

Materials Science and Technology

A Comprehensive Treatment

Edited by
R.W. Cahn, P. Haasen, E. J. Kramer

Volume 16



Processing of Semiconductors

Volume Editor
K. A. Jackson

vch

1 Silicon Processing

John G. Wilkes †

Formerly with Philips Components Ltd., Southampton, U.K.

List of Symbols and Abbreviations	2
1.1 Introduction	5
1.2 Metallurgical-Grade Silicon	7
1.3 Semiconductor Grade Polycrystal Silicon	11
1.3.1 The Chlorosilane Route	11
1.3.2 The Silane Route	14
1.3.3 The Polycrystal Silicon Market	16
1.4 Single Crystal Silicon	17
1.4.1 Float-Zoned Silicon	17
1.4.2 Neutron Transmutation Doped Silicon	19
1.4.3 Carbon and Nitrogen in Float-Zoned Silicon	20
1.4.4 Periodic Crystal Growth	22
1.5 Czochralski Silicon	25
1.5.1 Dislocation-Free Silicon	26
1.5.2 Constitutional Supercooling	28
1.5.3 The Incorporation of Carbon and Oxygen	30
1.5.4 Magnetic Czochralski Silicon	33
1.5.5 The Commercial Scaling of Czochralski Silicon	34
1.6 Slice Fabrication	36
1.6.1 Mechanical Damage in Silicon	36
1.6.2 Polishing and Cleaning	43
1.7 Oxygen in Czochralski Silicon	44
1.7.1 The Behavior of Oxygen in Silicon	44
1.7.2 The Precipitation of Oxygen in Silicon	45
1.7.3 Thermal Donors and Enhanced Diffusion	49
1.8 Crystal Engineering	51
1.8.1 Extrinsic Gettering in Silicon	51
1.8.2 Intrinsic Gettering in Silicon	53
1.9 Acknowledgements	59
1.10 References	60

List of Symbols and Abbreviations

A_n	Fourier series coefficient
a_0	lattice constant (for Si, $a_0 = 5.42 \text{ \AA}$)
A_0	constant
B	slice bow depth
C	concentration
C_H	crystal habit
C_l	concentration in liquid
C_p	concentration (of oxygen in oxide) in particle
C_s	concentration in solid
C_{es}	equilibrium solid solubility concentration
C_0	initial concentration
d	diameter
D	diffusion coefficient
E_c	activation energy for the formation of a particle of critical radius
f_i	cut-off frequency
\bar{f}_i	mean value of f_i
ΔF_v	volume free energy change of a precipitate
F_x, F_y, F_z	magnitudes of the forces generated at the edge during sawing
g	fraction of melt solidified
h	height
ΔH	enthalpy of reaction
$I(\theta)$	detector signal
I	interstitial
I_{cb}	collector-base current
k	Boltzmann constant
k_{eff}	effective distribution coefficient
k_0	equilibrium distribution coefficient
N	number of particles
N_c	number of particles of critical radius
n_f, n_t	fast neutron, thermal neutron
n_1	number of oxygen atoms in axial bonds
n_2	number of oxygen atoms in other bonds
$[O_i]$	bound interstitial oxygen concentration
Pr	Prandtl number
r	radius
R	radius of the total volume from which oxygen condenses into a precipitate
r_c	critical radius
r_0	radius of a final precipitate particle, small compared with R
t	time
T	absolute temperature
T_m	melting point (Si: 1412 °C)
$t_{1/2}$	half life of radioactive species
t_{Si}	thickness of a silicon slice

ΔT	temperature difference
v	velocity
V	vacancy
v_g	velocity of growth
W_c	intrinsic X-ray signal half width
W_m	measured X-ray signal half width
ΔW	rocking curve broadening
α	alpha particle
α_{\parallel}	absorption coefficient for polarized infrared light parallel to the stress axis
α_{\perp}	absorption coefficient for polarized infrared light perpendicular to the stress axis
γ	gamma particle
δ	boundary layer thickness
ε	strain
θ	test sample angle
θ_B	Bragg angle, X-ray reflection
λ_n	Fourier coefficient (with dimensions of inverse length)
λ_0	constant
ν	kinematic viscosity
σ	surface free energy
τ	relaxation time
τ_n, τ_0, τ_*	relaxation time constants
ω	angular velocity
AC	alternating current
ACR	advanced carbothermic reduction
ASTM	American Society for Testing Materials
BP	boiling point
CMOS	complementary, using both n- and p-type, metal–oxide–silicon device
CVD	chemical vapor deposition
CZ	Czochralski material
DC	direct current
DCS	dichlorosilane
DI	deionized
DRAM	dynamic random access memory
EBE	extended bulk epitaxy
EG	enhanced gettering
FZ	float zoned (material)
HF	high frequency
HI–LO	high temperature–low temperature (heat treatment)
LPCVD	low pressure chemical vapor deposition
MG-Si	metallurgical-grade silicon
MOS	metal–oxide–silicon (device) (n-MOS, p-MOS refer to the dopant type structure employed)

NFZ	nitrogen-doped float zone (material)
NTD	neutron transmutation doping
NTP	normal temperature and pressure
ppba	atomic parts per billion (10^9)
ppma	atomic parts per million
ppt	parts per trillion (10^{12})
psi	pounds per square inch
RF	radio frequency
rpm	rotations per minute
SANS	small angle neutron scattering
SIMS	secondary ion mass spectrometry
SRAM	static random access memory
TCS	trichlorosilane
TD	thermal donor
TIR	total integrated reading (of bow or warp)
UHF	ultrahigh frequency
ULSI	ultra large scale integration
UV	ultraviolet
VLSI	very large scale integration
WCA	water classified alumina

1.1 Introduction

Silicon today is a commodity, its price subject to all the forces of supply and demand in an intensely competitive market, and this has driven the development of high yield processes for the tight tolerance materials demanded. While discrete and power device manufacture calls for some float zoned, and neutron transmutation doped (NTD) silicon; the worldwide compass of integrated circuit manufacture consumes more than 75% of all the semiconductor silicon produced. The development of the product market distribution is shown in Fig. 1-1. Supply of this material is dominated by Czochralski crystal growth, the operational scale of which has increased from charges weighing a few hundred grams, around 1962, to the current units of 60 kilogram and more.

The evolution of the semiconductor industry as we now know it began in the 1950s, when many of the then large electrical companies became involved in the chemistry and metallurgy of Germanium. Their starting point was GeO_2 , the dioxide, which had to be reduced to metal powder, melted, zone refined, and crystals grown, before the machining operations which led to discrete devices. Germanium being an expensive rare element, the ma-

ching itself generated valuable byproduct sludges which had to be recovered. The extreme purity necessary led into problems in chemical and physical analysis, materials of containment, and in general chemical engineering.

In retrospect, very few of these electrical companies possessed either the resources or the experience needed for such work; so when, only shortly afterwards, silicon was introduced, almost all of them took the opportunity to withdraw from the chemical end of the business. Silicon is one of the most abundant elements, and so the sludges are of no economic importance. Henceforth their starting point became the ultrapure polycrystalline silicon from which they made their own single crystal. With time, the number of companies doing even this has steadily declined, until today few of the electronics manufacturers have any involvement in bulk material processing. Indeed most purchase polished slices, cleaned and packaged, furnace ready, for fabrication lines. A number of the device makers still carry out epitaxy and, to that extent alone, retain a residual materials activity.

In modern very large scale integration (VLSI) circuits, lithographic feature sizes have been reduced to $1\text{ }\mu\text{m}$ or less, and use multilevel interconnects to enable the production of high complexity devices of steadily rising chip area. Consequently, as the number of chips per wafer decreases, so there has been an accompanying call for ever larger wafer diameter – to reduce perimeter wastage, and to improve fabrication line yield and throughput, as shown in Fig. 1-2 – hence the continuous need to scale up crystal size, this demanding extremely heavy investment.

This scaling has not been at the expense of quality – in fact quite the reverse. As more has been learnt about the relation-

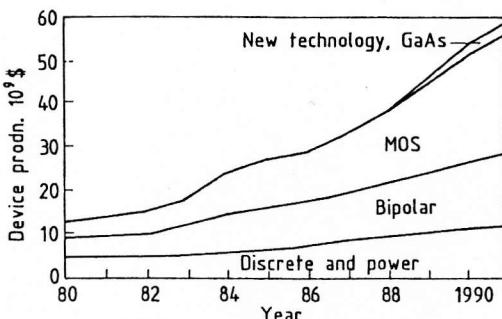


Figure 1-1. Development of the semiconductor product market.

