MKS INSTRUMENTS HANDBOOK

Semiconductor Devices and Process Technology

by the Office of the CTO

9/2017
© 2017 MKS Instruments, Inc.
All rights reserved.

Limit of Liability/Disclaimer of Warranty: The purpose of this book is to educate and should be used only as a general guide. The authors and publisher are providing this book and its contents on an “as is” basis and make no representations or warranties of any kind. Neither the authors nor the publisher shall be liable for any costs, damages and expenses whatsoever (including, without limitation, incidental, special and consequential damages) resulting from any use of or reliance on the information contained herein, whether based on warranty, contract, tort or any other legal theory, and whether or not they have been advised of the possibility of such damages.

This document shall not be copied, reproduced, distributed or published, in whole or in part, without the prior written consent of MKS Instruments, Inc.

AIRGARD®, ASTRON®, Baratron®, Conductron®, Convectron®, DynamYX®, Granville-Phillips®, I-Mag®, LIQUOZON®, Micro-Ion®, Mini-Convectron®, O_3MEGA®, Paragon®, Precisive®, R*evolution®, Resist-Torr®, SEMOZON®, SmartMatch®, SmartPower®, Spectra-Physics®, Spirit®, Stabil-Ion®, SurePower®, TOOLweb®, VQM® are registered trademarks, and Cirrus™, DELTA™, DualTrans™, Effluent Management Subsystem™, elite™, HybrYX™, KEINOS™, LoPro™, mksinst™, MicroPirani™, MultiGas™, Newport™, Process Eye™, Process Sense™, SenseLink™, Virtual Wall™, Vision 2000-C™, Vision 2000-E™, Vision 2000-P™, V-lens™ are trademarks of MKS Instruments, Inc. or a subsidiary of MKS Instruments, Inc.

DeviceNet® is a registered trademark of ODVA, Inc.
EtherCAT® is a registered trademark of Beckhoff Automation.
Inconel® is a registered trademark of Huntington Alloys Corporation.
Microsoft® and Windows® are registered trademarks of Microsoft Corporation.
PI PROFIBUS · PROFINET® is a registered trademark of PROFIBUS Nutzerorganisation e.V..

All other trademarks cited herein are the property of their respective owners.

MKS Instruments Handbook

Semiconductor Devices and Process Technology

by the Office of the CTO

This page intentionally left blank.

MKS Instruments Handbook

Semiconductor Devices and Process Technology

by the Office of the CTO

Office of the CTO

Josh Ding, Ph.D., Chief Technology Officer, Flow Solutions

Atul Gupta, Ph.D., Head of Technology, Plasma & Reactive Gas Solutions

Jim Kafka, Ph.D., Chief Technology Officer, Lasers

John TC Lee, Ph.D., Chief Operating Officer & Senior Vice President

Aaron Radomski, Chief Technology Officer, Power Solutions

Erik Saarmaa, J.D., Chief IP Counsel, Intellectual Property

Eric Snyder, Vice President, Quality & Reliability

Philip Sullivan, Ph.D., Fellow & Chief Technology Officer,
Pressure & Vacuum Measurement Solutions

Robin Swain, Ph.D., General Manager, Integrated Solutions Business

Eric Taranto, Vice President Analytical & Control Solutions & Integrated Solutions Business

Ruben Zadoyan, Ph.D., Chief Technology Officer, Newport Products

Former CTO Members

Gerardo Brucker, Ph.D., Chief Scientist, Pressure & Vacuum Measurement Solutions

Steve Walther, Ph.D., former Chief Technology Officer, Plasma & Reactive Gas Solutions

Kevin Wenzel, Ph.D., former Chief Technology Officer, Plasma & Reactive Gas Solutions

Peter Zemek, Ph.D., former Chief Technology Officer, Process & Environmental Analysis Solutions

Contributing Editors

Shannon Whidden, TechLink Writers Group

Tom K. Whidden, Ph.D., TechLink Writers Group

Special Thanks

Anne Devine, Graphic Designer

Jane Lawlor, Senior Marketing Specialist, Creative Design

MaryAnn Naddy, Senior Director, Corporate Marketing (retired)

Foreword

Semiconductor devices have become essential and ubiquitous parts of our everyday lives, and this has been enabled by an industry that has pursued a relentless reduction in cost per function for over 50 years. Today's devices contain billions of transistors interconnected by kilometers of "wires," yet all fitting in a fingernail-sized integrated circuit. These devices are enabled by manufacturing processes that continue to increase in complexity and require unprecedented innovation in physical structure, new materials, process control, and throughput.

Some of the most important unit processes required to sustain this pace of innovation include Etching, Deposition, Lithography, Metrology, and Wet Cleaning. In turn, the improvements in these unit processes are dependent on critical subsystem technologies and instruments that measure, control and sustain pristine vacuum systems, deliver critical gases, reactive species, and power to these systems, and provide state-of-the-art laser, optics and photonics capabilities to define and measure sub-nanometer features. MKS Instruments uniquely serves this market with the broadest portfolio of solutions in the industry.

