Surface Protection and Selective Masking during Diffusion in Silicon

C. J. Frosch and L. Derick

Bell Telephone Laboratories, Inc., Murray Hill, New Jersey

ABSTRACT

An apparatus is described for the vapor-solid diffusion of donors and acceptors into silicon at atmospheric pressure. It consists essentially of a fused silica tube extending through one or more controlled temperature zones. A gas such as nitrogen carries the vapors from the heated impurity element or one of its compounds past the heated silicon.

At temperatures above about 1000°C, gases such as helium or nitrogen are shown to cause serious pitting or erosion of the silicon surfaces. A thin vitreous silicon dioxide envelope enclosing the silicon during the high temperature heating operation is shown to provide complete protection of the underlying surface against damage. Methods of obtaining surface passivation are described.

In addition to surface protection, a silicon dioxide surface layer also is shown to provide a selective mask against the diffusion into silicon of some donors and acceptors at elevated temperatures. Data are presented showing the masking effectiveness of the silicon dioxide layer against the diffusion of several donors and acceptors into silicon.

The application of the masking technique to produce precise surface patterns of both n- and p-type is described. An example of its feasibility in device considerations is illustrated by the construction of a transistor by double diffusion. This transistor is unique in that both the emitter and base contacts are made at the surface in adjacent areas.

Finally a new predeposition technique is described for controlling the impurity levels in diffused layers over wide ranges. Data are presented to illustrate this technique.

The vapor-solid diffusion technique has become of considerable interest recently for the introduction of selected conductivity type determining impurities into semiconductors. The application of this technique in the design of Si semiconductor devices is illustrated by the solar battery (1), power rectifiers (2), and high frequency transistors (3). The feasibility of producing such structures by vapor-solid diffusion resulted largely from the investigations of Fuller and his co-workers (4, 5). Their investigations provided the necessary information on the high temperature diffusion parameters required for controlling the impurity concentrations and depths of the diffused layers.

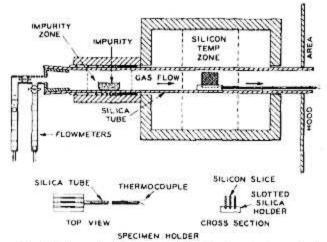
This paper describes: (a) a new apparatus for vapor-solid diffusion at atmospheric pressure, (b) some experiments illustrating surface protection of Si at high temperatures by oxidation, and (c) some experiments illustrating selective masking by SiO, surface layers against the diffusion of some impurities into Si at high temperatures. Methods for determining the electrical characteristics, surface concentrations, and diffusion depths are essentially the same as those described elsewhere (5) and are only identified in this paper.

Apparatus and Procedure

The apparatus for vapor-solid diffusion at atmospheric pressure is shown schematically in Fig. 1. It consists essentially of a fused silica tube extending through two controlled temperature zones. The first temperature zone serves to regulate the volatility of an impurity placed therein. The second temperature zone is the location of the Si samples. A rising temperature gradient between the two zones avoids the redeposition of the impurity vapor before reaching the Si samples. The temperatures are regulated to ±2°C by means of Pt-Pt 10% Rh thermocouples and automatic controllers.

The vapor from an impurity heated in the first temperature zone is carried past the Si heated in the second zone by carrier gases such as Hr, Nr, On and CO2 or mixtures thereof with or without the addition of water vapor. Flowmeters attached to one end of the furnace tube regulate the flow of the carrier gases. In this investigation flow rates of 1500 cc/min of the carrier gas were employed in furnace tubes of 2.5 cm ID. A temperature-controlled water bubbler is inserted between the flowmeters and the furnace tube when water vapor is being added to the carrier gas. Carrier gases were not saturated with water vapor under these flow conditions, but particular experimental conditions were found to be reproducible. The exit or open end of the furnace tube is inserted into a hood. This effectively removes the carrier gases which now probably contain residual amounts of 3rd and 5th column elements or their compounds. These in many cases are harmful poisons.

The Si samples are held vertically in a slotted fused silica plate as shown in Fig. 1. This assures a uniform contact of the carrier gas with both Si sur-



548

Fig. 1. Apparatus for vapor-solid diffusion at atmospheric

faces. A small silica tube fused to the sample holder is used to insert and remove the Si from the furnace. A Pt-Pt 10% Rh thermocouple inserted in the pusher tube is employed to measure the work temperature throughout the run.