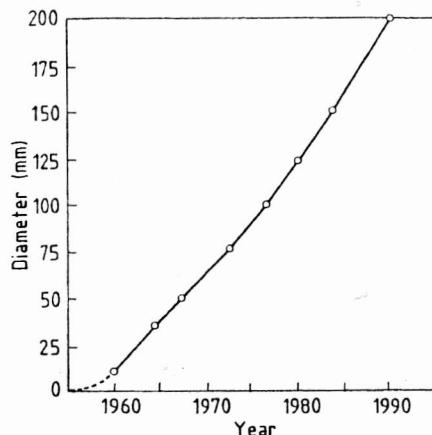


Figure 1-2. The year of introduction of the largest silicon wafers in production. (Note: This trend continues. A very small number of companies, mainly Far East, are now looking at 250 mm diameter possibilities.)

ship between materials properties and the device parameters, so the demand for better performance from the silicon has grown. If one compares a typical purchasing specification of even the mid 1970s, with that in force today for a similar application endproduct, the increase in the number of parameters specified, and the narrowing of virtually all tolerances, is marked. Contributory factors leading to this position include: Fine geometry lithography, needing slices of a flatness not even contemplated ten years ago; cassette, and now robotic, handling techniques, which call for close machining tolerances, and edge rounded slices, to prevent chipping and particles accumulating in ultraclean fabrication equipment. Research into the behavior of oxygen and carbon precipitation in bulk silicon under device furnacing conditions has led to the introduction of new specification parameters, new crystal processing methods, and to the concept of "crystal engineering". Controlled oxide precipitation in slices is carried out, prior

to their use in fabrication lines, to provide sites for the intrinsic gettering of unwanted fast diffusing electrically deleterious impurities, away from the surface layer where the MOS devices are made. Residual mechanical damage sites after the crystal machining provide similar extrinsic gettering sites. For many applications in "crystal engineering" today, combinations of controlled mechanical and oxide precipitate gettering are used together to achieve optimum performance from the silicon, to match the particular device requirements in MOS, CMOS, and bipolar configurations. To achieve this matching it is necessary to examine the total thermal inventory of the multistage fabrication process, in order to select the most appropriate structure.

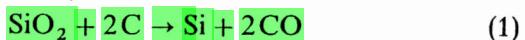
While the early 1990s have seen a general slowdown in world economies, the surge in the personal computer market, linked to the major developments in microprocessor chips, has meant that device production revenues have continued to grow, by 1993, to \$ 60–70 billion (10^9), and are projected to be in the range \$ 150–200 billion by the year 2000. These microprocessor chips are of ever increasing size and complexity; for example the Intel Pentium with around 3 million transistors, running at a speed of 100 million instructions per second, is about 0.5 sq. in. ($\sim 3 \text{ cm}^2$) in area. In memory chips the 16 Mbit DRAM is being followed by the 256 Mbit version, and since each DRAM cell needs one transistor, even with shrinking all dimensions the chip size is increasing. Charge coupled devices for displays also require large-area chips. Such ultra large scale integration (ULSI) applications today are leading the demand for Czochralski silicon wafers with diameters of 200 mm (8"), and inevitably still bigger will follow. For these diameters, the current pullers have to be scaled up

further towards 100–150 kg machines. The larger thermal masses will impose difficulties in the control and uniformity of the dopant, oxygen, and crystal defect concentrations. Finer dimension lithography needs ultra flat wafers, and particulate contamination levels of less than 10 particles greater than 0.1 μm per wafer are expected. Leakage control in large DRAMs requires metallic surface contamination to below 10⁹ atoms/cm. These are severe challenges at the final wafer cleaning and packaging stages.

Overall, larger slices, made to extremely close tolerances by rigidly defined processing, from silicon that conforms to tightly specified criteria with respect to uniformity, dopants, impurities, point defect precipitates, clean surface characteristics, and metrology, are required of the materials vendors. Yet as the market competition is fierce, all of this is wanted at the minimum possible price – a constraint that reflects back through every step in silicon material manufacture. Thus right back in the raw material sector, over the past twenty years the supply of silicon has steadily become concentrated into fewer large specialist merchant vendors, and usually these operate as a division within some much bigger general chemical corporation. Many of the same companies are now involved in the present ruthless shakeout of the parallel gallium arsenide material market, as this sector, in the 1990s, is becoming commercially more significant. Thus the highly competitive commodity environment, which can never be ignored, is the constant background influence against which this chapter is set.

grade silicon, manufactured by the carbo-thermic reduction of silica in an electric arc furnace. Silica, occurring naturally as quartzite, in vein quartz, and in sandstone, and as unconsolidated sands and gravels, is a common mineral with worldwide distribution. Silicon, after oxygen, is the second most abundant element, but does not occur naturally in its elemental form. Silica, either free as in quartz or in the many forms of silicate igneous rocks, constitutes about a quarter of the earth's crust. However, the silicon metal producers demand an ore purity of better than 99% SiO₂, and also place tight restrictions on the allowable concentrations of various impurities present – in particular arsenic, phosphorus, and sulfur – so that often only a small fraction of an ore deposit meets their purity specification. Geologically washed out gravel from river bed deposits, and similarly leached out quartz sands, are a source of very high purity silica. Vast new deposits, yielding quartzite ore of the highest purity available today, have been discovered in Arkansas, U.S.A., from which monocrystals weighing several tons apiece have been displayed in exhibitions worldwide.

In the traditional electric arc furnace process, which has been used for most of this century, chunky quartzite is reacted with carbon, as the reductant, in the forms of coal, coke, or charcoal, which can be a source of at least an order of magnitude greater impurity levels than present in the silica. The overall reaction appears simple:



However, as discussed by Healy (1970), the actual reaction sequence in the different temperature zones of the furnace is far more complex than this, as set out in the schematic diagram of Fig. 1-3.

1.2 Metallurgical-Grade Silicon

The source of the raw silicon used for semiconductor purposes is metallurgical-

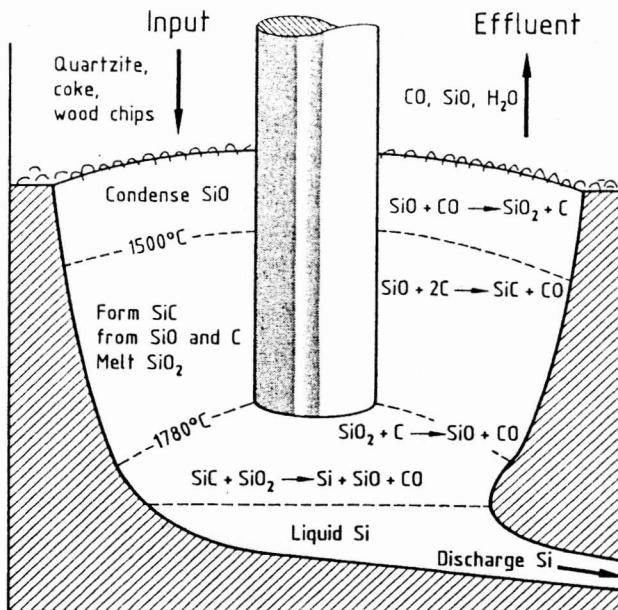
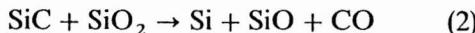
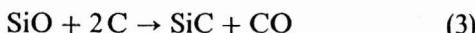


Figure 1-3. Schematic diagram of the submerged-electrode electric arc furnace for the production of metallurgical grade silicon.

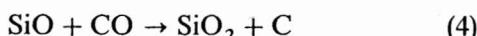
- (a) Towards the bottom of the furnace, in the region of the arc between the electrodes where the temperature can exceed 2000°C , silicon is produced by the reaction



- (b) Above this, at a somewhat lower temperature, around $1700-1500^{\circ}\text{C}$, the rising byproduct gases react to form the intermediate product silicon carbide by



- (c) Nearer to the top, where the temperature falls below 1500°C , as is expected thermodynamically, the reverse reaction predominates:



The input materials are fed into the top of the furnace, while liquid silicon is periodically tapped from the bottom and cast into ingots. If this casting is carried out directionally, under the conditions referred to

as normal freezing, impurity redistribution can be used to effect some purification, following the well equation by Pfann (1952, 1958):

$$C_s = k_{\text{eff}} C_0 (1 - g)^{(k_{\text{eff}} - 1)} \quad (1-1)$$

For the arc process to run properly, it is essential to maintain porosity throughout the charge to allow uniform SiO and CO gas flow, and to permit the escape of CO, some SiO, and H₂O from the top. To assist this wood chips may be included in the feedstock, and the silica must be of a form which does not readily crumble during initial heating in the upper part of the furnace, which could lead to premature fusion and crusting over, with the risk of a dangerous pressure buildup within the charge. Clearly the carbothermic reduction of silica is not a trivial process.

