

PATENT SPECIFICATION (11)

1 332 384

DRAWINGS ATTACHED



- (21) Application No. 28758/71 (22) Filed 18 June 1971 (19)
 (31) Convention Application No. 78806 (32) Filed 7 Oct. 1970 in
 (33) United States of America (US)
 (44) Complete Specification published 3 Oct. 1973
 (51) International Classification H01L 11/14
 (52) Index at acceptance

H1K 217 223 22Y 277 288 306 312 353 413 415 418 41Y
 422 42X 441 458 472 474 482 507 51Y 520 530
 54Y 551 55Y 578 579 581 599 611 615 618 61Y
 636 637 638 63X 650

(54) FABRICATION OF SEMICONDUCTOR DEVICES

(71) We, RCA CORPORATION, a corporation organised under the laws of the State of Delaware, United States of America, of 30 Rockefeller Plaza, City and State of New York, United States of America, do hereby declare the invention, for which we pray that a patent may be granted to us, and the method by which it is to be performed, to be particularly described in and by the following statement:—

This invention relates to the fabrication of insulated gate field effect transistors.

One type of insulated gate, field effect transistor comprises a body of semiconductor material having, at one surface thereof, source and drain regions of one type conductivity separated by a channel region of the other type conductivity. Covering the surface of the body over the channel region is a gate insulator of a dielectric material covered, in turn, by a gate electrode. Preferably, for optimum device performance, the gate electrode should be exactly aligned with the channel region, neither extending over nor falling short of the source and drain regions.

One process for fabricating such devices comprises providing a layer of a dielectric material on a surface of a body of doped semiconductor material and providing a gate electrode on the dielectric layer. Then, using the gate electrode as a mask, the dielectric layer is defined to provide the gate insulator, and conductivity modifiers are caused to diffuse into the semiconductor body on either side of the composite electrode-insulator gate structure to provide the source and drain regions. The portion of the semiconductor body covered by the gate structure is masked from the conductivity modifiers and comprises a channel region which is aligned with the gate electrode.

One problem with the above described process is that while the gate structure is used as a mask in the doping of the source and drain regions, thus providing substantial alignment of the channel region with the gate electrode, some lateral diffusion of the

conductivity modifiers under the gate structure does occur. This lateral diffusion results in an extension of the source and drain regions underneath the gate electrode and gives rise to an undesired amount of overlap between the gate electrode and the source and drain regions.

According to the invention, we provide a method of fabricating insulated gate field effect transistors comprising providing on a surface of a body of doped semiconductor material a gate structure comprising a first layer of a dielectric material covered by a second layer of heavily doped semiconductor material covering at least the lateral edges of said second layer with a protective coating, and thereafter, using said gate structure as a mask, driving conductivity modifiers into said body through portions of said surface.

In the accompanying drawings:

Figure 1 is a cross-sectional view of a semi-conductor material wafer workpiece at an early step in a device processing sequence in accordance with one embodiment of the instant invention;

Figure 2 is a cross-sectional view, on an enlarged scale, of a portion of the workpiece shown in Figure 1, but showing the workpiece at a later step in the processing sequence;

Figures 3, 4 and 5 are views similar to that of Figure 2 but showing successive steps in the processing sequence.

Figure 6 is a plan view of Figure 5; and

Figure 7 is a view similar to that of Figures 3, 4 and 5, but at a still later step in the processing sequence.

Shown in Fig. 1 is a body 10 of a doped semiconductor material, e.g. silicon, germanium, or the like. The body 10, in this embodiment, is a wafer of silicon cut from a silicon ingot in known fashion, the body 10 being doped with a P type conductivity modifier, e.g., boron, at a concentration of 1×10^{16} atoms/cm³. Covering one surface of the wafer 10 is a relatively thick, e.g. 10,000

5 Å, layer 11 of silicon dioxide having a number of openings 12 therethrough exposing surface portions 13 of the wafer 10. Semiconductor devices fabricated on the wafer 10, in accordance with the instant embodiment of the invention, are disposed within the openings 12, one device per opening. The purpose of the layer 11 is to provide electrical isolation of the various devices from one another.

10 Although only two openings 12 are shown, in actual practice, a much larger number of openings 12 are generally used.

15 Known techniques can be used to provide the workpiece shown in Figure 1.

Shown in Figure 2 is the bottom of one of the openings 12 (Figure 1) after the performance of certain steps on the workpiece. Thus, covering the previously exposed surface portion 13 of the body 70 is a first thin layer 14 of a dielectric material, e.g., an 800 Å thick layer of silicon dioxide, covered by a second layer 16 of a dielectric material which can serve as a protective mask for the first layer 14, e.g., a 250 Å thick layer of silicon nitride.