Before the Si samples are inserted for diffusion or oxidation, the flow of carrier gas is started and allowed to continue for several minutes to establish a steady-state condition. The samples then are introduced at a reproducible rate into the desired temperature zone. At the end of the heating operation the samples generally are withdrawn at the insertion rate. In these experiments the samples were inserted and withdrawn in about 1 min.

Samples

The Si samples were cut from approximately 5 ohm cm n-type (arsenic doped) or p-type (boron doped) single crystal ingots grown in the [111] direction. The ingots were sawed perpendicular to the growth direction into slices approximately 10-1 cm thick. These were machine lapped on both surfaces with No. 1800 emery powder to a thickness of about 7 x 10⁻⁷ cm. After cutting into squares of 0.75 x 0.75 cm, the samples were etched for 4-5 min in a mixture of 5 parts of concentrated HNO, to 1 part of concentrated HF to produce a mirror finish. This resulted in final sample thicknesses of about 5 x 10-2 cm. Reproducible surfaces of excellent appearance were obtained by maintaining the etchant temperature at about 20°C and employing mechanical agitation during etching.

Surface Protection

When Si is heated in the previously described diffusion apparatus above about 1000°C for relatively short times in commercial grades of nonoxidizing gases such as H, or He, serious erosion or pitting of the surface occurs. A significant loss in weight also takes place. On the other hand, Si surfaces are essentially unchanged after long heating periods well above 1000°C in oxidizing gases such as O2 and CO; or nonoxidizing gases containing sufficient concentrations of oxygen or water vapor.

The surface appearance of Si samples after heating in H, He, No, or O, gases is illustrated in Fig. 2. The surfaces were washed in concentrated HF after

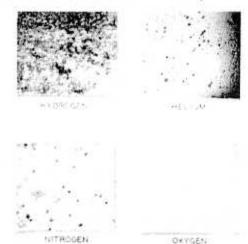


Fig. 2. Surface appearance of Si after heating for 1 hr at 1350°C in different carrier gases.

the heating operation to remove any surface layers such as SiO, which forms in the pure oxygen atmosphere. Note the damage on the Si surfaces after heating in H, and He. On the other hand after heating in oxygen the excellent surface appearance is essentially the same as that of the original etched surface. The crystalline surface appearance after heating in N, may be due to the formation of a Si,N,.

The effect of varying the water vapor content in a hydrogen carrier gas during heating on the surface appearance is illustrated in Fig. 3. The temperatures under the photographs represent the temperatures of the water bubbler. Note the decrease in the size of the surface pits as the water vapor content of the carrier gas is increased. Actually in this series of heating experiments at 1350°C, water bubbler temperatures between 30°-80°C resulted in excellent surfaces. However, a further increase in the water content of the carrier gas resulted in a large number of very small surface pits. The reappearance of surface pits might be explained by the failure to form a completely coherent oxide layer in the presence of high water vapor contents at this température. Silicon dioxide is known to be quite soluble in water at high temperatures.

Other examples of surface protection by oxidation might be given. For example, Si can be heated

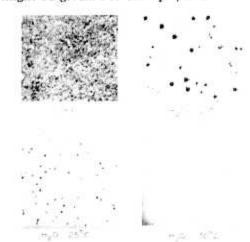


Fig. 3. Surface appearance of Si after heating for 1 hr at 1350°C in hydrogen containing different concentrations of water vapor

549

550

in CO₂ or in various mixtures of N₂ and O₂ including air without apparent surface damage. The latter suggests many process simplifications for producing diffused junction layers. However, the oxidizing potential of the carrier gas must be sufficient so that the thermodynamic equilibrium favors the formation of SiO, at each processing condition. On the other hand optimum surface appearance does not necessarily result from maximum oxidizing conditions. Generally it is preferable to establish optimum conditions experimentally for each processing condition.

SURFACE PROTECTION & SELECTIVE MASKING

Surface protection of Si in oxidizing atmospheres is due to the formation of a continuous protective envelope of nonvolatile SiO₃. This prevents not only the evaporation of the underlying Si but also its complete oxidation. The oxide layer is easily removed by washing in HF. In this work the SiO₂ layers generally varied in thickness between about 1500 to 3000Å depending on the temperature, time and carrier gas composition. The layers exhibit brilliant interference colors. The constant thickness of the oxide layer on a particular Si surface is indicated by the uniformity of these interference colors.