Crossman and Baker (1977) have given a very interesting comparison of the impurities present in typical quartzite and the carbon used, related to the spectrographic analysis of more than 2000 tons of the met-

allurgical-grade silicon produced. Their data, collected into Table 1-1, indicated total impurity levels in the quartzite of around 750 ppm; in the carbon 8000 ppm, and in the resulting metallurgical-grade silicon (MG-Si) 4000–4500 ppm. Within this analysis the two predominant impurities are seen to be aluminum and iron, largely originating from the carbon, and taken together accounting for over 80% of that in the silicon product. Since these results referred to MG-Si to be used for the production of semiconductor grade polycrystal silicon, the importance of the purity of the carbon source is underlined.

Recent developments have focused on improved and cleaner processes, better quality carbon, and efforts to develop quartz sands as an alternative low cost and high purity source. Maintaining charge porosity constitutes the most serious restriction in the operation of the submerged arc furnace, and much attention has been focused on how to meet, or circumvent this problem. In work aimed to reduce drastically the impurities in arc furnace silicon, Dosaj et al. (1978) working at Hemlock

Semiconductor Corp. U.S.A. reported using a high purity silica source together with carbon black powder, pelletized with pure sucrose binder, to obtain MG-Si at 99.99% purity. Although the boron content of the material was relatively low, this particular element tends to be more persistent through the later stages of semiconductor silicon manufacture, and therefore recently there has been interest in exploiting the lower boron content of carbon obtained from pet coke.

The pelletization of upgraded quartz sands can provide very pure silica in a suitable form. This material then has to be agglomerated to lumps, either separately or mixed with carbon powder. This approach has been studied by several groups, including Elkem A/S, Norway, the largest European silicon metal producer, but until now it has only been taken to a development stage. The Siemens advanced carbo-thermic reduction (ACR) process has recently been described by Aulich et al. (1985), in which high purity pelletized quartz sand is reduced by carbon granules, prepared from carbon black briquettes, which had been leached with hot HCl to a purity comparable to that of the silica. Since in an arc furnace about 10% of the carbon comes from the electrode, the effective carbon impurity level was somewhat higher. Nevertheless a substantial overall impurity reduction was achieved.

A more radical approach to overcoming the porosity problem has been the application of DC plasma-arc techniques to the production of ferrosilicon alloys and silicon metal. The most important feature of the plasma-arc furnace here is that it can process ore fines directly, without prior briquetting or pelletization. The potential of this route is supported by the extremely efficient plasma purification of normal MG-Si, by factors of up to 100 000, re-

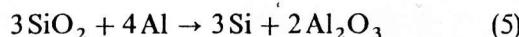
Table 1-1. Impurities in silica, carbon, and metallurgical-grade silicon.

Impurity	Quartzite (ppma)	Carbon ^a (ppma)	MG-silicon ^b (ppma)
Al	620	5500	1570±580
B	14	40	44±13
Cr	5	14	137±75
Fe	75	1700	2070±510
P	10	140	28±6
Others	10	600	—
Mn			70±20
Ni			47±28
Ti			163±34
V			100±47

^a Weighted; ^b average value ± standard deviation.

ported by Armouroux et al. (1986). Developed up to the stage of full scale production of ferrochromium ore, in South Africa, the plasma-arc technology is now being assessed in other countries.

The great evolution of heat from the oxidation of aluminum forms the basis of the Thermit process for the reduction of refractory oxides, such as Cr_2O_3 , and MnO_2 . By the application of this technique to silicon, an entirely new manufacturing route has resulted from the extensive work by Dietl and Holm [see, e.g., Dietl et al. (1981) and Dietl and Holm (1986)] at Wacker Hellotronic, Germany, on the aluminothermic reduction of quartz sand in a liquid flux system ($\text{CaO}-\text{SiO}_2$) at a temperature of 1600–1700°C:



The flux serves simultaneously as a solvent for the byproduct aluminum oxide, and as a liquid–liquid extraction medium. As the silicon is released it is immiscible in the flux and so separates. Since the silicon is of lower density, if floats as the upper layer and at intervals can be poured off into a mold, where controlled normal freezing further separates low segregation coefficient impurities. The silicon made by this novel semicontinuous process is of relatively high purity compared with normal MG-Si. It is characterized by low boron and carbon levels, and after subsequent grinding, acid leaching, and liquid-gas extraction, provides a material that is suitable for solar-cell applications.

During the past decade, the most important economic trend in silicon metal production has been abandonment of the earlier small scale multi-unit plants, having limited productivity, for the use of very much larger electric arc furnaces, commonly dedicated to a particular product, which operate with lower unit costs. A

modern commercial submerged-electrode arc furnace built in a three phase, three electrode configuration, each of these 1.25 m in diameter, and driven from a 24 MW power source, can produce 8000–10000 t/a metallurgical-grade Si at an energy consumption of 12–14 kWh/kg.

The demand for metallurgical-grade silicon is dominated by the iron-steel and aluminum alloy industries, which require 98% purity metal. A somewhat higher quality, of 99%+ purity, is required for conversion into chlorosilanes, the key intermediates in the synthesis of organo-silicon compounds for the silicone industry, leading to products such as oils, resins, lubricants, and water repellants. Although the semiconductor industry wants the highest purity it can obtain, the amount needed still represents only a very small fraction of the world's output. For example the global production of MG-Si in 1986 was just under 600 000 t, from which the organo-silicones consumed about 20%. By contrast, in that same year the production of semiconductor-grade, ultrahigh purity, polysilicon reached about 6000 t – representing a consumption of less than 2% of total MG-Si output, and, significantly, only a fraction of the capability of a single-arc furnace. Western Europe accounts for over half the world capacity, led by Elkem, Norway, and Pechiney, France at 100 000, and 75 000 tons per annum, respectively.

Thus, in summary, while the manufacturers of semiconductor silicon can have only limited influence over the quality of metallurgical silicon, there have been improvements in this product. While these have probably been driven more by the much larger organo-silicon purity needs, some of the companies in the MG-Si industry have established strong links in the semiconductor market, and their contributions are of greater significance when set

against the global background of silicon metal production.

1.3 Semiconductor Grade Polycrystal Silicon

As shown in Table 1-1, metallurgical-grade silicon of 99%+ purity contains, in addition to carbon, the major impurities iron and aluminum at more than 1000 ppma, various transition and other metals – titanium, chromium, etc. – at around 100 ppma, and lesser impurities, including boron and phosphorus, at lower levels of 10's ppma. Semiconductor polysilicon specifications require that all of these be reduced to parts per billion (ppba: 1 in 10^9) levels. For the producers this extremely demanding task has to be achieved economically, to meet the very competitive market pricing pressures which dominate the industry. Early polysilicon plants were built with a capacity of around 100 t/a, but today, to meet price targets, the latest plants are built with capacities around 1000 tons, or more.

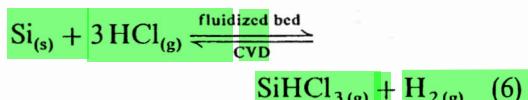
Two main routes are available for the production of semiconductor-grade silicon from MG-Si; either via chlorosilanes (principally trichlorosilane, SiHCl_3), or via silane (SiH_4). The former has been predominant since the late 1950s, but is now being challenged.

1.3.1 The Chlorosilane Route

This process, developed at Siemens (Bischoff, 1954), rapidly superceded the earlier SiCl_4/Zn method, which had been the principal silicon source until that time. There are three basic key steps in the process:

- Reaction between powdered MG-Si and hydrogen chloride gas in a fluidized bed reactor to form trichlorosilane (TCS).
- Fractional distillation of the TCS to provide it in an ultrapure, ppba, form.
- Reduction of the ultrapure TCS by hydrogen in a chemical vapor deposition (CVD) reaction to yield the desired product – semiconductor-grade polycrystalline silicon.

The fundamental, reversible, reaction is



The layout of a fluidized bed unit is shown in Fig. 1-4. However, again the actual reactions are more complex, and between this and the later stages of the process there is considerable recycling. The overall flow design of the plant, the efficient use of heat exchangers, and precise control of the recycling of intermediates are crucial factors in the operating costs.

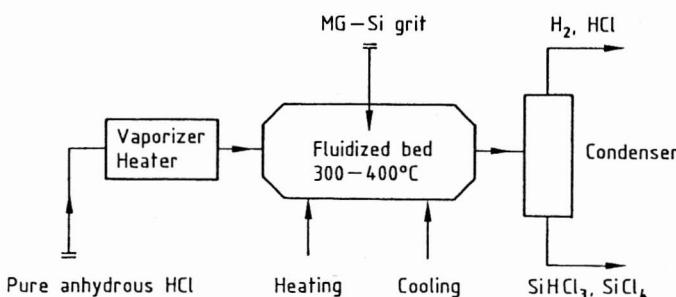


Figure 1-4. Layout of a fluidized bed reactor. The high degree of recycling in a chlorosilane plant is similar to that in a silane plant – see also Fig. 1-5.

