JOURNAL OF MATERIALS SCIENCE 22 (1987) 2887–2892

Oxidation of silicon by oxygen: a rate equation

ROBERT H. DOREMUS, ANDREW SZEWCZYK*
Materials Engineering Department, Rensselaer Polytechnic Institute, Troy, New York 12180, USA

A new rate equation fits data on the oxidation of silicon by oxygen. It can be derived from a model of molecular diffusion in the oxide layer and the influence of strain on this diffusion.

1. Introduction

The oxidation of silicon is a vital part of making electronic components; a better understanding of the mechanisms of oxidation would improve the processing and properties of devices. A widely used relation between the thickness L of an oxide layer and oxidation time t at a constant temperature is the linear–parabolic equation

$$L^2 + AL = Bt \quad (1)$$

in which A and B are parameters not dependent on thickness or time. In their pioneering work Deal and Grove [1] found good agreement between Equation 1 and the oxidation of silicon by water, and other workers have confirmed these results.

When silicon is oxidized by dry oxygen, Equation 1 is not followed, especially at early times. To account for this discrepancy Deal and Grove [1] introduced an additional parameter τ, giving an equation with three adjustable parameters:

$$L^2 + AL = B(t + \tau) \quad (2)$$

However, thicknesses of silica films formed on silicon by oxygen do not fit Equation 2 at early times, as shown in Fig. 1. Furthermore, there has been no satisfactory explanation of the experimental values of τ.

The main purpose of the present paper is to propose an equation with three adjustable parameters that can be fitted to all experimental results on the oxidation of silicon. This equation is

$$L^2 + AL - AC \ln\left(1 + \frac{L}{C}\right) = Bt \quad (3)$$

with parameters A, B and C. Methods of fitting experimental data to this equation are discussed in the next section, and then some calculated parameters for selected experimental results are listed. A theory of oxidation is then described that results in Equation 3. A detailed consideration of theoretical parameters and their comparison with experimental data is not included in this paper.

2. Fitting equations to data

A useful graph for comparing experimental data with Equation 3 is that of $\log L$ against $\log t$. Some experimental data on such a plot are in Fig. 2, and calculations from Equations 2 with particular parameters A, B and C are in Fig. 3. In both experimental and calculated results the $\log L$ against $\log t$ plots are linear. This comparison gives support to the validity and usefulness of the empirical functional form of Equation 3. Fig. 3 shows that the slope of the $\log L$ against $\log t$ plot depends on the ratio A/C of parameters in Equation 3. Thus the $\log L$ against $\log t$ plots provide a convenient way to determine this ratio, and can reduce the number of parameters to be determined by direct fitting to two. A convenient form of Equation 3 for calculation of these slopes is in terms of the dimensionless thickness L/C:

$$\left(\frac{L}{C}\right)^2 + \frac{A}{C}\left(\frac{L}{C}\right) + \frac{A}{C}\ln\left(1 + \frac{L}{C}\right) = \frac{Bt}{C^2} \quad (4)$$

Calculations in this form are in Fig. 3. Values of slopes on $\log L$ against $\log t$ plots for different values of A/C, as calculated with Equation 4, are in Table I. A graph based on Table I can be used to determine A/C from experimental plots of $\log L$ against $\log t$.

Fitting of Equation 3 to experimental data is not simple, even if the A/C ratio is known, because the equation is non-linear in the parameters to be determined (B and A or C). Various direct non-linear regression methods [2] were tried to fit the data, but it was not possible to arrive at reliable values of the parameters with the available experimental data, because of the lack of sensitivity of the minimization process on the parameters.

