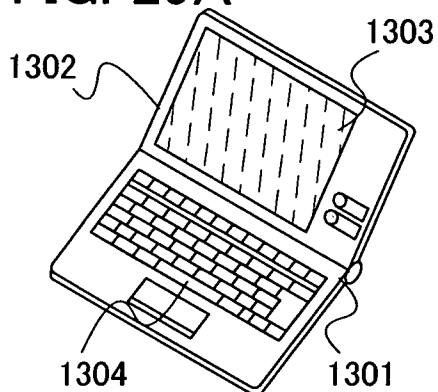


FIG. 29B

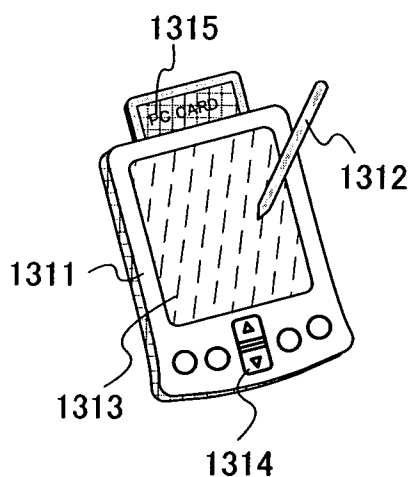


FIG. 29C

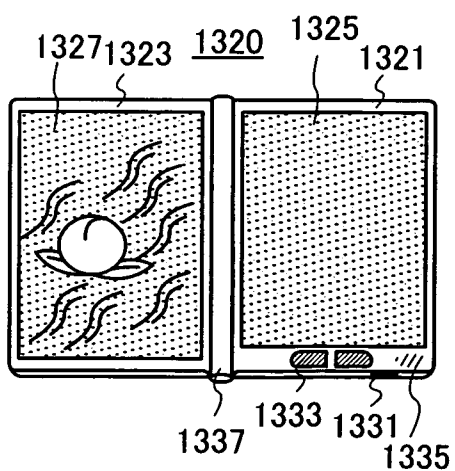


FIG. 29D

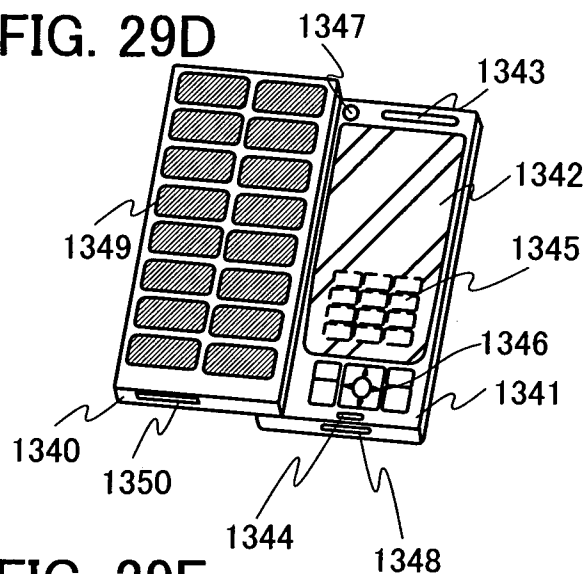


FIG. 29E

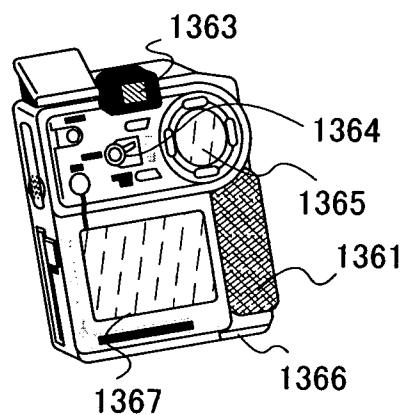
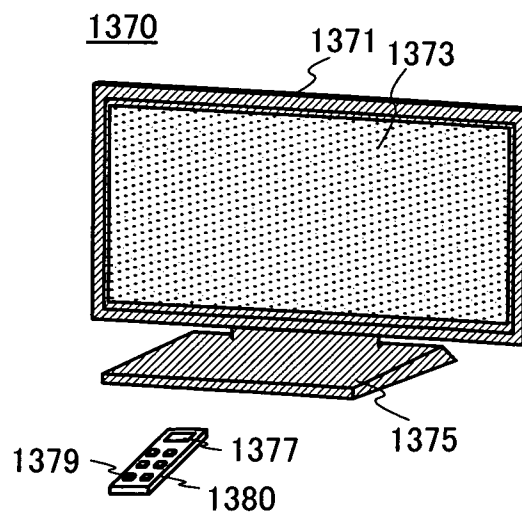