Although the use of the protective layer 16 is preferred, this layer, as described hereinafter, can be omitted.

30 Covering the two layers 14 and 16 is a third, relatively thick layer 18, e.g., 1 micron, of a heavily doped semiconductor material, e.g., highly doped silicon or germanium, upon which an oxide of the material can be thermally grown. In this embodiment, the layer 18 comprises polycrystalline silicon of .001 ohm-cm P type conductivity. Layers 18 having other conductivity or type of conductivity characteristics can be used depending upon the particular device being fabricated.

45 Techniques for providing the workpiece shown in Figure 2 are known. The silicon dioxide layer 14, along with the silicon nitride layer 16, form, as described hereinafter, the gate insulator of the transistor to be formed. Accordingly, the layer 14 is preferably provided by known thermal growth techniques, such techniques, as known, providing a layer of good insulating quality. The silicon nitride layer 16 can be provided, for example, by known deposition techniques.

50 Then, using known photolithographic techniques, the third layer 18 is defined (Figures 3 and 6) to provide a rectangular gate electrode 18¹ of the transistor to be formed, and the workpiece is heated to a temperature of about 900°C. in a water vapor atmosphere to provide a relatively thick layer 20, e.g., 8,500 Å, of thermally grown silicon dioxide enveloping the otherwise exposed sides of the gate electrode 18¹. During this thermal growth process, the layer 16 of silicon nitride protects the silicon body 10 from being ox-

idized and additionally serves as a diffusion mask to prevent the conductivity defining atoms in the gate electrode 18¹ from diffusing through the silicon dioxide layer 14 and into the body 10.

70 The enclosing of exposed sides of the gate electrode prior to the formation of the source and drain regions of the transistors differs from the prior art practice wherein the gate protective oxide layer is provided after the source and drain regions are formed. As described hereinafter, this change from the prior art practice provides important advantages.

75 Thereafter, using the silicon dioxide coated gate electrode 18¹ as an etch mask, first the previously exposed portions of the silicon nitride layer 16 are removed, using hot phosphoric acid, and then the underlying portions of the silicon dioxide layer 14 thus exposed are also removed, using buffered hydrofluoric acid. While the etchant used to remove the silicon dioxide layer 14 also attacks the silicon dioxide layer 20 covering the gate electrode 18¹, the etching process is discontinued as soon as the much thinner layer 14 is removed, whereby a substantial thickness of the layer 20, e.g., 7,500 Å, remains. Thus, no masking of the layer 20 is required during these etching steps. At the conclusion of these steps, as shown in Figure 4, a composite gate structure 22 is provided comprising the layer 14 of silicon dioxide, the layer 16 of silicon nitride, the gate electrode 18¹, and the gate electrode covering layer 20 of silicon dioxide. The gate electrode 18¹ has a pair of oppositely disposed edges 44 disposed laterally inwardly of oppositely disposed edges 42 of the composite gate structure 22.

80 The source and drain regions of the transistor are next formed. This is accomplished, for example, by depositing on the workpiece a layer 24 (Figure 5) of silicon dioxide having a high concentration, e.g., in the order of 1×10^{20} atoms/cc of N type conductivity impurities, e.g., phosphorus, and, using known photolithographic techniques, defining the layer 24 to cover part (Figure 6) of the gate structure 22 and to extend beyond two opposite sides thereof. Thereafter, the workpiece is heated in an inert atmosphere, e.g., argon, at 1,100°C., to drive the impurities in the layer 24 into the silicon body. Where the impurities diffuse into the body 10, i.e., within regions 30 and 32 on opposite sides of the masking gate structure 22 (Figure 5), the conductivity of the body, originally of P type, is changed to N type. The two regions 30 and 32 comprise the source and drain regions of the transistor being fabricated, the region 34 between the regions 30 and 32 being the transistor channel region, and the

boundaries 40 between the region 34 and each region 30 and 32 comprising a P-N junction.

Alternatively, the source and drain regions 30 and 32 can be provided by placing the workpiece in a diffusion furnace where the conductivity defining impurity is provided from either liquid, solid, or gaseous sources.

One advantage of the use of the protective layer 20 is that no doping of the gate electrode 18¹ occurs during the formation of the source and drain regions. Thus, the conductivity characteristics of the gate electrode 18¹ are not affected by the conductivity characteristics of other portions of the device and can be selected as desired. This provides great flexibility in the design of devices made in accordance with the instant invention.