An alternative method is to calculate values of t for measured values of L for many assumed sets of parameters with a computer. A software package [3] was used to make such calculations, and then the set of parameters for the sum of deviation squares were chosen. If $\hat{t}_i$ is a time calculated from Equation 3 for an experimental L_i value, and t_i is the experimental time, then the sum of squares is

$$\sum_i (\hat{t}_i - t_i)^2 \quad (5)$$

A necessary assumption for the validity of a regression calculation is that the variance in the testing variable is constant [4]. The variance in time values is estimated by $(\hat{t}_i - t_i)^2$ for each time t_i at which thickness is measured. A sample of calculations is given in Table II for a particular set of parameters A, B and C. The data in the table were taken from two separate investigations [5, 6] of the oxidation of silicon in 1 atm

*Current address: IBM, East Fishkill Facility, Hopewell Junction, New York 12583.

0022–2461/87 $03.00 + .12 © 1987 Chapman and Hall Ltd.

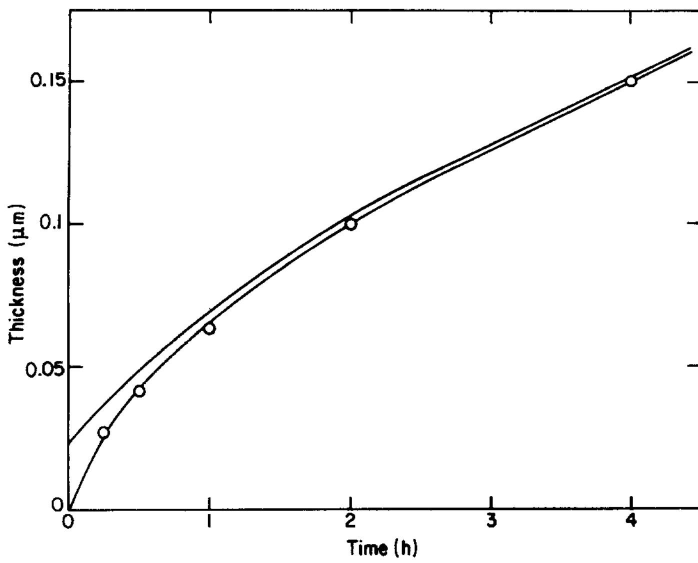

Figure 1 Thickness of SiO_2 layer on silicon with time of oxidation in 1 atm oxygen at 1000° C. (○) Data of Deal and Grove [1]; upper line, calculated from Equation 2 and parameters given by Deal and Grove [1]; lower line, from Equation 3 with parameters in Table III.

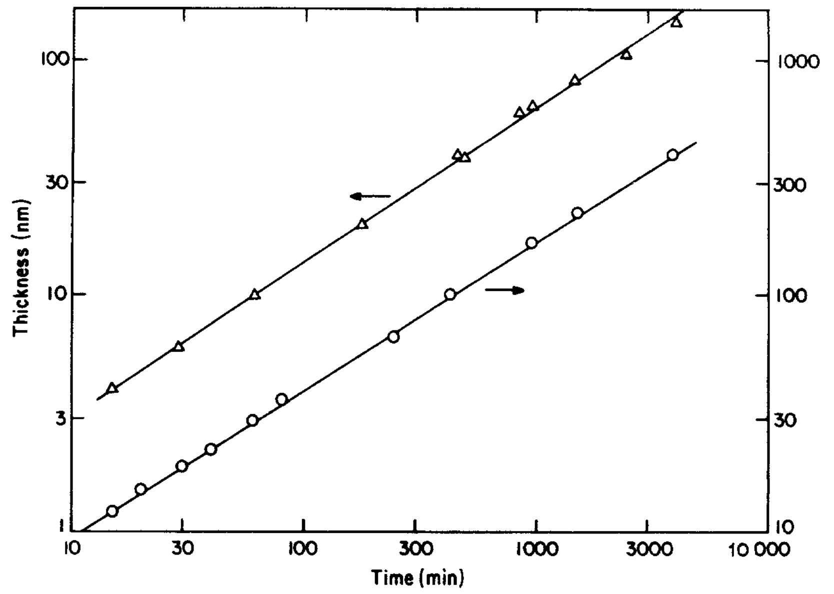

Figure 2 Log (thickness) of SiO_2 layers on silicon as a function of log (time) during oxidation in 1 atm oxygen. (○) At 900° C data from Table II; (△) at 800° C, data from Massoud [6] and Deal and Grove [1].

