Solid-State Electronics Pergamon Press 1971. Vol. 14, pp. 499–505. Printed in Great Britain

CONVERSION OF SILICON NITRIDE INTO SILICON DIOXIDE THROUGH THE INFLUENCE OF OXYGEN

I. FRÄNZ and W. LANGHEINRICH
AEG-Telefunken, Forschungsinstitut und Röhrenwerk Ulm, Germany

(*Received* 3 *August* 1970; *in revised form* 22 *October* 1970)

Abstract—When thin films of silicon nitride are annealed in the presence of oxygen, thin films of silicon dioxide are formed on the nitride layers while silicon nitride is consumed. The thickness of the oxide layers which are built up by reaction with dry oxygen, wet oxygen and by a mixture of oxygen and nitrogen in the presence of phosphorus pentoxide, are given as a function of time and temperature. The results are compared with the thermal oxidation of silicon.

Résumé—Lorsque de minces pellicules de nitrure de silicium sont recuites en présence d'oxygène, il se forme de minces pellicules de bioxyde de silicium sur les couches de nitrure alors que le nitrure de silicium est utilisé. L'épaisseur des couches d'oxyde qui se forment par réaction avec l'oxygène sec, l'oxygène humide et avec un mélange d'oxygène et d'azote en présence de pentoxyde de phosphore, est donnée en fonction du temps et de la température. Les résultats sont comparés à l'oxydation thermique du silicium.

Zusammenfassung—Tempert man dünne Siliziumnitridschichten bei Gegenwart von Sauerstoff, so bilden sich auf den Siliziumnitridschichten unter Verbrauch von Siliziumnitrid dünne Siliziumdioxidschichten. Die bei Einwirkung von trockenem Sauerstoff, feuchtem Sauerstoff sowie eines Sauerstoff-Stickstoffgemisches bei gleichzeitiger Anwesenheit von Phosphorpentoxid gebildeten SiO_2-Schichtdicken wurden als Funktion von Zeit und Temperatur gemessen. Die Ergebnisse werden mit der thermischen Oxydation des Siliziums verglichen.

1. INTRODUCTION

SINCE Newman *et al.*[1] reported on the use of silicon nitride in microelectronics, a number of papers [e.g. 2–6] have been published about the preparation and use of this material as a thin amorphous film for the purpose of diffusion masking and surface passivation in the manufacture of semiconductor devices. It has already been reported that silicon nitride masks against the diffusion of oxygen[3, 6]. From applications in the ceramics industry we also know that silicon nitride is resistent to oxygen[7]. Tombs and Sewell[5] have already used the conversion of silicon nitride into silicon dioxide in the presence of oxygen at high temperatures to produce a thin silicon dioxide film which acted as an etching mask. Their data indicate that the conversion process is very slow, but more accurate data about this process are not available in the literature. These data are, however, necessary for semiconductor technology where very thin silicon nitride films are used and where it is frequently necessary to anneal in the presence of oxygen. It was therefore the purpose of this work to investigate more closely the reaction between oxygen and silicon nitride. The standard free energy of reaction as calculated from Janaf thermochemical tables[8] according to the equation

$$Si_3N_4 + 3O_2 \longrightarrow 3SiO_2 + 2N_2$$

is shown in Fig. 1 in the range of 1000 to 1600°K. The reaction therefore should be possible in this range, which is in accordance with the results.

In the described experiments the oxidation was performed with dry and wet oxygen. The evaluation of the conversion of nitride also included a simultaneous analysis of the silicon oxidation. Though the thermal oxidation of silicon has been thoroughly examined, the results of the various authors do not always agree, a fact which is probably attributable to the different experimental conditions.

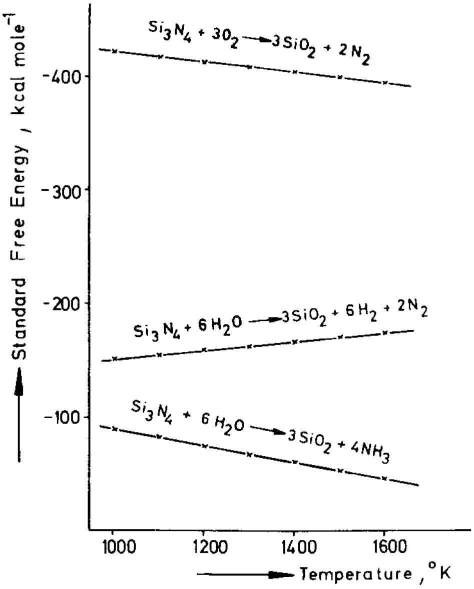

Fig. 1. Standard free energy as function of temperature.