Because the impurities diffuse laterally, as well as downwardly, the junctions 40 of the regions 30 and 32 with the region 34 are disposed inwardly of the edges 42 of the gate structure 22.

As previously noted, a problem with prior art fabricated devices is that the lateral diffusion of the source and drain conductivity modifiers gives rise to excessive overlap between the gate electrode and the source and drain region. This is undesirable, as known, since it increases the inter-electrode capacitance of the device, thereby reducing the device high frequency performance.

In accordance with the invention, however, the edges 44 of the gate electrode 18¹, as previously noted, are disposed inwardly of the edges 42 of the gate structure 22, the location of the edges 42 relative to the edges 44 being determined by the thickness of the gate protective oxide layer 20. Thus, knowing the amount of lateral diffusion which will occur, this being substantially equal to the depth of the source and drain regions 30 and 32, e.g., 1 micron (determined by the desired device characteristics), a thickness of the layer 20 is selected which will result in the junctions 40 being disposed just inwardly of the gate electrode 18¹ edges 44. Although exact alignment of the junctions 40 with the gate electrode edges 44 would be preferred, owing to unavoidable processing variations, exact positioning of the junctions cannot be obtained on a reproducible basis, and the process is designed to produce some overlap, e.g., 2,500 Å, of the gate electrode 18¹ with the source and drain regions. With respect to device performance, such a condition of overlap is preferred to a condition in which the source and drain regions 30 and 32 lie outside a projection of the gate electrode 18¹.

To complete the device, the doping layer

24 is removed, as with buffered hydrofluoric acid, an opening 46 (Figure 7) is made through the protective oxide 20 to expose a surface portion of the gate electrode 18¹, a layer of metal, e.g. a 10,000 Å thick layer of aluminum, is deposited on the workpiece, and, using known photolithographic steps, the metal layer is defined to provide contact 58 for the gate electrode 18¹. Similarly, the metal layer also provides contacts 52 and 54 for the source 30 and drain 32 regions respectively.

A further advantage of the hereindescribed process is that since the gate electrode 18¹ is already completely enclosed by the insulating layer 20 prior to the application of the metal layer, there is comparatively little danger of shorting either the source 30 or drain 32 regions with the gate electrode 18¹, via the metal contacts provided on the workpiece surface. For example, as shown, the metal contacts 52 and 54 can even partially overlap the gate structure 22 with little ill effect, the relatively thick oxide layer 20 preventing excessive capacitive coupling between the metal contacts 52 and 54 and the gate electrode 18¹. By thus protecting the gate electrode 18¹, the various metal contacts can be defined with comparatively little regard for accuracy, thus simplifying the process and reducing the cost thereof.

Additionally, contrary to the prior practice, no protective insulating layer is provided over the source and drain regions 30 and 32. In the prior process, in which the gate electrode is not enclosed in a protective coating early in the process, as in the instant process, the source and drain regions are provided with a protective layer simultaneously with the provision of the protective layer over the gate electrode. The protective layer covering the source and drain regions is left in place, and contact openings are made through the layer to expose surface portions of the source and drain regions. The opening forming process requires a fair degree of process accuracy (thus adding expense to the process) since if the openings are not properly spaced from the gate electrode, the contact openings can expose both the gate electrode and one of the adjacent regions through the same opening. Thus, the metal contacts extending into the contact opening would cause a shorting of the device. Additionally, owing to the requirement that the source and drain contact openings be spaced from the gate electrode, larger devices must be made than are required with the hereindescribed process in which the source and drain contacts, as described, can even overlap the gate electrode. Thus, greater packing density of devices, in accordance with the instant invention, can be provided.

As described, one function of the silicon

nitride layer 16 (Figure 3) on the body 10 is to prevent oxidation of the body 10 during the process of thermally growing the silicon dioxide layer 20 on the gate electrode 18¹. As previously noted, however, the nitride layer 16 can be omitted. In one such embodiment of the invention, not illustrated, the process of thermally growing the layer 20 on the electrode 18¹ causes a corresponding increase in thickness of the silicon dioxide layer 14 covering the body 10, the result being that the layer 14 is even thicker than the layer 20. Thereafter, in order to remove portions of the layer 14 to allow doping of the body 10 thereunder while not simultaneously removing the layer 20, the layer 20 is covered with a protective masking layer. The use of the silicon nitride layer 16 is thus preferred since, for one thing, it avoids the need, as described, for this protective masking operation.