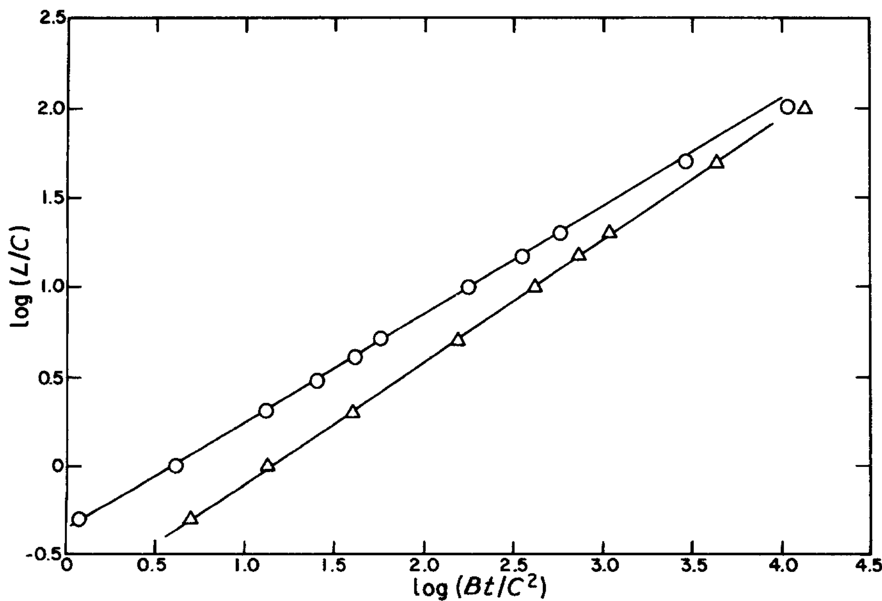

Figure 3 Dimensionless thickness L/C as a function of dimensionless time, log–log plot. From Equation 3 (○) $A/C = 10$, (△) $A/C = 40$.

TABLE I Slopes of $\log L$ against $\log t$ plots as a function of A/C

A/C	Slope
2	0.53
5	0.57
10	0.60
15	0.62
20	0.64
40	0.69

oxygen at 900° C to provide values over a wide range of times and oxide thicknesses. At one hour the two investigations found layer thicknesses of 29.3 and 29.5 nm, respectively. Table II shows that the residual values $(\hat{t}_i - t_i)$ are approximately proportional to $\hat{t}_i$ values. Thus the variance in t is not constant, but is roughly proportional to $(\hat{t}_i)^2$. For this functionality the time variables must be transformed to $\log t$ to give a constant variance [4]. The residuals $(\log \hat{t}_i - \log t_i)$ from such a transformation are also shown in Table II and are plotted in Fig. 4, which shows that for this transformation the residuals and therefore the variance $(\log \hat{t}_i - \log t_i)^2$ has no trend as a function of $\log \hat{t}_i$. Thus the sum of squares that must be minimized for a best fit is

$$\sum_i (\log \hat{t}_i - \log t_i)^2 \qquad (6)$$

Experimental measurements of oxide layers on silicon show that an oxide layer about 1 to 2 nm thick is present on cleaned silicon surfaces [6]. This initial thickness L_0 can be included in Equation 3 as follows:

$$L^2 - L_0^2 + A(L - L_0) - AC \ln \left[\frac{1 + (L/C)}{1 + (L_0/C)}\right] = Bt$$

or

$$L^2 + AL - AC \ln \left(1 + \frac{L}{C}\right) - \left[L_0^2 + AL_0 - AC \ln \left(1 + \frac{L_0}{C}\right)\right] = Bt \qquad (7)$$