The simultaneous performance of the experiments eliminated this shortcoming, making it possible to compare these results with those in the literature. With limitations, of course, the conversion of silicon nitride by means of oxygen may also be compared with the conversion of silicon nitride by oxygen in the presence of phosphorus pentoxide. Although, in the latter case, glass consisting of phosphorus pentoxide and silicon dioxide is produced, the ultimate result here also is the conversion of silicon nitride into silicon dioxide. As the discussion of the results will be based on the model given by Deal and Grove[9] for dry and wet oxidation of silicon, this model is given in the appendix. When the model was established, a lot of work had already been done on oxidation of silicon, in which empirically the parabolic relationship equation[11] had been found. Besides Deal and Grove several other authors e.g. [10–12] found a linear-parabolic relationship for the oxidation of silicon in dry oxygen or steam.

Silicon nitride was deposited on one side of polished silicon slices using the SiH_4–NH_3 system and then tempered at 1200°C. Subsequently the wafers were exposed to molecular sieves dried oxygen or to wet oxygen (H_2O 95°C) flowing at a rate of 80 l/hr through a tube of diameter 40 mm. Conversion in the presence of phosphorus pentoxide was performed by passing a gas consisting of 50 l/hr nitrogen and 5 l/hr oxygen first through a bubbler filled with phosphorus tribromide and then into the furnace[13]. The thicknesses of the glassy layers on both sides were measured interferometrically[14–16] as a function of temperature and time. From the reflection measurements required for this purpose, the thickness of single SiO_2 films can be calculated directly. However as will be shown later, a double layer of silicon nitride/silicon dioxide on silicon is obtained on the formation of silicon dioxide from silicon nitride. In this case the thickness of the underlying Si_3N_4 film must be determined. For this purpose the upper silicon dioxide film is successively etched off with buffered hydrofluoric acid. Since silicon dioxide is etched off about 100 times faster than silicon nitride, the nitride film can be exposed undamaged and its thickness calculated from the reflection measurement. Using this procedure the thickness of the silicon dioxide film may be checked through the overall time needed to etch it off because the etching rate of silicon dioxide is known. However, to get the thickness of the underlying Si_3N_4 film, it is not necessary to remove the entire upper layer as has been shown by the authors[17]. All steps of the process were checked by infra-red spectroscopy.

3. RESULTS

When silicon nitride is annealed in oxygen, the absorption peaks of silicon dioxide appear in the infra-red spectrum, from which it can be derived that silicon dioxide is formed. The absorption peaks of a wafer coated with silicon nitride are shown in the range of 600–1200 cm^{-1} before and after oxidation in Fig. 2. That both peaks occur in this range, the one characteristic for SiO_2 and the other for Si_3N_4, indicates that Si_3N_4 and SiO_2 exist in two separate phases, because only one absorption peak is observed from $Si_xO_yN_z$ films (limits SiO_2 and Si_3N_4), which exist in one homogeneous phase.

Additionally, during successive etching in hydrofluoric solution, the absorption peak of silicon dioxide disappears, and only after it has completely disappeared, the absorption peak of the silicon nitride gradually vanishes. This indicates that a film sequence $Si/Si_3N_4/SiO_2$ is built up during oxidation. Reflection curves obtained during etching are shown in Fig. 3. The normalized reflection

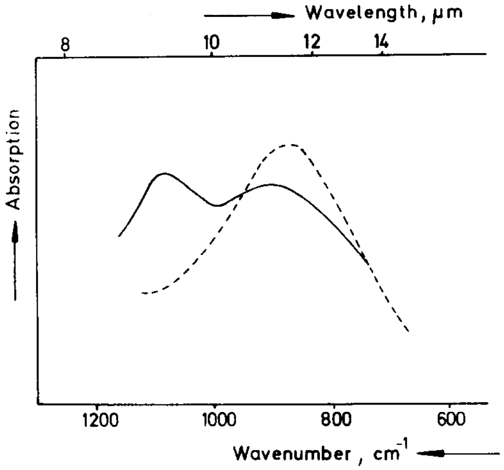

Fig. 2. Infra-red absorption of Si_3N_4 and a double film of Si_3N_4/SiO_2.