Selective Masking

In addition to surface protection, a SiO₈ surface layer provides a selective mask at high temperatures against the diffusion of some donors and acceptors into Si. Only a few typical examples of selective masking will be given.

Masking is studied most effectively by preoxidizing etched Si slices and removing the oxide layer from one surface with HF. This provides an oxidized and unoxidized surface on the same sample for electrical, diffusion depth, and other measurements. In the examples reported here the preoxidation was done by heating at 1200°C for 1 hr in N, bubbled through water at 30°C. This gives an oxide layer about 2500Å thick.

Fig. 4 illustrates the masking of As by a SiO₂ layer at three diffusion temperatures. The first figure is a cross-sectional view of a p-type Si slice with an oxide layer on one surface. The next figure shows that arsenic trioxide is heated at 235°C and the carrier gas is N₂ bubbled through water at 30°C for each processing condition. In the next column the diffusion times and temperatures are given for

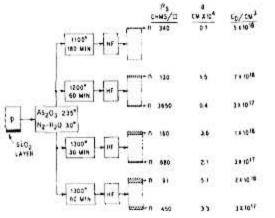


Fig. 4. Masking by an exide layer against the diffusion of As at different diffusion temperatures.

each of the four processing conditions. The fifth column shows a cross-sectional view of the processed samples after a wash in HF which removes only the oxide envelope. The cross-hatching illustrates the relative thickness of the n-type conversion layer. The average sheet resistance values (ρ_i) and the diffusion depths (a) were measured by the four point probe and the 5 degree angle lapping-staining technique, respectively. These measuring techniques are described in some detail elsewhere (5). From the ρ_i and (a) values the surface concentrations (C_0) of the impurities given in the last column were calculated by assuming an error function distribution of the impurities through the diffused layers.

Note the absence of a conversion layer on the oxidized surface after the 1100°C diffusion shown in Fig. 4. This indicates that the surface concentration of the As on this surface is less than 3 x 10° cm², the background p-type doping in the original crystal. At the two higher diffusion temperatures a conversion layer is formed on the oxidized surface. However, the C_{*} values remain relatively constant and well below those for the unoxidized surfaces. The effect of the SiO_{*} surface layer and the diffusion temperature on surface concentration and hence diffusion depth is clearly evident from the data.

Fig. 5 illustrates the effect of carrier gas type on the diffusion of boron. The representations have the same significance as those described previously. Because of its very low volatility, the boron trioxide source is heated at the same temperature as the Si. Note the absence of a conversion layer on the oxidized surface and the highly doped conversion layer on the unoxidized surface after heating in oxygen. An appreciable masking effect by the oxide layer also is evident after heating in the nitrogen-water carrier gas. On the other hand no masking by the preoxidized surface is evident after heating in the hydrogen-water carrier gas. The doping level also is appreciably higher for the wet hydrogen than for the unoxidized surfaces in the other two carrier gases. The reason for this effect of carrier gas type and composition on the masking by SiO, surface layers is not known. However, it does suggest some difference in the molecular specie of the diffusant vapor in the various carrier gases which alters its reactivity with, diffusion through, or solubility in a SiO, layer. A complete study of the effect of carrier gas composition on the masking of a SiO, surface layer against the diffusion of donors and acceptors

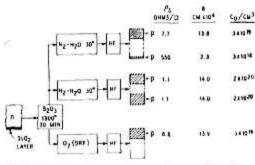


Fig. 5. Effect of carrier gas type on the masking by an oxide layer against the diffusion of B.

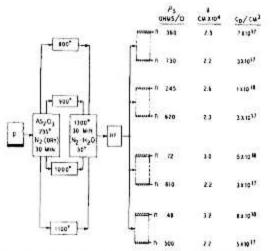


Fig. 6. Effect of predeposition temperature on masking and surface concentration.

into Si at high temperatures should be very useful in the design of semiconductor devices. A few experimental results for wet hydrogen and nitrogen as well as oxygen are summarized in Table I.