The reaction between powdered silicon and anhydrous hydrogen chloride gas in the fluidized bed, held at 300 to 400°C is highly exothermic, producing a mixed output which contains about 90% of the wanted SiHCl_3 [Boiling point (BP): 31.8°C], about 10% SiCl_4 (BP: 57.6°C), and also a little dichlorosilane, SiH_2Cl_2 (BP: 8.5°C); together with hydrogen, unreacted HCl, and some volatile impurity metal chlorides. For this conversion high purity anhydrous HCl gas is essential, and a complex purification plant is needed to guarantee the <1 ppma level specified for this stage. Phosphorus trichloride (BP: 76°C), and boron trichloride, which is a gas at room temperature, are the two principal electrically active impurities carried over from the MG-Si; arsenic, as AsCl_3 (BP: 130°C) is also present to a lesser degree, together with small amounts other volatile metal chlorides, such as AlCl_3 ; but the fluidized bed stage does reduce the incoming impurity levels quite substantially.

At the next stage conventional high performance multiplate fractional distillation is employed to refine the TCS, separating it from other chlorosilanes and Si—H—C—Cl species present, and reducing the undesirable metals to ppba levels. The fractional distillation is backed up by the use of selective adsorption techniques to reach the very highest purity possible. This stage is pure chemical engineering, akin to that to be seen at any petrochemical refinery. Provided that they are kept completely dry, chlorosilanes, and also anhydrous hydrogen chloride, are chemically inactive in both liquid and gaseous form, and therefore can be moved and transported in conventional carbon-steel pipelines and tanks, but special valves and pumps are needed to maintain a totally leak-free environment. Thus the final ultrapure TCS is relatively easy to handle onwards to the next, silicon deposition, stage.

The quality of the semiconductor-grade TCS obtained at this stage is seen in Table 1-2, which shows the low levels of all impurities typically achieved. There has had to be much development of new analytical techniques in order to be able to quantify these impurities. Even using the highly sensitive method of ultraviolet spectroscopy, all metals are normally at a level below their respective detection limits. Special grades of even higher purity are available, for example, for epitaxial deposition. At this level often the only way to discriminate between two source materials is by their comparative performance under rigorously controlled conditions. When the data in Table 1-2 are compared with earlier published results, as for example those

Table 1-2. Impurities in semiconductor-grade trichlorosilane and polycrystalline silicon^d.

Impurity	Trichlorosilane	Polycrystalline silicon
SiHCl_3	99.9% min	—
SiH_2Cl_2	<0.2% max	—
SiCl_4	0.01% max	—
Resistivity	—	>1000 Ωcm
Carbon	<0.5 ppma ^a	<0.5 ppma ^b
Oxygen	not quoted	not quoted
Donors	<0.02 ppba	<0.10 ppba ^c
Acceptors	<0.03 ppba	<0.05 ppba ^c
Metals	below detection limits for UV spectroscopy, e.g.:	by neutron activation analysis:
Fe	<5 ppba	Fe <0.4 ppba
Ca	<1 ppba	Cr <0.05 ppba
Al	<1 ppba	Zr <0.2 ppba
Mg	<0.02 ppba	other metals individually <0.02 ppba

^a Hydrocarbons by chemical techniques; ^b by IR absorption; ^c by electronic measurements on the deposited polycrystal silicon.

^d Compiled from standard trichlorosilane and polysilicon specifications available from major suppliers, 1989.

given by Crossman and Baker (1977) (their studies of MG-Si have already been mentioned), the improvements of the last decade stand out, particularly with respect to the reductions in carbon, and residual donor levels.

The third stage of the process is silicon deposition, where the Siemens chemical vapor deposition technique is used worldwide by all the major producers. Again today's reactors are much larger, but the configuration is still essentially that first proposed in the mid 1950s. Two thin rods of ultrapure silicon, about 5 mm in diameter, and, today, up to 2–3 m in length are attached to heavy electrodes at their lower ends, while at their upper ends they are joined together with a shorter bridging piece of silicon thin rod – in an inverted “U” configuration. By passing an electric current through the silicon, the rods are heated to a temperature of about 1100°C in a trichlorosilane–hydrogen mixture, when reaction at the hot surface deposits ultrapure silicon. The process is continued for some 200–300 h until the rods reach a diameter of 150–200 mm. The preferred input material is TCS, rather than SiCl₄, which was used in the 1960s by some companies trying to evade the Siemens patent, because the reaction using TCS proceeds at a lower temperature and a faster rate. Again more complex than given by Reaction (6), the intermediate high temperature compound silicon dichloride, SiCl₂, plays an important role, both in the adsorption–deposition step at the surface and in the formation of byproduct SiCl₄. Since chlorosilane epitaxy uses the same process to deposit thin single crystal silicon layers, this reaction has been studied extensively, the work of Bloem, Gilling, and their co-workers over a number of years being notable (Bloem and Gilling, 1978; Bloem and Classen, 1980, 1983/1984).

In such a deposition process, any homogeneous nucleation, followed by particle growth in the gas phase, would lead to powder deposits downstream all over the system, which would be disastrous. The TCS process operating parameters are chosen to ensure a heterogeneous surface controlled reaction. Under these conditions the rate of deposition with time (kg/h) is directly proportional to the surface area, rising as the rod diameter increases. Clearly it is desirable to grow the rod to as large a diameter as possible, and hence the scaling up of these reactors has contributed very markedly to reducing process costs, as does recycling.

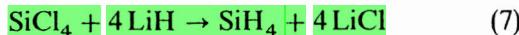
Only a fraction of the TCS fed through the reactor cells is converted to silicon. The outgoing gas contains unreacted SiHCl₃, SiCl₄, HCl, H₂, and other lesser constituents, all of which are separated, repurified as necessary, and fed back to the appropriate earlier stage, the HCl, for example, being returned right back to the first fluidized bed stage. Some of the SiCl₄ byproduct from the overall operation is not recycled but finds a market, having various uses, in epitaxy, for some silicon and silicone products, and more recently in the manufacture of optical fibers.

In summary, while there are operating complications, the Siemens process, in the three basic steps of dissolution, highly efficient distillation, and redeposition, converts MG-Si of ppma purity to ultrapure ppba semiconductor silicon, in a very cost-effective manner, and far superior to the earlier SiCl₄/Zn Dupont process (Lyon et al., 1949) that it displaced completely. By 1959 the laboratory growth of undoped float zone silicon crystals was being reported which had near intrinsic resistivity values, in excess of 100 000 Ω cm (Hoffman et al., 1959). Today's analytical data for typical electronic grade polysilicon, col-

lated from latest vendor specifications, as shown in Table 1-2, reflect the steadily continuing quality improvements, linked to those of its precursor trichlorosilane – the material which has underpinned the advances in the device field for over two decades.

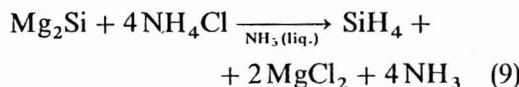
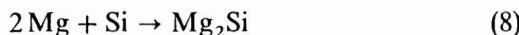
1.3.2 The Silane Route

Although research into the potential semiconductor uses of silane, SiH_4 , started early, silane became commercially important only in the later 1960s when its planar fabrication applications emerged, used in the chemical vapor deposition of silicon dioxide and silicon nitride dielectrics, and for polysilicon interconnect layers in devices. Much of this silane has been made using the lithium hydride, high temperature flux $\text{LiCl}(45\%)/\text{KCl}(55\%)$ electrolytic cell system (Sundermeyer, 1957), in which the key reaction at $390\text{--}430^\circ\text{C}$ is



The merchant market demand for silane for these purposes reached about 110 tons in 1986 and, following the continuing rapid silicon integrated circuit device expansion, has since more than doubled. However, here we are concerned with the use of silane as a route to bulk polysilicon.

For many years Komatsu Ltd., Japan, has operated a silane plant (Yusa et al., 1975; Taylor, 1987) which uses a process based on the reactions



Powdered magnesium and silicon are melted together to form the silicide, which is then reacted with ammonium chloride

in a liquid ammonia solvent at a temperature of 0°C . Since boron forms a stable $\text{BH}_3\cdot\text{NH}_3$ addition compound, it is reduced to extremely low levels, 0.01–0.02 ppba, at this stage. Phosphorus, as phosphine, PH_3 , is a much more difficult impurity to remove from the output silane gas. Therefore the next step is multiplate fractional distillation, under reduced pressure at very low temperature, below the boiling point of silane, -112°C . Since silane forms a spontaneously explosive mixture with air, the equipment is complicated and expensive, and the product could still contain too much phosphorus. Further purification by selective adsorption onto modified A-type zeolites, at a temperature of $20\text{--}30^\circ\text{C}$, removes PH_3 to extremely low levels, well below 0.01 ppba (Yusa et al., 1975). Work on improved modified zeolite adsorbants has continued, for example, Showa Denko KK (1984).