TABLE II Deviations of times calculated from Equation 3 and experimental times for oxidation of (1 1 1) silicon in 1 atm oxygen at 900° C

Experimental		Calculated time, $\hat{t}$ (min)*	$\hat{t}_i - t$ (min)	$\log \hat{t} - \log t$
Time t_i (min)	Film thickness (nm)			
Data of Massoud [6]				
0	1.2			
5	5.8	4.1	−0.9	−0.086
10	9.0	9.1	−0.9	−0.041
15	12.1	15.1	+0.1	+0.003
20	15.0	21.6	+1.6	+0.039
30	19.0	31.9	+1.9	+0.027
40	22.2	40.9	+0.9	+0.010
50	26.4	54.0	+4	+0.033
60	29.3	63.5	+3.5	+0.025
80	36.6	90.0	+10	+0.051
Data of Deal et al. [5]				
60	29.5			
240	66.2	224	−16	−0.030
420	100	415	−5	−0.005
960	166	(960)		
1500	224	1571	+71	+0.020
3890	398	4238	+398	+0.052

*Fitting parameters in Equation 3: $A = 0.15\,\mu$m, $B = 0.003\,\mu\text{m}^2\,\text{h}^{-1}$, $C = 0.01\,\mu$m, $L_0 = 0.0012\,\mu$m. B found by matching calculated and experimental data at 960 min.

where the term in square brackets is not a function of L or t. In fitting the equation the term in square brackets must of course be calculated for each pair of A and C values. This term is necessary only for thickness measurements below about 0.05 μm.

Logarithmic terms have been included in previous rate equations for the oxidation of silicon [7–9], especially for oxidation by water and oxygen mixtures [8, 9]. The form of these equations was different from that of Equation 3, and Irene and Ghez [9] concluded that the logarithmic term " accounted for only a few percent". The logarithmic term in Equation 3 can be of substantial importance, especially for the thinner films.

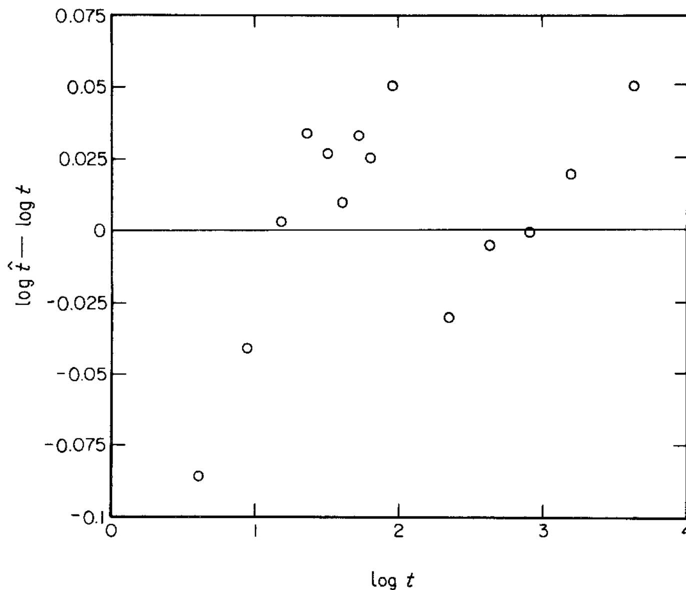

Figure 4 Residuals $(\log \hat{t} - \log t)$ of experimental times t and times $\hat{t}$ calculated from Equation 3 with the parameters in Table II.