Predeposition of Impurities

When a SiO₃ layer masks against the diffusion of an impurity, the simultaneous oxidation and addition of such an impurity may cause a partial masking effect on the unoxidized surface. Fig. 6 illustrates a technique for avoiding oxidation effects during the addition of the impurity. It consists essentially of predepositing the impurity in nonoxidizing gases such as N_n, He, or H_n at low temperatures and for very short times. Under these processing conditions surface erosion is not appreciable. The diffusion of the impurity into the Si then may be carried out at higher temperatures and for longer times in oxidizing carrier gases.

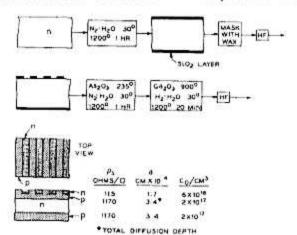


Fig. 7. Simultaneous production of single and double diffused layers of controlled structure by selective masking.

In Fig. 6 As₂O₈ was predeposited for 30 min in dry nitrogen at 800°, 900°, 1000°, and 1100°C, respectively. The samples then were diffused for 30 min at 1300°C in N₂ bubbled through water at 30°C. The samples usually were washed in HF after the predeposition to remove any soluble deposits of the diffusant on the unoxidized surface and the preoxidized layer which now may contain some of the diffusant. This washing procedure was found to give more reproducible results.

As shown by the data in Fig. 6 the SiO₂ layer masks against the diffusion of As in the case of all four predeposition temperatures. The degree of masking, which is defined as the ratio of the surface concentration of the unoxidized surface to that of the oxidized surface, increases with predeposition temperature over the range studied. The data for the unoxidized surfaces also illustrate the usefulness of the predeposition technique for controlling the surface concentration over wide limits. The applica-

Table 1. Masking by axide surface layers against the diffusion of same donors and acceptors*

JOURNAL OF THE ELECTROCHEMICAL SOCIETY

Diffusant		Diffusion		Surface type		ns (ohms/m)		(em × 10+)		Cu/cm ²	
Type	Temp.	Time.	Temp,	0	υ	O	U	o	U	0	υ
					N ₂ -1	LO-30°C (arrier ga	is			
As,O,	235	180	1100	P	N	***	340		0.7	8000m2	5×10"
As ₂ O ₁	235	60	1200	P N	N	3650	130	0.4	1.5	3×10^{11}	7×10°
As ₂ O ₄	235	30	1300	N N P	N	680	160	2.1	3.6	3×10^{17}	1×10"
B ₂ O ₃	1250	30	1250	N	P	-	76	Administra	5.1	22	7×10
B ₂ O ₄	1300	30	1300	P	P	550	7.7	2.3	13.8	3×10^{14}	3×10"
P(Red)	220	180	1100	N	N	2560	195	0.5	2.4	4×10^{17}	1×10°
P(Red)	220	30	1300	N	N	885	170	3.6	7.4	8×10^{14}	4×1011
Sb ₂ O ₄	900	30	1300	N	N	1370	140	1.0	4.9	3×10^{17}	6×10°
					11,-1	LO-30°C C	Carrier go	us		31-51 NESCON	87,56,00
As ₂ O ₃	235	60	1200	P-N	N		330	See	1.3	200	1×10"
As _e O ₁	235	30	1300	N	N	680	86	2.1	3.7	3×10"	$2 \times 10^{\circ}$
B ₂ O ₃	1300	30	1300	N P	P	1.1	1.1	14.0	14.0	2×10 ²⁰	$2 \times 10^{\infty}$
P(Red)	220	30	1300	N	N	745	150	3.6	5.6	1×10"	4×10"
Ga ₂ O ₄	900	40	1200	N P	P	1890	1835	5.8	5.8	6×1010	6×10°
Ga ₂ O ₁	900	30	1300	P	P	675	675	9.1	9.1	1×10"	1 × 101
Sb ₂ O ₄	900	30	1300	P P	N		160	-	5.6	777	4×10°
						bry O, car	rier gas				
As ₂ O _a	235	30	1300	N	N	545	235	2.0	3.1	3×10^{17}	5×1011
B _v O _a	1300	30	1300	N	P	1	6.8		13.9	4114	3×1010
P ₂ O ₅	220	180	1100	N	N	1390	25	0.9	3.8	3×10^{17}	7×10°
P ₂ O ₃	220	30	1300	N	N	0.45	0.44	19.1	20.1	3×10^{10}	3×10^m

^{*}Original samples about 5 ohm cm, preoxidized 1 hr at 1200°C in Ng-HgO-30°C, U-oxide layer removed from this surface with HF before diffusions. O-oxide layer on this surface before diffusions.