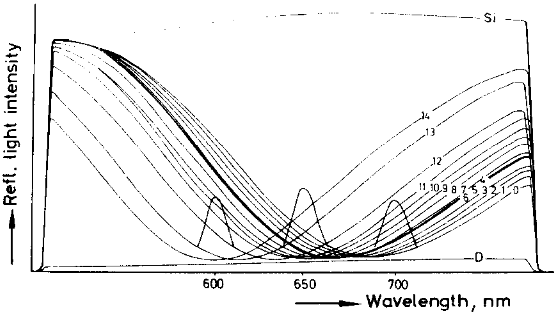

Fig. 3. Reflection measurements of a film sequence $SiO_2/Si_3N_4/Si$. The normalized reflection of silicon is designated by 'Si'. Etching time was 3 sec. in each case between curves 0–6. The etching time between the curves with higher numbers amounted to minutes.

of silicon is designated by 'Si'. The curves 0–6 refer to thicknesses of films resulting from succesive etchings of 3 sec. The curves 4–6 virtually coincide, i.e. after 3 sec etching time the film is unaffected. Here, the upper silicon dioxide film has been etched away and the silicon nitride film appears. The thickness of the silicon nitride film is reduced only after much longer etching times. It is obvious from Fig. 3 that a gap in the etching rate is obtained by chemical removal. Thus, if the etching rate of the etching solution is known for SiO_2, the thickness of the silicon dioxide film may be determined by means of the etching time. The gap in the etching rate is shown in Fig. 4 as a function of the distance from the surface. The etching solution used in the experiments etches silicon dioxide away at a rate of approx. 10 Å/sec, and silicon nitride approx. 100 times slower. From Fig. 3 the refractive indices of the two films may also be obtained. Here, too, the refractive index of silicon dioxide follows for the upper film.

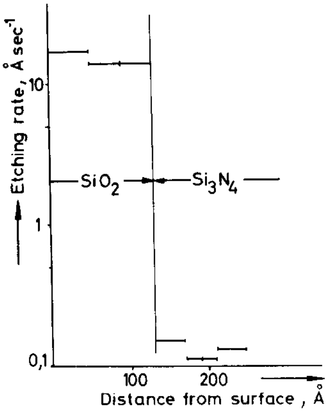

Fig. 4. Etching rate of a double film of SiO_2/Si_3N_4 on silicon as a function of distance from the SiO_2 surface.

The thickness of the silicon dioxide films formed, when Si_3N_4 is annealed in dry oxygen are shown in Fig. 5 as function of time at various temperatures in the range of 1000–1260°C. The values were calculated from interferometric measurements. The film thicknesses determined from the etching times agree with these values within an accuracy of 10 per cent. The silicon dioxide film thicknesses obtained by conversion of silicon nitride in wet oxygen are shown in Fig. 6 as function of oxidation time at various temperatures in the range of 800–1100°C. Here too, the stated silicon dioxide film thicknesses determined interferometrically agreed within an accuracy of 10 per cent with the film thicknesses obtained from the etching times. It will be noticed that the thicknesses of silicon dioxide films formed in a wet oxygen athmosphere are much larger than those of films formed in dry oxygen, other conditions being equal.

As was already mentioned, a phosphorus glass consisting of silicon dioxide and phosphorus pentoxide is formed when silicon nitride is exposed to oxygen and phosphorus pentoxide at the same time,

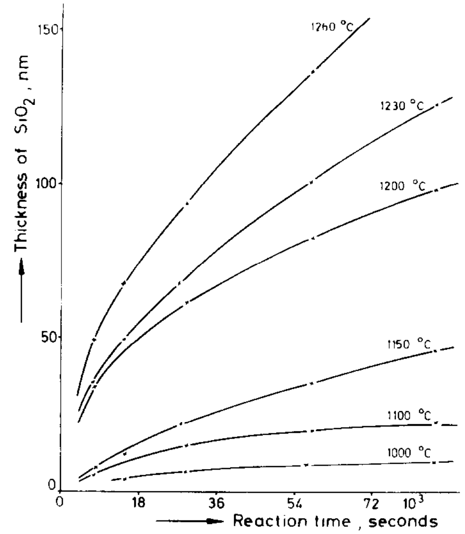

Fig. 5. Thickness of the SiO_2 film formed on Si_3N_4 on exposure to dry oxygen as a function to time with the temperature as parameter.