While the invention has been described in connection with the use of a semiconductor material workpiece, the invention also has utility in the fabrication of semiconductor devices formed on insulating substrates. That is, instead of the source and drain regions being formed within a wafer of semiconductor material, as above described, the semiconductor material comprises a thin film on an insulating substrate, e.g., a film of doped silicon on a sapphire substrate, and the source and drain regions are formed in the thin film. Further processes to fabricate the devices on the insulating substrate are substantially identical to the above-described processes used to fabricate the devices on a semiconductor material wafer.

WHAT WE CLAIM IS:—

1. A method of fabricating insulated gate field effect transistors comprising providing on a surface of a body of doped semi-

conductor material a gate structure comprising a first layer of a dielectric material covered by a second layer of a heavily doped semiconductor material, covering at least the lateral edges of said second layer with a protective coating as a masking layer against the conductivity modifiers, and thereafter, using said gate structure as a mask, driving conductivity modifiers into said body through portions of said surface.

2. A method as claimed in claim 1, including providing a further masking layer intermediate said first two layers of said gate structure and covering said surface portions of said body, said further masking layer being effective to prevent the formation of a protective layer on said surface portions during the formation of said protective layer on said second layer, and using said gate structure as a mask, removing said masking layer over said surface portions.

3. A method as claimed in either one of claims 1 or 2 wherein said conductivity modifiers are caused to diffuse laterally under said gate structure a distance at least equal to the thickness of said protective coating.

4. A method as claimed in any one of claims 1 to 3 including the steps of providing conductive contacts covering and in direct contact with, substantially the entire extent of said surface portions.

5. A method of fabricating a semiconductor device substantially as described herein with reference to the accompanying drawings.

For the Applicants:
MATTHEWS, HADDAN & CO.,
Chartered Patent Agents,
Haddan House,
33 Elmfield Road,
Bromley, BR1 1SU.



4
5 nitride layer 16 (Figure 3) on the body 10 is to prevent oxidation of the body 10 during the process of thermally growing the silicon dioxide layer 20 on the gate electrode 18¹. As previously noted, however, the nitride layer 16 can be omitted. In one such embodiment of the invention, not illustrated, the process of thermally growing the layer 20 on the electrode 18¹ causes a corresponding increase in thickness of the silicon dioxide layer 14 covering the body 10, the result being that the layer 14 is even thicker than the layer 20. Thereafter, in order to remove portions of the layer 14 to allow doping of the body 10 thereunder while not simultaneously removing the layer 20, the layer 20 is covered with a protective masking layer. The use of the silicon nitride layer 16 is thus preferred since, for one thing, it avoids the need, as described, for this protective masking operation.

10 While the invention has been described in connection with the use of a semiconductor material workpiece, the invention also has utility in the fabrication of semiconductor devices formed on insulating substrates. That is, instead of the source and drain regions being formed within a wafer of semiconductor material, as above described, the semiconductor material comprises a thin film on an insulating substrate, e.g., a film of doped silicon on a sapphire substrate, and the source and drain regions are formed in the thin film. Further processes to fabricate the devices on the insulating substrate are substantially identical to the above-described processes used to fabricate the devices on a semiconductor material wafer.

40 WHAT WE CLAIM IS:—

1. A method of fabricating insulated gate field effect transistors comprising providing on a surface of a body of doped semi-

conductor material a gate structure comprising a first layer of a dielectric material covered by a second layer of a heavily doped semiconductor material, covering at least the lateral edges of said second layer with a protective coating as a masking layer against the conductivity modifiers, and thereafter, using said gate structure as a mask, driving conductivity modifiers into said body through portions of said surface.

2. A method as claimed in claim 1, including providing a further masking layer intermediate said first two layers of said gate structure and covering said surface portions of said body, said further masking layer being effective to prevent the formation of a protective layer on said surface portions during the formation of said protective layer on said second layer, and using said gate structure as a mask, removing said masking layer over said surface portions.

3. A method as claimed in either one of claims 1 or 2 wherein said conductivity modifiers are caused to diffuse laterally under said gate structure a distance at least equal to the thickness of said protective coating.

4. A method as claimed in any one of claims 1 to 3 including the steps of providing conductive contacts covering and in direct contact with, substantially the entire extent of said surface portions.

5. A method of fabricating a semiconductor device substantially as described herein with reference to the accompanying drawings.

For the Applicants:
MATTHEWS, HADDAN & CO.,
Chartered Patent Agents,
Haddan House,
33 Elmfield Road,
Bromley, BR1 1SU.