551

SURFACE PROTECTION & SELECTIVE MASKING

Multiple Layer Structures

A more interesting example of selective masking by a SiO, layer is illustrated in Fig. 7. After the preoxidation of an n-type Si sample in wet nitrogen at 1200°C, the oxide layer is removed from one surface to give alternate oxidized and unoxidized bands. This is done by applying alternate bands of wax, washing in concentrated HF, and then removing the wax with a solvent. An As diffusion followed by a Ga diffusion and an HF wash results in the structure illustrated by the top and cross-section views shown in the bottom left of Fig. 7. Note the alternate p- and n-type surfaces as well as the alternate single and double diffused layers. These correspond to the alternate oxidized and unoxidized bands, respectively, on the top face before the diffusion operations. The bottom surface, which is completely covered with an oxide layer before the diffusion treatments, is entirely p-type with a single diffused layer. A similar structure results if Sb₀O, heated at 900°C is substituted for the arsenic trioxide.

It is not necessary to remove the oxide layer after the As diffusion since Ga is not masked by the SiO, layer under these processing conditions (see Table I). Actually the oxide layers covering the n-and p-type surfaces before the final HF wash are distinctly different in color. The color difference probably is due to a thickness difference between the two oxide layers. This is indicated since the same colors are observed when the identical heating sequence is employed without the addition of the As and Ga. The color difference of the oxide layers covering the n- and p-type Si surface areas is useful in device fabrication as an aid in locating these regions after diffusion.

A transistor (6) has been made by these techniques which had an alpha of 0.91 at a collector voltage of 10 v and an emitter current of 1 ma. The unique feature of this transistor structure was the ability to make a direct contact on the same surface to both the diffused base layer and the diffused emitter layer. In previously described double diffused transistors (3), the base contact was made by alloying Al through the emitter layer. The direct contact to the base layer described in this paper also should allow a higher surface concentration in the emitter than is permissible in the Al alloy base contact structure.

The structure obtained in Fig. 7 can best be explained by referring to Fig. 8. This illustrates the appreciably lower surface concentration of As in the oxidized surface as compared to the unoxidized surface. Since the Ga is not masked by the SiO-layer under these processing conditions, its surface concentration is equivalent for both surfaces. Thus the As overdopes the Ga on the unoxidized surface, while the reverse is true on the oxidized surface. In other words a np diffused structure is produced in the unoxidized surface and a single p diffused layer is produced in the preoxidized surface. The surface concentration of the diffused As layer in the oxidized

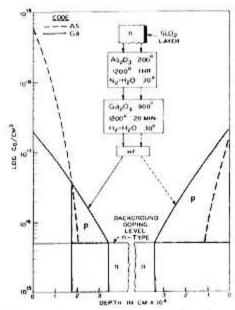


Fig. 8. Plot to illustrate the farmation of single and double diffused layers by selective masking.

surface was not determined in this experiment and is drawn merely for illustration.

Fig. 9 shows a photograph of an actual diffused structure which was produced by the selective masking technique just described. The diffused layers on one surface were exposed by lapping at a 5 degree angle (5) perpendicular to alternate p- and n-type surface bands similar to those illustrated in Fig. 7. The p region was darkened by the HF-HNO₃ staining technique (5). Note the single p diffused layer in the center region with double np diffused layers on either side. While not evident from the photograph, the dark surface or p region indexes very closely with the original preoxidized surface.

Controlled structures with either n- or p-type single diffused layers and either np- or pn-type double diffused layers have been produced by the selective masking procedure when the proper diffusants and processing conditions are employed. Moreover, both structural types have been produced with B and P by employing the predeposition technique. This combination of impurities also can be employed for producing low resistance surface layers for collector contacts in transistor construction. The essential processing steps for producing either an np-

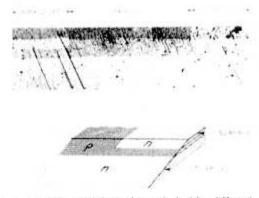
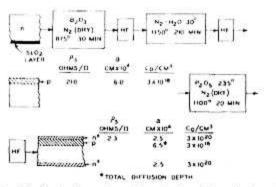


Fig. 9. Photograph of single and double diffused layers produced by selective masking.