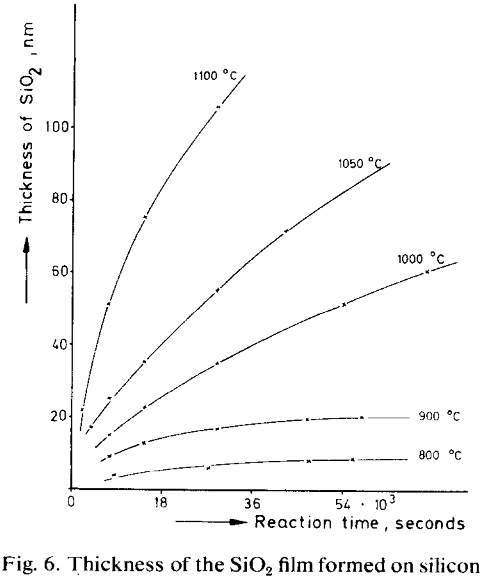

Fig. 6. Thickness of the SiO_2 film formed on silicon nitride on exposure to wet oxygen as a function of time with the temperature as parameter.

that means that in this case silicon nitride is also converted into silicon dioxide. The thickness of the converted Si_3N_4 film was determined and from it the corresponding thickness of SiO_2 was calculated in order to get a direct comparison with the oxidation rates in pure oxygen. As will be noted from Fig. 7, the thickness of the silicon dioxide film and thus converted Si_3N_4 film is a linear function of time in the range between 5 and 20 min and temperatures of 900, 1000, 1100 and 1200°C.

The experiments on the oxidation of silicon in dry oxygen, performed simultaneously with the experiments on conversion of silicon nitride in order to have a direct comparison, were evaluated with the aid of equation (11), which is valid for large oxidation times. The rate constants *B* are listed in Table 1. They agree well with the values found by Deal and Grove[9] and other authors.

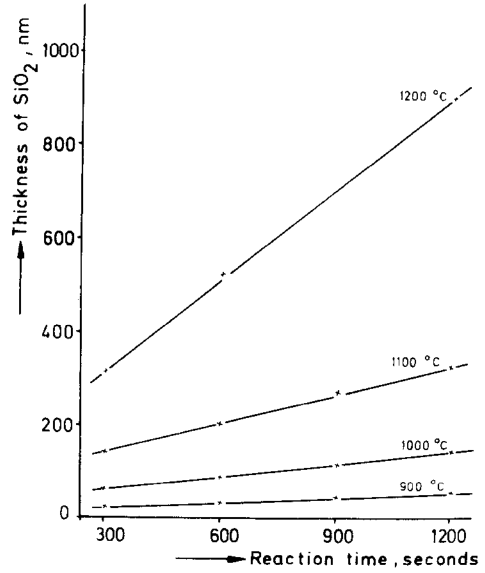

Fig. 7. SiO_2 film thickness calculated from the converted Si_3N_4-thickness as a function of time at various temperatures, on exposure to a mixture of 5 l/h O_2 and 50 l/h N_2 saturated with PBr_3 at 25°C.

4. DISCUSSION

The conversion of silicon nitride is a process, which can be related, with limitations of course, to the oxidation of silicon. As the linear-parabolic relationship for silicon oxidation, which results from the model of Deal and Grove[9] is well established now, nitride conversion will be pointed out in view of this model. Thereby it must be taken into account that the error in thickness measurement of double layers is larger than in the measurement of SiO_2-monolayers on silicon and that the