FIG. 29F



EXPLANATION OF REFERENCE

100: substrate, 102: protective layer, 104: semiconductor region, 106: element isolation insulating layer, 108a: gate insulating layer, 110a: gate electrode, 110b: electrode, 112: gate insulating layer, 114: impurity regions, 116: channel formation region, 118: sidewall insulating layer, 120: high-concentration region, 122: metal layer, 124: metal compound region, 126: interlayer insulating layer, 128: interlayer insulating layer, 130a: source or drain electrode, 130b: source or drain electrode, 130c: electrode, 132: insulating layer, 134: conductive layer, 136a: electrode, 136b: electrode, 136c: electrode, 136d: gate electrode, 138: gate insulating layer, 140: oxide semiconductor layer, 142: conductive layer, 142a: source or drain electrode, 142b: source or drain electrode, 144: protective insulating layer, 146: interlayer insulating layer, 148: conductive layer, 150a: electrode, 150b: electrode, 150c: electrode, 150d: electrode, 150e: electrode, 152: insulating layer, 154a: electrode, 154b: electrode, 154c: electrode, 154d: electrode, 154e: electrode, 156: insulating layer, 158a: electrode, 158b: electrode, 158c: electrode, 158d: electrode, 160: transistor, 164: insulating layer, 164a: insulating layer, 164b: insulating layer, 166: gate insulating layer, 168: insulating layer, 170: interlayer insulating layer, 172: interlayer insulating layer, 178: gate electrode, 200: bottom substrate, 202: insulating layer, 206: oxide semiconductor layer, 206a: oxide semiconductor layer, 208a: source or drain electrode, 208b: source or drain electrode, 212: gate insulating layer, 214: gate electrode, 216: interlayer insulating layer, 218: interlayer insulating layer, 250: transistor, 300: bottom substrate, 302: insulating layer, 304: oxide semiconductor layer, 304a: oxide semiconductor

31/31

layer, 305: oxide semiconductor layer, 306: oxide semiconductor layer, 306a: oxide semiconductor layer, 308a: source or drain electrode, 308b: source or drain electrode, 312: gate insulating layer, 314: gate electrode, 316: interlayer insulating layer, 318: interlayer insulating layer, 350: transistor, 400: nonvolatile latch circuit, 400a: nonvolatile latch circuit, 400b: nonvolatile latch circuit, 401: data holding portion, 402: transistor, 402a: transistor, 402b: transistor, 404: capacitor, 404a: capacitor, 404b: capacitor, 411: latch portion, 412: first element, 413: second element, 414: wiring, 415: wiring, 431: switch, 432: switch, 1301: main body, 1302: housing, 1303: display portion, 1304: keyboard, 1311: main body, 1312: stylus, 1313: display portion, 1314: operation keys, 1315: external interface, 1320: e-book reader, 1321: housing, 1323: housing, 1325: display portion, 1327: display portion, 1331: power button, 1333: operation keys, 1335: speaker, 1337: hinge, 1340: housing, 1341: housing, 1342: display panel, 1343: speaker, 1344: microphone, 1345: operation keys, 1346: pointing device, 1347: camera lens, 1348: external connection terminal, 1349: solar cell, 1350: external memory slot, 1361: main body, 1363: eyepiece portion, 1364: operation switch, 1365: display portion B, 1366: battery, 1367: display portion A, 1370: television set, 1371: housing, 1373: display portion, 1375: stand, 1377: display portion, 1379: operation keys, 1380: remote controller.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2010/070753

A. CLASSIFICATION OF SUBJECT MATTER

Int.Cl. See extra sheet

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Int.Cl. H01L21/8242, H01L21/8234, H01L27/08, H01L27/088, H01L27/10, H01L27/108, H01L29/786

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Published examined utility model applications of Japan 1922-1996
 Published unexamined utility model applications of Japan 1971-2011
 Registered utility model specifications of Japan 1996-2011
 Published registered utility model applications of Japan 1994-2011

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	JP 5-110392 A (Hitachi Ltd.) 1993.04.30, [0045]-[0051], Fig.1 (No Family)	1-29
Y	JP 7-147530 A (Mitsubishi Electric Corporation) 1995.06.06, [0055]-[0063], Fig.12-14 (No Family)	1-29
Y	JP 1-66899 A (Ramtron Corporation) 1989.03.13, page4-7, Fig.1A & US 4809225 A & EP 297777 A2 & EP 530928 A2 & DE 3855877 T & DE 3886600 T & DE 297777 T	15-27, 29



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search

15.02.2011

Date of mailing of the international search report

01.03.2011

Name and mailing address of the ISA/JP

Japan Patent Office

3-4-3, Kasumigaseki, Chiyoda-ku, Tokyo 100-8915, Japan

Authorized officer

Shigeki Komori

Telephone No. +81-3-3581-1101 Ext. 3462

4M

4 1 4 5

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2010/070753

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
	& DE 3855877 D & DE 3886600 D & AU 1556388 A & AU 588505 B	
Y	JP 2006-165532 A (Canon Inc.) 2006.06.22, [0043]-[0045],[0255]-[0259],[0289]-[0296], Fig.5-6,24-25 & US 2006/0113565 A1	1-29
P,A	JP 2010-34710 A (NEC Electronics Corporation) 2010.02.12, Full text, all drawings & US 2010/0019839 A & US 2010/0019839 A1	1-29

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2010/070753

CLASSIFICATION OF SUBJECT MATTER

H01L21/8242 (2006.01) i, H01L21/8234 (2006.01) i, H01L27/08 (2006.01) i,
H01L27/088 (2006.01) i, H01L27/10 (2006.01) i, H01L27/108 (2006.01) i,
H01L29/786 (2006.01) i