TABLE III Calculated parameters A, B and C in Equation 3 for the oxidation of silicon by oxygen

Temperature (° C)	A/C	A (μm)	C (μm)	B ($\mu m^2 h^{-1}$)	Data references
1200	small	–	–	0.042	[1]
1100	4	0.11	0.027	0.022	[1]
1000	13	0.09	0.007	0.0086	[5]
900	17	0.07	0.004	0.0024	[5, 6]
800	26	0.07	0.0027	0.00054	[1, 6]

3. Calculations from experimental data

To show the validity and usefulness of Equation 3, experimental data for the oxidation of silicon by oxygen at different temperatures were fitted with the equation, and the resulting parameters compared with those of previous investigations (see Table III). At 800 and 900° C the data of different investigators were combined to give a wider range of layer thickness. At these temperatures there was overlap and close agreement at intermediate times and thicknesses, as shown in Table II. The parameters in Table III were taken from minimum values of the sum of squares of Equation 6.

The values of the coefficient B are plotted in Fig. 5. At the highest temperatures (1200 and 1100° C) the slope corresponds to an activation energy of about 26 kcal mol^{-1} (109 kJ mol^{-1}), which is close to the measured value of 27 kcal mol^{-1} (113 kJ mol^{-1}) by Norton [10] for the diffusion of molecular oxygen in vitreous silica, and the absolute value of B corresponds to the calculated one [11]. At lower temperatures the values of B become lower than expected, and are also lower than the values reported by Deal and Grove [1]. A very similar anomalous decrease in B values occurred for the oxidation of silicon by water

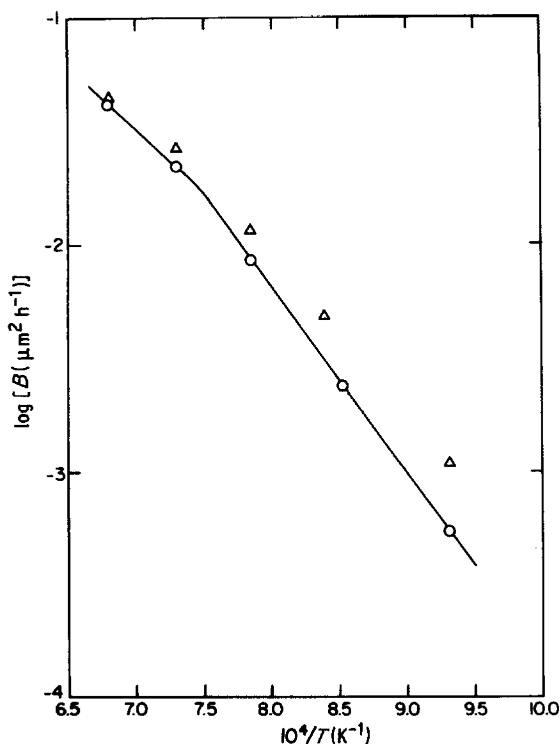

Figure 5 Log B, the parabolic rate parameter, as a function of $1/T$: (○) calculated in the present work, (△) reported by Deal and Grove [1].

[12], and was explained as resulting from strain in the silica layer [13]. It seems likely that a similar explanation is valid for oxidation of silicon by oxygen.

The values of the parameter A in Table III increase somewhat with increasing temperature, and are perhaps nearly constant. The ratio B/A is often considered as the "linear coefficient" and is found to give a linear plot of $\log(B/A)$ against $1/T$. The results in Table III do not give this temperature dependence of B/A, and raise doubts about the significance of activation energies calculated from B/A values.

The values of the parameter C increase with temperature by about an order of magnitude from 800 to 1100° C. Detailed discussion of the theoretical significance of the parameters A and C is reserved for future work. A start is given in the next section in which Equation 3 is derived from a model of the influence of strain on molecular diffusion in the silica layer.

It is difficult to fit the experimental data reliably to a non-linear equation such as Equation 3; many sets of parameters A, B and C give nearly the same sum of squares of deviations. It would be very desirable to have experimental results of a particular experimental arrangement over a wide range (10 nm to 5 μm?) of layer thicknesses.