Table 1. *Parabolic rate constants for the formation of* SiO_2

Temperature (°C)	$Si \rightarrow SiO_2$ O_2 dry B, cm² sec⁻¹	$Si_3N_4 \rightarrow SiO_2$ O_2 dry B, cm² sec⁻¹	$Si_3N_4 \rightarrow SiO_2$ O_2 wet B, cm² sec⁻¹
1000	$3{\cdot}4 \times 10^{-14}$		
1100	$6{\cdot}7 \times 10^{-14}$		$3{\cdot}8 \times 10^{-15}$
1200	$1{\cdot}15 \times 10^{-13}$		
1230	$1{\cdot}4 \times 10^{-13}$	$1{\cdot}8 \times 10^{-15}$	
1260	$1{\cdot}6 \times 10^{-13}$	$3{\cdot}2 \times 10^{-15}$	

thicknesses to be measured are much smaller, due to the very slow proceeding conversion process. Additionally, another gas is produced in the interface reaction, in the case of conversion in dry oxygen probably nitrogen, which must diffuse through the oxide to the surface, in opposite direction to the oxidant.

From the results presented in Figs. 5 and 6, those obtained at 1230°C and 1260°C with dry oxygen and those obtained at 1100°C with wet oxygen, can be evaluated by means of equations (8) and (11) clearly, as is shown in Figs. 8 and 9. The values of B, as calculated from Figs. 8 and 9 are listed in Table 1. According to equation (10) B depends on the product of the equilibrium concentration C^* of the oxidant in the oxide and the effective diffusion constant D_{eff}. In pointing out data of Flint[18] on silicon oxidation in mixtures of oxygen and argon, Deal and Grove could also show that the coefficient A and hence the effective diffusion constant D_{eff} is independent of the oxygen partial pressure. This implies that the reduction of the rate constant B in the oxidation of silicon by oxygen-inertgas mixtures is caused by the reduction of C^* through the influence of the inertgas. Assuming now that nitrogen and argon behave likewise, that means the nitrogen diffusing to the surface has no influence on the diffusion constant of the oxidant, and comparing the values of B, it follows from equation (10) that the equilibrium concentration of the oxidant in the oxide C^* is much lower in the case of nitride conversion than in the case of silicon oxidation.

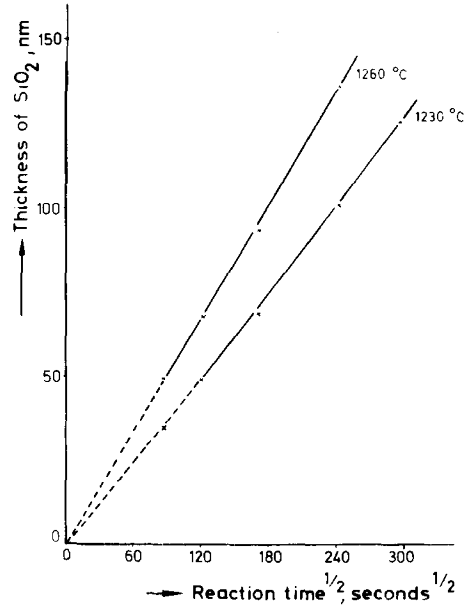

Fig. 8. Film thickness as a function of $\sqrt{t}$ for the dry oxidation of Si_3N_4 at temperatures $T = 1230°C$ and 1260°C.

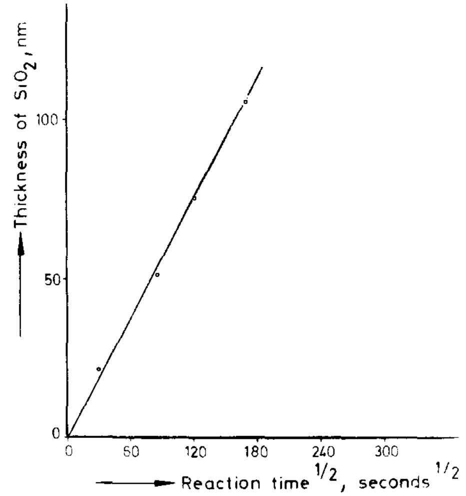

Fig. 9. Film thickness as a function of $\sqrt{t}$ on conversion of Si_3N_4 in wet oxygen. $T = 1100°C$.

From equation (6) it follows that the growth rate of the oxide decreases with decreasing equilibrium concentration C^*. Thus it may be understood that the nitrogen which is formed during the chemical reaction at the interface oxide/nitride and which must diffuse through the oxide to the surface lowers the concentration of the oxidant in the oxide, and thus the growth rate of oxide in a pure oxygen ambient is lower in the conversion of the nitride than in the oxidation of silicon.