The final stage of deposition, carried out in equipment as described above for the Siemens process, but operated at a somewhat lower temperature, around 900°C , in order to avoid unwanted homogeneous reaction, is a simple thermal pyrolysis, requiring no other reactant, unused silane and hydrogen being the only byproducts for recovery. The polysilicon produced is of a very high quality, but it is more expensive than that made by the trichlorosilane route. The capacity of the Komatsu plant is quite small (150 tons in 1986), and the premium grade silicon product finds its use in the market for high-resistivity float-zoned, and neutron transmutation doped single crystals (see Sec. 1.4.2).

Within the U.S. Department of Energy funded solar energy program of the mid 1970s, various potential routes to silicon were explored or re-examined. Examples include the reduction of SiCl_4 by Na or Zn, and routes involving SiF_4 . These have

been well reviewed by Dietl et al. (1981). Out of this program one process has reached full maturity, in a plant that added significantly to the world's semiconductor silicon capacity. The Union Carbide, Moses Lake, Washington, plant produces silane as the precursor to deposition, but combines the merits of both TCS and silane in the intermediate stages (Taylor, 1987). Manufacture again starts from MG-Si fed into a fluidized bed reactor, but operated quite differently from that described in the Siemens process, in that the silicon is mixed with large amounts of recycled SiCl_4 , and hydrogen in the reaction



Operated at a high pressure of 500 psi (≈ 35 bar) and at a temperature of 500°C , this process is nearly thermoneutral, and yields a single-pass efficiency of 30–35%. It is quite unlike the highly exothermic Reaction (6) of the Siemens system. This stage is followed immediately by distillation to separate out and recycle the excess SiCl_4 .

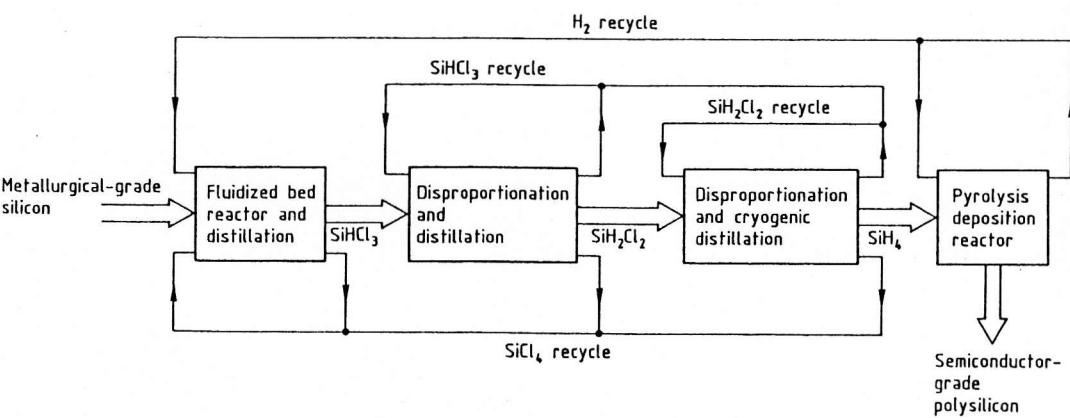
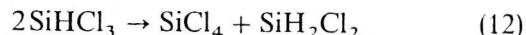
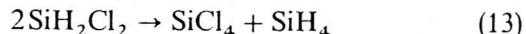


Figure 1-5. Flow diagram of the Union Carbide ultrahigh purity silane–silicon process. Noting the low single-pass stage efficiencies, control of the recycling is a key feature of this process, even more so than for the chlorosilane route. To keep this diagram simple the final recycle of unused SiH_4 has been omitted, as have the various condensers, etc., in each earlier feedback loop.

The next two stages, leading to silane, are successive catalytic disproportionation steps:



to dichlorosilane (DCS) and then



both are carried out at a pressure around 50 psi (≈ 3.5 bar) and at a temperature in the range of 60 – 80°C , using a tertiary amine ion-exchange resin catalyst; and each achieves a single-pass efficiency of around 10–12%. Again, as shown, the components are separated by distillation between the stages, the silane being finally purified by cryogenic fractionation. As revealed by the stage single-pass efficiency factors, there is a high degree of recycling throughout the process. The schematic flow diagram of the Union Carbide process set out in Fig. 1-5 shows the high level of feedback between stages.

The development of the purification methods required to achieve the highest purity silane has demanded improved and

new analytical techniques, such as deep level transient spectroscopy and photoluminescent spectroscopy, where sensitivities down to the parts per trillion (ppt) level are now routinely reported (Taylor, 1988).

The final pyrolysis at 900°C is as described above, with the byproduct hydrogen, and unused SiH₄ (not shown) being recycled.

The quality of the polycrystal silicon material available now from the latest silane technology is of somewhat higher purity than that from the established chlorosilane route, particularly with respect to ultralow levels of boron. The differences, however, are fairly marginal, and, even in the new, very large installations, present indications are that production costs are still a little higher than those associated with chlorosilane material.

1.3.3 The Polycrystal Silicon Market

While this chapter is concerned with technology, this is impacted by the market, and a few comments are appropriate. In order to compete, polysilicon plants have to be large. The initial installed capacity at Moses Lake in 1984 was 1200 t/a, which could be enlarged further to 3000 t/a. Another silane technology plant built in the U.S.A. by the Ethyl Corp., onstream in 1986 at 1000 t/a, was planned for 2000 t/a. The older Wacker, Germany, chlorosilane plant has been steadily modernized, and expanded to over 3000, and could reach 5000 t/a, while Tokuyama, in Japan is on a similar scale with room for expansion. The present western world installed capacity is difficult to state precisely but is in the range of 12 000 t/a, while the eastern bloc nations (around 2200 t/a in 1987) are now passing 3000 t/a.

It is important to recognize that this is the installed capacity, which makes no

allowance for maintenance downtime, breakdowns, and yields falling below design criteria; historically actual production output has been 70–80% of capacity. The cyclic and volatile nature of the component business has added to the competitive problems, as companies strive to hold, or increase their share; also Far East penetration into the silicon materials field increased substantially during the 1980s. Therefore it is not surprising that, notwithstanding increased energy costs and inflation, the pressure has meant that prices have stayed almost constant throughout this decade. Figure 1-6 shows the price/time changes for polysilicon, and their link to monosilicon prices, discussed later. As a result of the introduction of the new plants, the fraction of this silicon made from silane rather than chlorosilane has increased markedly.

Moving on, at the end of 1990 the Union Carbide, Moses Lake, plant was sold to

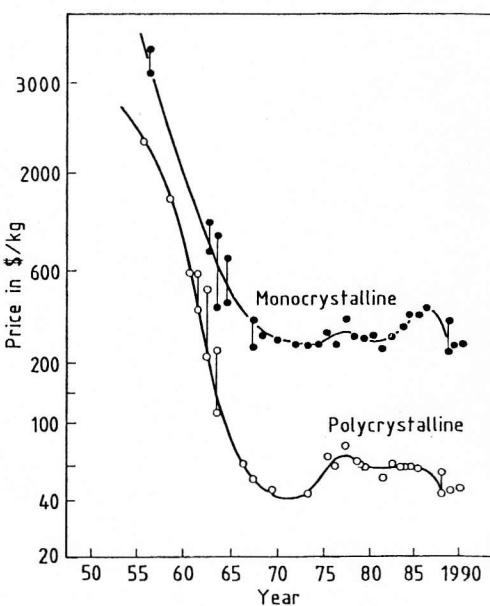


Figure 1-6. Prices of polycrystal and monocrystal silicon, 1956–1990.

Komatsu, Japan, a company with long experience of silane technology, as detailed above, which had the effect of further shifting market share towards the Far East. This operation has been renamed A. Si. M.: Advanced Silicon Materials. Following from the present and forecast device market trends, analysts are once again expressing some concern that there could be a silicon shortfall by the year 2000. Whether this develops or not, what is certain is that there will have to be further major capital investment in this sector, and it takes a considerable time to build a 1000 ton polysilicon plant. It remains to be seen who will lead, and where.

1.4 Single Crystal Silicon

The manufacture of monocrystalline silicon from the ultrapure polycrystal material is practiced on a much wider base than polycrystal production, where some of the metal refiners take their process no further, and do not enter the semiconductor part of the operation. It involves major vendors that only partially resource, or have no in-house polysilicon capacity, and also includes a few of the largest device companies, whose single crystal production is augmented by the purchase of either ingot, or later stage, for example, polished slice, products. Those who retain materials activities, are generally those who have been involved historically, and consume in excess of, say, 200–300 t/a; even so they always second-source from the large merchant vendors. Thus, over the years, crystal growth and slice fabrication have moved out from the device companies into the vendors with their specialized materials science and chemical engineering expertise.