4. Equations for strain-influenced oxidation

The linear–parabolic Equation 1 can be derived by assuming that there are two processes involved in oxidation, diffusion through the oxide film and a chemical reaction at an oxide interface. For water or oxygen as oxidants many experiments have established that the oxide grows at the silicon–oxygen interface, so the oxidant must diffuse through the oxide layer to the silicon. This diffusion process has been modelled as resulting from the diffusion of molecular water or oxygen through the amorphous silica layer [11]. This mechanism has been confirmed by a variety of independent experiments, and is widely accepted [14, 15]. In this model the diffusion coefficient D of oxidant through the oxide is related to the parameter B by

$$B = 2C_L D/\varrho \tag{8}$$

in which C_L is the concentration of molecular oxygen or water dissolved in the silica at the silica–gas interface and ϱ is the density of the oxide in the same concentration units. Strain in the oxide layer influences the molecular diffusion of oxygen or water in the layer. Modified expressions for A and B for water and oxygen are derived below, and an additional parameter C for the oxidation of silicon in oxygen is introduced in place of τ in Equation 2. References and more details for the derivation are given elsewhere [13].

It is assumed that there is a compressive strain in the oxide layer that decreases linearly from the silicon–oxide interface to the oxide–gas surface, where the strain is negligibly small. The strain ε is then

$$\varepsilon = \varepsilon_0\left(1 - \frac{X}{L}\right) \tag{9}$$

where ε_0 is the strain at the oxide–silicon interface

($X = 0$) and X is the distance from this interface in the oxide layer.

It is also assumed that the diffusion coefficient D of molecular oxygen or water in the oxide depends on the exponential of the strain:

$$D = D_L \exp(-k\varepsilon) \quad (10)$$

where D_L is the diffusion coefficient when there is no strain ($X = L$), and k is a parameter measuring the sensitivity of the diffusion coefficient to strain. In transition state theory $k = EV^*/RT$, where E is Young's modulus and V^* is an "activation volume".

The distance-dependent diffusion coefficient of Equation 9 is now used in the equations for the diffusion-controlled growth of an oxide film. It is assumed that the concentration of oxygen or water dissolved in the oxide at the oxide–gas surface is held constant at the value c_L by oxidant in the gas phase, and the concentration at the silica–silicon interface is small, effectively zero.

Furthermore a steady-state diffusion profile results, because the film grows slowly compared with the diffusion time in it. The constant flux J of oxidant through the film is

$$J = D\frac{dc}{dX} = D_L \exp\left[-k\varepsilon_0\left(1 - \frac{X}{L}\right)\right]\frac{dc}{dX} \quad (11)$$

where c is the concentration at distance X from the oxide–silicon interface. To find the concentration profile in the oxide, Equation 11 is integrated with the boundary condition $c = c_L$ at $X = L$ to give

$$\frac{c}{c_L} = \frac{1 - \exp(-k\varepsilon_0 X/L)}{1 - \exp(-k\varepsilon_0)} \quad (12)$$

This equation gives an exponential profile of concentration with distance X in the oxide. The flux is

$$J = \frac{c_L D_L k\varepsilon_0}{L[\exp(k\varepsilon_0) - 1]} \quad (13)$$

If $k\varepsilon_0 \leqslant 1$, the exponential can be expanded and only the first three terms used, so that

$$J \approx \frac{c_L D_L k\varepsilon_0}{L[k\varepsilon_0 + (k\varepsilon_0)^2/2]} = \frac{c_L D_L}{L[1 + (k\varepsilon_0/2)]} \quad (14)$$

Since the flux $J = \varrho(dL/dt)$, the differential equation for film growth with strain is

$$\frac{dL}{dt} = \frac{c_L D_L}{L[1 + (k\varepsilon_0/2)]\varrho} \quad (15)$$

For cuprous oxide (Cu_2O) films on copper, Borie *et al.* [16] found that the value of ε_0 in Equation 9 became smaller as the film thickness increased; ε_0 was inversely proportional to film thickness for films of thickness from 1.4 to 4.4 nm. It is assumed therefore that ε_0 depends on film thickness in this way:

$$\varepsilon_0 = 2q\left(1 + \frac{l}{L + C}\right) \quad (16)$$

where q is a dimensionless coefficient and l and C have dimensions of length. Substitution of Equation 16 into Equation 15 gives

$$\frac{dL}{dt} = \frac{c_L D_L/\varrho}{L(1 + kq) + [kqlL/(L + c)]} \quad (17)$$