552

Fig. 10. Production of a $p^*p\pi\pi^*$ transistor structure by predisposition and selective masking.

or pn-type transistor structure of this type with B and P are shown in Fig. 10 and 11.

In the np structure illustrated in Fig. 10, the predeposition of the B_sO_a at 875°C in dry N, gives a surface concentration (C_a) on the unoxidized surface which is well below that eventually reached by the P. The processing conditions are chosen to give the desired final diffusion depth of the p-layer into the originally unoxidized surface. Because of the masking effect of the oxide layer during the predeposition of the boron trioxide, the C_a of the oxidized surface is reduced to well below that of the unoxidized surface. Actually no B conversion layer is formed in the oxidized surface under these processing conditions. Thus the structure illustrated by the last cross section in Fig. 10 results after the final phosphorus diffusion.

The pn structure illustrated in Fig. 11 is produced in a similar manner except that the P is masked and its surface concentration is reduced by predeposition below that eventually reached by the B. In this case a conversion layer is formed in the oxidized surface by the P. However, the partial masking by the oxide layer is sufficient to reduce the diffusion depth below that finally attained by the B layer. Because of this only a p' layer is formed on the original oxidized surface as shown in Fig. 11. At these low doping levels, red P as a source of P vapor gives more reproducible results than P₂O₅. Red P also generally results in lower surface concentrations than the phosphorus oxide for otherwise similar processing conditions.

In addition to the structural possibilities of masking, it also can be applied in conjunction with the

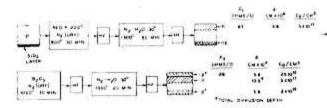


Fig. 11. Production of p'npp' transistor structure by predeposition and selective masking.

predeposition technique to provide a wide range of controlled impurity levels in diffused layers. For example, by modifying the processing conditions illustrated in Fig. 10 and 11, the doping levels of the base and emitter layers can be varied over wide

It is evident from the few examples of selective masking described that a large variety of structures can be produced by the proper choice of processing conditions. It is hoped that the examples given illustrate the essential procedures for producing such structures.

Summary

The use of oxidizing atmospheres has been shown to provide surface protection of Si surfaces during high temperature heating operations. The protection results from the enclosure of the Si in a SiO, envelope. The latter also provides a means for masking against the diffusion of some impurities into Si at high temperatures. The masking effectiveness of the oxide layer is determined by several process parameters such as time, temperature, carrier gas composition, impurity type, impurity compound, and preoxidizing conditions. Thus far, processing parameters have been established for masking B, As, Sb, and P. Precise surface structures of both n- and p-type as well as the simultaneous production of single and multidiffused junction layers are possible by this masking technique. Finally, a wide range of controlled impurity levels in diffused layers can be obtained by means of a new predeposition technique.

Acknowledgment

The authors wish to thank several people for their contribution to this investigation. In particular, the authors are grateful to C. S. Fuller, M. Tanenbaum, J. L. Moll, and F. M. Smits for their helpful discussions; to G. Backenstoss for the calculations of the surface concentration values; and to H. E. Bridgers for supplying the crystal ingots. Thanks also are due to R. M. LeBach for some of the measurements.

Manuscript received Nov. 29, 1956. Portions of this material were presented at the Semiconductor Conference of the I.R.E. in June 1956 and portions at the Cleveland Meeting of the ECS, Sept. 30 to Oct. 4, 1956. Any discussion of this paper will appear in a Discussion Section to be published in the June 1958 JOURNAL.

REFERENCES

- D. M. Chapin, C. S. Fuller, and G. L. Pearson, J. Appl. Phys., 25, 676 (1954).
- G. L. Pearson and C. S. Fuller, Proc. Inst. Radio Engrs., 42 (April 1954).
- M. Tanenbaum and D. E. Thomas, Bell System Tech. J., 34, 105 (1955).
- C. S. Fuller and J. A. Ditzenberger, J. Appl. Phys., 25, 1439 (1954).
- C. S. Fuller and J. A. Ditzenberger, ibid., 27, 544 (1956).
- M. Tanenbaum, Private communication.