Two techniques are used to manufacture monocrystalline silicon – free float zoning,

and Czochralski crystal pulling from a silica crucible. The material produced has different properties, and very different device applications. Today float zoning addresses a much smaller part of the market, and is discussed first.

1.4.1 Float-Zoned Silicon

Floating zone refining was introduced independently by several groups of workers in the early 1950s (Keck and Golay, 1953; Theurer, 1952, 1956), and very quickly became a widely used crystal growth technique.

In float zoning a cylindrical rod is held vertically and heated in argon by a radio frequency induction coil, to establish a molten droplet between the lower end of the rod and a single crystal seed rod mounted coaxially beneath, the two rods being rotated in opposing directions. The molten zone, retained in place between the polycrystal and seed rods only by surface tension, is then moved gradually upwards through the length of the polycrystal, converting it to monocrystal form.

The input material is a 2–3 m long cylindrical rod, obtained by limiting the final diameter at the deposition stage. The surface of an as-deposited rod has a coarse granular structure on a millimeter scale, which is removed by rotary diamond grinding to a tight diameter tolerance. The rod is then deep etched to a bright smooth finish, to ensure that all the damage and impurities introduced into the surface by the machining are removed, and also to reach the diameter required for the zoner. Finally the rod is washed in very high purity deionized water before zoning.

As already stated, the silicon deposition rate in the polycrystal reactor is proportional to diameter, and so limiting the process to making smaller diameter material

increases costs and reduces the reactor plant capacity. Only true straight rod sections can be used for zoning, and so, when the additional machining and preparation are taken into account, the input material is considerably more expensive than the simple broken poly lump suitable for melting in a crucible.

Float zoning is an inherently difficult technique because of the problems in controlling such a free molten zone. Compared to Czochralski pulling, where melt convection can be influenced by the heater shape and its position with respect to the melt, and by baffle design, and the crucible and crystal rotations can be chosen at will, in the zoner there are far fewer degrees of freedom to choose from to adjust its performance. A great deal of work has centered on induction coil design, and the use of auxiliary coils, to set the desired temperature gradients.

Hence developments to reach larger diameters lagged years behind the availability of equivalent Czochralski material. It was found almost impossible to maintain a large diameter molten zone whose volume was such that surface tension alone could hold it in place without the liquid falling out. The problem was eventually solved by

a new approach, known as the "needle-eye" technique (Keller, 1959), see Fig. 1-7. During the rod preparation machining stage, described above, a modification was introduced whereby, after the rod had been ground to a parallel cylinder, its bottom end was further ground to a tapered cone, as matching section to the seed. The RF induction coil was made at a smaller diameter than the machined rod, so that as the zone was established and passed up through the polysilicon rod, with the coil now just inside the rod diameter, the smaller molten zone volume involved could then be successfully moved through the length of the rod without prematurely falling out. Establishing this as a reliable production procedure took much effort, and the design of highly specialized zoners.

By this crucible-free technique standard crystal product, zero dislocation density silicon up to 150 mm diameter, is made at a purity close to that of the input polysilicon. Resistivities are available up to 5000 $\Omega \text{ cm}$, and in limited amounts even higher, well beyond that attainable by Czochralski growth.

The control of doping to specific resistivity bands in float zoned (FZ) silicon is approached in three ways. Firstly, by gas

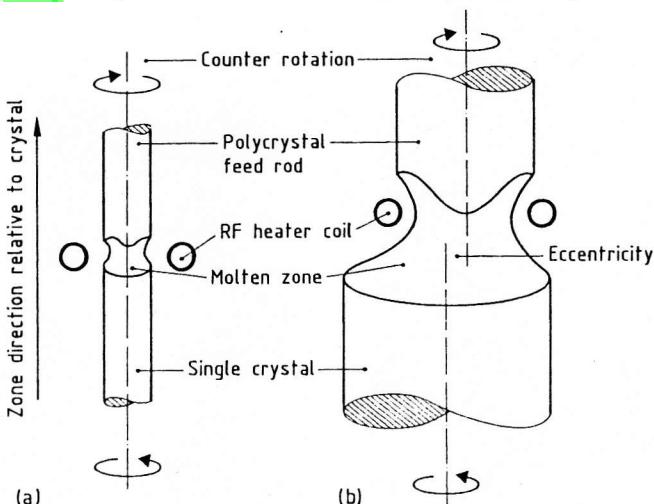


Figure 1-7. Floating zone growth configurations. (a) Keck, small diameter; (b) Needle-eye, large diameter.

phase doping during crystal growth, adding diluted phosphine, PH_3 , or diborane, B_2H_6 , to the argon gas flow through the zoner, a wide range of both n- and p-type specifications can be made. The higher resistivities are more difficult to meet, and the tolerances quoted get wider. The equilibrium segregation coefficients for boron and phosphorus in silicon are 0.8 and 0.35, respectively, and so doping uniformity, both axially and radially, is somewhat easier to achieve with boron.

The second doping method used is only available to the polysilicon producers. The thin rods used to construct the inverted "U" structures, for the reactors in which the silicon deposition occurs, are made by fast pulling from a pure silicon melt. By adding phosphorus or boron to the melt, doped thin rods can be made to various specifications and stocked, to be used later when the deposition rods are planned for float zoning. Then as the molten zone is passed through the polysilicon rod the dopant in the core is released. Very precise resistivity control can be achieved by this method, with high run-to-run reproducibility. The third method, only possible for phosphorus, n-type, material, is neutron transmutation doping (NTD) which has become very important for power applications.

1.4.2 Neutron Transmutation Doped Silicon

By float zoning undoped material under very clean conditions crystal can be produced in which the resistivity is of the order of $5000 \Omega \text{ cm}$ or greater, and with a very low residual phosphorus and boron. If this ingot is placed into a nuclear reactor, transmutation doping generates phosphorus in an extremely uniform distribution, avoiding the growth striation phenomena common to both the float zone and

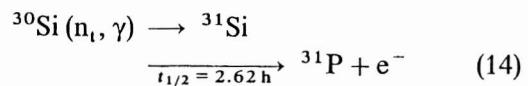
Czochralski growth methods, to be discussed later. This method is particularly suitable for making the high resistivity silicon required for power device applications, where the other doping techniques cannot match NTD material in its ability to meet very close tolerances.

Normal elemental silicon consists of three stable isotopes with abundancies as:

^{28}Si	92.21%
^{29}Si	4.70%
^{30}Si	3.09%

The possibility of doping silicon by transmuting the ^{30}Si isotope into ^{31}P was first recognized by Lark-Horowitz (1951). Later, Tanenbaum and Mills (1961) made detailed experiments to verify that the scheme was potentially useful, but this work lay on one side and only re-emerged when Herrmann and Mücke (1973) published their power device study. Since then the major developments have taken place, leading to a series of international conferences and many papers on this sole topic.

The principal nuclear reaction upon which the whole process depends is



The stable ^{30}Si isotope captures a thermal neutron to form ^{31}Si with the emission of γ -rays. For this isotope of natural abundance 3.09%, the capture cross-section for a thermal neutron is $0.11 \times 10^{-20} \text{ cm}^2$, or 0.11 barn. In its turn ^{31}Si is unstable and decays with a half life of 2.62 h to the stable phosphorus isotope ^{31}P with the emission of an electron of energy 1.47 MeV.

During neutron irradiation other nuclear reactions occur, some of which must be taken into account:



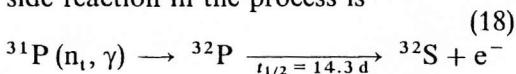
and



The thermal neutron capture cross-sections for these are 0.08 and 0.28 barn, respectively, and the only real effect on the process arises from ingot heating by the emitted γ -rays. However, two other reactions occur, which are much more important:

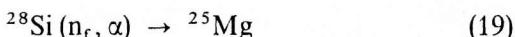


In this reaction a total energy release of 2.5 MeV is associated with the α and lithium particles, which leads to considerable short range lattice damage. Boron has a very high thermal neutron capture cross-section of >755 barn, but, provided the boron concentration in the silicon is kept low, the effect is small. The most serious side reaction in the process is



Since the capture cross-section here is only 0.2 barn the amount of sulfur produced is minute in doping terms, but the long half life for the decay of ^{32}P can impose restrictions on the handling of low resistivity NTD silicon.