Integrating Equation 16 from $L = 0$ at $t = 0$ gives Equation 3, with the coefficients A, B and B/A equal to

$$B = \frac{2l_L D_L}{\varrho(1 + kq)} \quad (18)$$

$$A = \frac{2kql}{1 + kq} \quad (19)$$

$$\frac{B}{A} = \frac{c_L D_L}{\varrho kql} \quad (20)$$

At a particular temperature the coefficients B and B/A are proportional to c_L in Equations 18 and 20, and hence to the partial pressure of oxygen or water in the gas phase as found experimentally for the oxidation of silicon.

If $kq \ll 1$ (strain effects are small), $B = 2c_L D_L/\varrho$ and $A = 2kql$. These relations are valid for the oxidation of silicon above about 950° C for water oxidation and about 1100° C for oxidation by oxygen. In these temperature ranges the activation energy of B for oxidation of silicon by water or oxygen is equal to the activation energy for molecular diffusion of these gases in bulk vitreous silica [2].

At lower temperatures the factor B for the oxidation of silicon by water [12] and oxygen (present work) is lower than expected from Equation 8; in bulk vitreous silica there is no corresponding change in the activation energy for the diffusion of water. Furthermore there is no anomalous decrease in B/A for water at these temperatures [12]. Equations 18 and 20 give an explanation for these results. At lower temperatures strain effects become important and the factor kq becomes significant compared with unity. Then B is smaller than expected without the kq factor, but B/A still has the same temperature dependence. A quantitative comparison between experimental and calculated values of B is given elsewhere [13] and shows good agreement.

Acknowledgement

This work was supported by NSF-MRG Contract No. DMR-8510617.

References

1. B. E. DEAL and A. S. GROVE, *J. Appl. Phys.* **36** (1965) 3770.
2. N. R. DRAPER and H. SMITH, "Applied Regression Analysis", (Wiley, New York, 1981) Ch. 10.
3. B. J. KORITES, "Data Plotting, Software for Micros" (Kem Publications, Duxbury, Massachusetts).
4. D. C. MONTGOMERY and E. A. PECK, "Introduction to Linear Regression Analysis", (Wiley, New York, 1982) p. 88ff.
5. B. E. DEAL, D. W. HESS, J. D. PLUMMER and C. P. HO, *J. Electrochem. Soc.* **125** (1978) 339.
6. H. Z. MASSOUD, PhD thesis, Stanford University (1983).
7. J. R. LIGENZA and W. G. SPITZER, *J. Phys. Chem. Solids* **14** (1960) 131.
8. K. HIRABAYASKI and J. IWAMURA, *J. Electrochem. Soc.* **120** (1973) 1595.

9. E. A. IRENE and R. GHEZ, *ibid.* **124** (1977) 1757.
10. F. J. NORTON, *Nature* **191** (1961) 701.
11. R. H. DOREMUS, *J. Phys. Chem.* **80** (1976) 1773.
12. R. R. RAZOUK, L. N. LIE and B. E. DEAL, *J. Electrochem. Soc.* **128** (1981) 2214.
13. R. H. DOREMUS, *Thin Solid Films* **122** (1984) 191.
14. N. F. MOTT, *Proc. R Soc.* **A376** (1981) 207.
15. *Idem, Phil. Mag.* **45** (1981) 323.
16. B. BORIE, C. J. SPARKS and J. V. CATHCART, *Acta Metall.* **10** (1962) 691.

Received 22 September
and accepted 15 December 1986