It should be mentioned additionally that a lower oxide growth could be caused also by consumation of part of the oxidant through the formation of nitrogen oxides. But it would be not possible to explain the low conversion rate quantitatively by this effect.

Comparing the value of B for conversion of silicon nitride in wet oxygen with the value of B as given by Deal and Grove[9] for wet oxidation of silicon ($1{\cdot}4 \times 10^{-12}$ cm² sec⁻¹ at 1100°C) it becomes obvious that the lower reaction rate in this case depends on the lower equilibrium concentration of oxidant in the oxide C^* also. The gas which is formed at the nitride/oxide interface and which is the reason for the lower value of C^* is H_2 and N_2 or NH_3, according to one of the following equations

$$Si_3N_4 + 6H_2O \longrightarrow 3SiO_2 + 6H_2 + 2N_2$$

$$Si_3N_4 + 6H_2O \longrightarrow 3SiO_2 + 4NH_3$$

from which the first reaction is thermodynamically preferred, as Fig. 1 shows.

Taking into account that the error in the measured SiO_2-thicknesses on silicon nitride is larger than in the measurement of SiO_2 on silicon because of the more complicated system, it is possible that the formation of thin films at low temperatures obeys the same linear-parabolic relationship. But it is obvious that in the beginning the conversion of silicon nitride is quite different from oxidation of silicon and its possible that another relationship is valid. More can be said, when the formation of thin films in nitrogen–oxygen mixtures is investigated.

Quite different are the results obtained under the influence of phosphorus pentoxide. The conversion is much faster though the partial pressure of oxygen is lower in comparison to the conversion of nitride in dry oxygen. As was established earlier [13] the formation of silicon dioxide under the influence of P_2O_5 takes place through the intermediate formation of silicon phosphide:

$$3Si_3N_4 + 2P_2O_5 \longrightarrow 5SiO_2 + 4SiP + 6N_2$$

Silicon phosphide cannot be isolated in the presence of oxygen, because it is immediately oxidized to form SiO_2 and P_2O_5 i.e.

$$4SiP + 9O_2 \longrightarrow 2P_2O_5 + 4SiO_2$$

The formation rate is dependent on the oxygen partial pressure, as a comparison with the authors' previously obtained data[13] shows. By a flux of about 2·5 l/hr O_2 and 50 l/hr N_2 175 nm SiO_2 are produced in 10 min whereas 500 nm were formed in a gas mixture of 5 l/hr O_2 and 50 l/hr N_2.

Diffusion of oxygen or phosphorus pentoxide apparently not predominates in controlling the reaction rate as can be derived from the linear relationship. This becomes more obvious by the fact that the phosphorus concentration is homogeneous in the forming glass[13]. The rate limiting step may be a reaction at one of the two phase boundaries, possibly at the interface nitride/glass. A further interesting point is that positive thickness values, different for the different temperatures, would be obtained on extrapolation to time $t = 0$. This means no doubt, that a different mechanism controls the reaction rate at the beginning.

SUMMARY

Conversion of silicon nitride into silicon dioxide in dry oxygen proceeds slower than conversion in wet oxygen. By the influence of phosphorus pentoxide the conversion is rather accelerated. In comparison to the thermal oxidation of silicon the rate of silicon dioxide formation is slower when silicon nitride is converted, in dry oxygen as well as in wet oxygen.

In pointing out the results in view of the model given by Deal and Grove for silicon oxidation it follows that this effect possibly is due to a lower concentration of oxidant in the oxide which may be caused by nitrogen produced in the interface reaction.

Acknowledgements—The authors are indebted to Prof. Dr. K. Fränz, Dr. H. Kallweit and Dr. A. F. Bogenschütz for the promotion of this work. They also thank Miss R. Hannawald for the performance of numerous experiments. This work was supported by the Ministry of Education and Science.