All the reactions so far have referred to thermal neutrons, that is, neutrons which have already been scattered by sufficient collisions within the pile that their energy has reached thermal equilibrium with the medium before intersecting the silicon. Such neutrons, at room temperature, have an average energy of only 0.025 eV and a velocity of 2200 m/s. However, fast neutrons in the pile, with energies in excess of 1 MeV, also reach the silicon. These give rise to much of the lattice damage generated during transmutation doping, and are also responsible for reactions of the type



producing a high energy α -particle. Even without reaction, the head-on collision of a 1 MeV neutron with a silicon atom will knock out about 200 silicon atoms from their lattice sites. Thus the slow-to-fast neutron ratio in the nuclear reactor is critical, and it is for this reason that heavy water reactors, with slow-to-fast ratios around 1000:1 (a much higher ratio than available in light water and other reactors), have proved so successful for NTD processing.

The subject of neutron irradiation damage has proved to be a matter of great complexity, beyond the scope of this chapter. Much lattice disarray is introduced immediately after irradiation high resistivity values are found, and at this point most of the phosphorus formed is interstitial. Therefore the post-irradiation annealing process is crucial and has been studied in depth. The resistivity values, expected from the total reactor neutron flux, are fully realized after about 1 h at 600 °C; however defect studies indicate the need for a higher temperature, and poor minority carrier lifetime has been a problem (Meese, 1978). The producer's postanneal processes which at the introduction of NTD products were at moderate temperature, sometimes relying on the high temperature semiconductor fabrication to complete the anneal, subsequently moved to higher temperatures and more complex time-temperature schedules. The exact details of vendor anneal schedules are generally sensitive proprietary information.

1.4.3 Carbon and Nitrogen in Float-Zoned Silicon

For many years it was thought that an residual carbon in silicon was of little importance; it is an isoelectronic group 4 element, it occupies substitutional sites in th

lattice, and silicon carbide is an insulator. The first indication of device linked effects came when Akiyama et al. (1973) reported a correlation between high carbon concentrations, reduced breakdown voltage, and increased reverse current leakage in rectifier diodes. Because this work used silicon with a very high carbon level (between 1 and 2×10^{17} atoms/cm³) close to the solid solubility limit, and only appeared as a short communication, its significance was generally overlooked at the time. Carbon was not listed in most purchasing specifications.

The common requirement of trichlorosilane for organosilicon and semiconductor use has already been noted in Sec. 1.2, and many of the world's polysilicon plants are cosited with, or close to a silicones plant. Thus, when in mid 1975, accidentally and undetected, a quantity of high-carbon float zone silicon reached device lines, serious yield problems were met in rectifier, thyristor, and power transistor manufacture. Subsequent research showed that, while not affecting the breakdown voltage, lower levels of carbon still degrade the reverse leakage, as is shown in the plot of Fig. 1-8. Recognizing a severe problem, polysilicon producers made major plant overhauls to remove carbonaceous sources from pumps, valves, glands, etc., set new low carbon operating standards, and instituted strict test procedures – in single crystal the substitutional carbon has an infrared absorption at 16.6 μm (605 cm⁻¹) measured by differential (double-beam) spectrophotometry, ASTM Standard F123. Today carbon levels are rigorously controlled by all silane, TCS, and polysilicon producers, to ensure final silicon levels below around 2×10^{16} atoms/cm³.

The role of nitrogen in silicon is quite different. Unlike other group 5 elements, such as P, or As, nitrogen does not behave

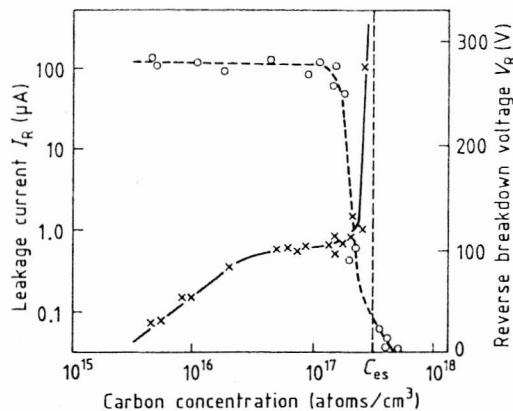


Figure 1-8. Rectifier diode failure as a function of carbon concentration: o breakdown voltage; x reverse leakage current. Note that, even when the breakdown voltage has been restored, leakage effects persist to much lower carbon levels. The dashed vertical line marks C_{es} , the carbon solid solubility saturation value.

as a donor impurity. An electronic center deep in the band gap has been reported by Tokumaru et al. (1982), but in general nitrogen does not appear to be electrically active in melt doped silicon. Another distinguishing feature is its low solid solubility: 3×10^{15} atoms/cm³ at the melting point of silicon (Yatsurugi et al., 1973). This is much lower than for other light elements, such as carbon, $N = 3.5 \times 10^{17}$ cm⁻³, or oxygen, 1.7×10^{18} atoms/cm³. In float zoned silicon the equilibrium solid solubility is often exceeded, the excess concentration being proportional to the zone velocity, where typically the values met may be up to 5×10^{15} atoms/cm³ (Yatsurugi et al., 1973).

The low electrical activity of nitrogen in silicon is useful because it has a major attribute, in that nitrogen doping at low concentrations, limited by its solid solubility, effectively inhibits dislocation generation and propagation, as first reported by Abe et al. (1981). Second-phase hardening is a well known metallurgical phenomenon,

softer bed pad, uses a gentler slurry of near neutral pH, and operates at lower pressure and around room temperature. After a somewhat shorter second polishing period, around 10 min, only a small extra amount of silicon is removed (1–2 µm), but this delivers the final defect-free surface. In the third stage on the same machine, pressure is relieved so that the slices are almost free floating and the second stage slurry is replaced by a "haze suppressant", a complex organo-silicate, diluted in a glycol–water medium, which contributes to hydrophobicity, and clean liquid drainage from the slice surfaces, leaving no marks as they are removed at the end of the process. The fine tuning of the polishing process, to reach the very high standards demanded, involves considerably more variables and detail than given in this outline.

This chemical–mechanical polishing combines both the mechanical process necessary to achieve flatness to micrometer tolerances, and a chemical process to ensure a final crystallographic perfection with undisrupted atomic structure free of all damage.

The final cleaning and packaging are similar to polishing in their emphasis on the materials used and the practice. Again in outline, the first step after polishing is to remove any traces of polishing slurries, waxes, and organics by means of degreasants, and/or organic destruct routes, using, for example, hot nitric acid, followed by washing in deionized (DI) water. The emphasis in the final clean is on the removal of every trace of surface contaminants, metals, etc., and particles of whatever form. Kern and Puotinen (1970) first described the, subsequently widely used, RCA cleaning process. All the chemicals employed have to be of electronic grade, and membrane filtered to very low submicrometer particle counts, as is the DI water. Much

depends on the performance of soft-p slice scrubbers, and single slice spinners operating in a cassette-to-cassette mode while robotics are finding increasing use. Class 10, and lower, clean surroundings, routes where hydrophobic slices are required, an HF dip is inserted, but when the, more normal, hydrophilic surface is wanted, the ammonia–peroxide treatment has become common, before the final DI water scrub–rinse–spin-dry schedule. Cassette packaging in multicomponent ultraclean boxes completes the wafer processing.

The impact of ULSI requirements on these stages has been as demanding as elsewhere. At the polishing stage much attention has been paid to the machine bed construction to ensure that slices do not have their flatness degraded during the process. Then the process itself has been modified essentially with the aim of reducing the temperatures and pressures used at the stages to achieve gentler processing. Particles and slurries have been improved; now instead of feeding slurry onto the machine in small quantities, flooding from a recirculating tank is used to ensure that adequate compound reaches the central area of larger diameter slices. The comments on cleanliness given above are reinforced, and even more attention has been given to assay methods to test for impurities in chemicals to ultralow levels of detection in order to meet the very low surface metal targets of below 10^9 atoms/cm².

1.7 Oxygen in Czochralski Silicon

1.7.1 The Behavior of Oxygen in Silicon

The dissolution, evaporation, and flow, dynamic balance, controlling the incorporation of oxygen into silicon is described in Secs. 1.5.3 and 1.5.4, above.

softer bed pad, uses a gentler slurry of near neutral pH, and operates at lower pressure and around room temperature. After a somewhat shorter second polishing period, around 10 min, only a small extra amount of silicon is removed (1–2 μm), but this delivers the final defect-free surface. In the third stage on the same machine, pressure is relieved so that the slices are almost free floating and the second stage slurry is replaced by a “haze suppressant”, a complex organo-silicate, diluted in a glycol–water medium, which contributes to hydrophobicity, and clean liquid drainage from the slice surfaces, leaving no marks as they are removed at the end of the process. The fine tuning of the polishing process, to reach the very high standards demanded, involves considerably more variables and detail than given in this outline.