REFERENCES

1. R. Newman, N. C. Tombs, H. A. R. Wegener *Sperry eng. Rev.* **2**, 2 (1966).
2. S. M. Hu, *J. electrochem. Soc.* **113**, 693 (1966).
3. V. Y. Doo, *IEEE Trans. Electron Devices* **13**, 561 (1966).
4. F. K. Heumann, D. M. Brown, E. Mets, *J. electrochem. Soc.* **115**, 99 (1968).
5. N. C. Tombs, F. A. Sewell, *J. electrochem. Soc.* **115**, 101 (1968).
6. F. Morandi, *Elektronikpraxis* **5**, 2 (1970).
7. D. A. Oliver, *High Temperatures in Aeronautics*, Proc. Symp. p. 377 (1962).
8. JANAF Interim Thermochemical Tables, The Dow Chemical Co. Midland, Michigan.
9. B. E. Deal, S. S. Grove, *J. appl. Phys.* **36**, 3770 (1965).
10. W. A. Pliskin, *IBM Jl. Res. Dev.* **10**, 198 (1966).
11. P. J. Burkhardt, L. V. Gregor, *Trans. Metall. Soc. A.I.M.E.* **236**, 299 (1966).
12. A. G. Revesz, K. G. Zaininger, R. J. Evans, *Appl. Phys. Letters* **8**, 57 (1966).
13. I. Fränz, W. Langheinrich, *Solid-St. Electron.* **12**, 955 (1969).
14. I. Fränz, W. Langheinrich, K. H. Löcherer, *Telefunkenzeitung* **37**, 194 (1964).
15. I. Fränz, W. Langheinrich, *Solid-St. Electron.* **11**, 59 (1968).
16. I. Fränz, W. Langheinrich, *Solid-St. Electron.* **11**, 987 (1968).
17. I. Fränz, W. Langheinrich, *Solid-St. Electron.* **13**, 807 (1970).
18. P. S. Flint, Spring Meeting of the Electrochem. Soc., Los Angeles (1962). Abstract **94**.

APPENDIX

Deal and Grove[9] have shown that the oxidation of silicon can be described by a physical model which includes three fluxes

$$F = F_1 = F_2 = F_3 \tag{1}$$

when the process is considered under steady state conditions, that means when an initial oxide film had already been built up. The flux of the oxidant from the gas to the vicinity of the outer surface is taken to be

$$F_1 = h(C^* - C_0) \tag{2}$$

where h is a gas phase transport coefficient
C_0 is the concentration of the oxidant at the outer surface of the oxide given at any time, and
C^* is the equilibrium concentration of the oxidant in the oxide.

The flux of the oxidant across the oxide layer is assumed to be given by Fick's law.

$$F_2 = -D_{\mathrm{eff}}(\partial C/\partial x) \tag{3}$$

at any point x within the oxide layer, where D_{eff} is the effective diffusion coefficient and $\mathrm{d}C/\mathrm{d}x$ is the concentration gradient of the oxidising species in the oxide. As in the steady state $\mathrm{d}F_2/\mathrm{d}x = 0$, F_2 is given by

$$F_2 = D_{\mathrm{eff}} \frac{(C_0 - C_i)}{x_0} \tag{4}$$

where C_i is the concentration of the oxidant near the oxide silicon interface. The flux, corresponding to the oxidation reaction is expressed by the relation

$$F_3 = K \times C_i \tag{5}$$

If $N1$ is the number of oxidant molecules incorporated into a unit volume of the oxide layer the growth rate is

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \frac{F}{N1} = \frac{kC^*/N1}{1 + k/h + kx_0/D_{\mathrm{eff}}} \tag{6}$$

It is now assumed that the total oxide thickness x_0 consists of two parts, an initial thickness x_i which is built up before the above treatment becomes valid, and the thickness of the step under consideration, that means

$$x_0 = x_i \text{ at } t = 0 \tag{7}$$

From equation (6) then follows

$$x_0^2 + Ax_0 = B(t + \tau) \tag{8}$$

where

$$A = 2D_{\mathrm{eff}}(1/k + 1/h) \tag{9}$$

$$B = 2D_{\mathrm{eff}}C^*/N1 \tag{10}$$

and

$$\tau = (x_i^2 + Ax_i)/B$$

At relatively large times equation (8) reduces to

$$x_0^2 \cong B \times t \tag{11}$$

At relatively small times, equation (8) reduces to

$$x_0 \cong \frac{A}{B}(t + \tau) \tag{12}$$