This chemical–mechanical polishing combines both the mechanical process necessary to achieve flatness to micrometer tolerances, and a chemical process to ensure a final crystallographic perfection with undisrupted atomic structure free of all damage.

The final cleaning and packaging are similar to polishing in their emphasis on the materials used and the practice. Again in outline, the first step after polishing is to remove any traces of polishing slurries, waxes, and organics by means of degreasants, and/or organic destruct routes, using, for example, hot nitric acid, followed by washing in deionized (DI) water. The emphasis in the final clean is on the removal of every trace of surface contaminants, metals, etc., and particles of whatever form. Kern and Puotinen (1970) first described the, subsequently widely used, RCA cleaning process. All the chemicals employed have to be of electronic grade, and membrane filtered to very low submicrometer particle counts, as is the DI water. Much

depends on the performance of soft-pad slice scrubbers, and single slice spinners, operating in a cassette-to-cassette mode, while robotics are finding increasing use, in Class 10, and lower, clean surroundings. In routes where hydrophobic slices are required, an HF dip is inserted, but where the, more normal, hydrophilic surface is wanted, the ammonia–peroxide treatment has become common, before the final DI water scrub–rinse–spin-dry schedule. Cassette packaging in multicomponent ultraclean boxes completes the wafer processing.

The impact of ULSI requirements on these stages has been as demanding as elsewhere. At the polishing stage much attention has been paid to the machine bed construction to ensure that slices do not have their flatness degraded during the process. Then the process itself has been modified, essentially with the aim of reducing the temperatures and pressures used at all stages to achieve gentler processing. Pads and slurries have been improved; now instead of feeding slurry onto the machine in small quantities, flooding from a recirculating tank is used to ensure that adequate compound reaches the central area of large diameter slices. The comments on cleanliness given above are reinforced, and even more attention has been given to assay methods to test for impurities in chemicals to ultralow levels of detection in order to meet the very low surface metal ULSI targets of below 10^9 atoms/cm².

1.7 Oxygen in Czochralski Silicon

1.7.1 The Behavior of Oxygen in Silicon

The dissolution, evaporation, and fluid flow, dynamic balance, controlling the incorporation of oxygen into silicon described in Secs. 1.5.3 and 1.5.4, above ig-

nored its distribution coefficient. Under such growth conditions this is not easy to determine, and past estimates have varied widely, from $k_0 = 0.5$ to 1.4. Current views that it is less than 1.0 are supported by a recent study (Jackson, 1990), which arrived at 0.7.

While carbon enters the lattice as a substitutional impurity occupying a silicon site, oxygen does not, but instead enters as a bound interstitial impurity, bonding between two adjacent silicon atoms, in a structure which permits more complex vibrational modes (Newman, 1973). The broad 9 μm infrared absorption band, seen at room temperature, arises from a number of vibrational modes of similar energies. The concentration of bound-interstitial oxygen in silicon is measured by the 9 μm

absorption (ASTM Standard F-121), and if any oxygen is precipitated within the crystal, by heating in the range 1050–600 °C, the absorption decreases. Reheating at a high temperature, >1100 °C, disperses the precipitates and restores the absorption. However, if the temperature is held at around 450 °C, any unprecipitated interstitial oxygen present forms “thermal donors”, which cause major resistivity changes in the crystal. This thermal behavior pattern was first established by Kaiser et al. (1956) and then expanded (Kaiser, 1957; Kaiser et al., 1958).

Long Czochralski crystals, which are grown over a period of many hours, slowly withdrawing into a cooler chamber, experience a different thermal history between the seed and tail ends, depicted in Fig. 1-25.

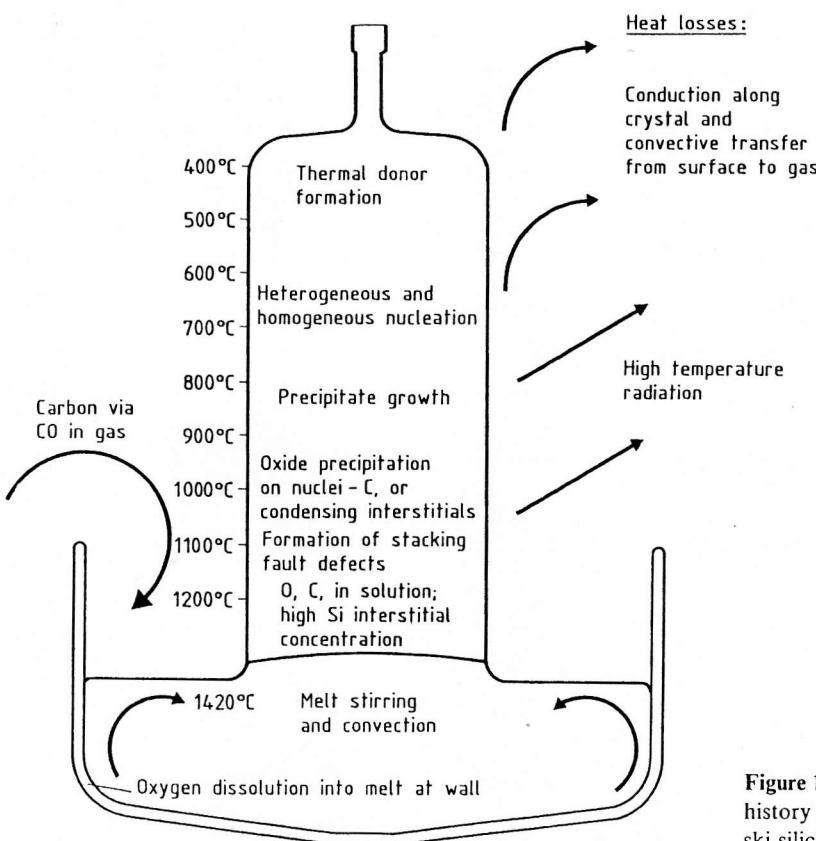


Figure 1-25. The variable thermal history of an as-grown Czochralski silicon crystal.

This has direct device effects; for example, a difference in leakage current yield losses of 16 k DRAM chips has been linked to the slice cutting location, reported by Steinbeck (1980a, b).

Other studies have shown device failures associated with crystal defects, either present at the start of the fabrication process or formed during it, and also linked to the oxygen status. From defect etching studies, many observers noted that where a high density of surface defect features (e.g., oxidation induced stacking faults, seen after the first furnace step) was found on one side of a slice, the opposite face had a very low density. In one direction, this was soon linked to residual damage remaining after slice polishing. Similar work demonstrated the relation between oxidation-induced stacking faults, the slice heat treatment temperature, and oxygen precipitation (Matsushita, 1982). Much device engineering research has explored the generation and suppression of oxidation-induced stacking faults during fabrication (Stimmel, 1986), but to work bulk silicon it is necessary to understand the basic precipitation mechanism.

1.7.2 The Precipitation of Oxygen in Silicon

The maximum amount of oxygen grown into a crystal at its top (seed) end, is such that, as it cools below about 1200°C, the concentration exceeds the solid solubility limit (Carlberg, 1986), and precipitation of a second, silicon oxide, phase may be expected at any lower temperature. Solid state reactions proceed at far slower rates than in liquids or gases, so as the crystal cools, over many hours, the amount of the supersaturated oxygen which actually precipitates is so small as to be almost undetectable. It is only on heating slices cut

from the crystal, for many hours at elevated temperatures, that substantial precipitation occurs, growing from tiny initial nuclei.

Research into bulk crystallization from liquids, to produce, for example, fertilizers and salts, has contributed much to nucleation concepts, and in particular the particle of critical radius r_c . In a supersaturated liquid, or solid, at the outset tiny atomic clusters form and redisperse in a highly dynamic situation, but some merge and grow, until, reaching a certain critical radius, they become stable, and from then on will not redissolve. In such a process there is an initial induction period during which sufficient nuclei reach r_c , then faster precipitation, which dies away as the equilibrium solubility is approached.

Many systems exhibit this behavior, including the solid state precipitation of oxygen in silicon, where at 750°C, the process has still not reached equilibrium after over 1000 h – solid state reactions are very slow. In this approach it should be expected that the nuclei formed by other impurities present will affect the initial nucleation induction step. Thus in the silicon case, the distribution of oxide precipitates across a slice after heat treatment closely maps the grown in carbon distribution shown in Fig. 1-11 (Wilkes, 1983), and also influences the actual precipitation kinetics (Kishino et al., 1979; Craven, 1981; Shimura et al., 1985; Barraclough and Wilkes, 1986).

After nucleation, the main precipitation process reduces the bound interstitial oxygen concentration, developing different numbers and sizes of particles according to the temperature employed. A simple model can be used to predict the qualitative behavior correctly, and provides a basis for understanding the theoretical approach.

Suppose two similar, adjacent, samples of the same impurity content, and with the