The Semiconductor Devices and Process Technology handbook presents the fundamental device physics, materials, and fabrication processes used to manufacture semiconductors, as well as the technologies, instruments, and equipment that are used to monitor, control, and automate the fabrication processes. This handbook was initially conceived as training material for our employees. As it has evolved, we want to share it with you, our customers, suppliers, and partners in the semiconductor industry, with the hope that you will find it informative and of some value to your work.

This handbook came about thanks to what we learned from working alongside you for over half a century. In that spirit, we welcome your feedback and input to future editions. There's no denying that rapidly evolving technology will continue to shape and transform this industry in the years to come. MKS Instruments looks forward to facing the opportunities and challenges this will present—together.

Gerald G. Colella
Chief Executive Officer & President

About MKS Instruments

MKS Instruments, Inc. is a global provider of instruments, subsystems and process control solutions that measure, control, power, monitor, and analyze critical parameters of advanced manufacturing processes to improve process performance and productivity. Our products are derived from our core competencies in pressure measurement and control, flow measurement and control, gas and vapor delivery, gas composition analysis, residual gas analysis, leak detection, control technology, ozone generation and delivery, RF & DC power, reactive gas generation, vacuum technology, lasers, photonics, sub-micron positioning, vibration isolation, and optics. Our primary served markets include semiconductor capital equipment, general industrial, life sciences, and research. Additional information can be found at www.mksinst.com.

Table of Contents

Section A - Semiconductor Device Physics, Materials and Fabrication 1

- I. Semiconductor Physics and Basic Device Structures 3
  - A. Semiconductor Physics 3
    - 1. Electrical Characteristics of Solids 3
    - 2. Electrical Conduction in Semiconductors 4
  - B. Basic Device Structures 9
    - 1. P-N Junction 9
    - 2. Diodes 10
    - 3. Bipolar Junction Transistor 11
    - 4. MOSFET 12
    - 5. FinFETs 14
    - 6. Other Common Semiconductor Devices 15
- II. Semiconductor Materials 19
  - A. Bulk Crystalline Silicon and Compound Semiconductors 19
    - 1. Electronic Grade Polycrystalline Silicon (Polysilicon) 19
    - 2. Single Crystal Silicon Wafer Manufacture 21
    - 3. Compound Semiconductor Wafer Production 23
  - B. Thin Films for Device Fabrication 25
    - 1. Polycrystalline Silicon Thin Films 25
      - a. Overview 25
      - b. Material and Electrical Properties 25
      - c. Deposition of Polycrystalline Silicon Thin Films 27
    - 2. Epitaxial Thin Films 31
      - a. Silicon Based Epitaxial Thin Films 31
      - b. Compound Semiconductor Epitaxial Thin Films 37
    - 3. Dielectric Thin Films 41
      - a. Silicon Dioxide 41
      - b. Silicon Nitride 48
      - c. High-k and Low-k Dielectrics 49
      - d. Other Dielectrics 50

4. Metal Thin Films for Contacts and Interconnects 52
a. Metal Materials Requirements 52
b. Aluminum 54
c. Tungsten 56
d. Copper 58
e. Barrier Layers 60
III. Example Fabrication Process: CMOS Devices 63
A. Wafer Surface Cleaning 63
1. Contaminant Types and Solution Cleaning Methods 65
a. Particles 65
b. Metals 65
c. Chemical Contamination 66
B. CMOS Device Structure 67
C. CMOS Process Steps 69
1. Pad Oxide (Thermal Oxidation) 69
2. Silicon Nitride 69
3. Shallow Trench Isolation (STI) Process 70
4. N- and P-Well Process 71
5. Gate Process 72
6. Contact Process 72

Section B - Process Equipment, Technology, and MKS Instrument's Product Applications . 75
I. Vacuum Technology 77
A. Basic Concepts 77
B. Creating a Vacuum 79
C. Measuring Vacuum 85
D. Vacuum/Pressure Control 98
E. Residual Gas Analysis 104
II. Substrate Surface Cleaning 109
A. Piranha, SC1, SC2, RCA and DHF 109
B. “Dry” Substrate Cleaning 114
III. Ion Implantation 117

IV. Thermal Processing 119
  A. Oxidation 119
  B. Diffusion 123
  C. Annealing 124
  D. Rapid Thermal Processing 124
V. Thin Film Deposition 127
  A. Chemical Vapor Deposition and Thin Film Formation 128
    1. Low Pressure Chemical Vapor Deposition (LPCVD) Systems 132
    2. Plasma Enhanced CVD (PECVD) Systems 135
    3. Sub-Atmospheric Pressure CVD Systems 136
  B. Atomic Layer Deposition (ALD) 137
  C. Physical Vapor Deposition (PVD) Systems 138
VI. Photolithography 143
  A. Basic Principles and Procedures 143
    1. Deep UV (193 nm) Technology 146
    2. Extreme Ultraviolet (EUV) Lithography at 13.5 nm 152
VII. Etch Technology 157
  A. Basic Processes 159
  B. Isotropic Radical Etching 161
  C. Reactive Ion Etching 162
  D. Physical Sputtering and Ion Milling 164
VIII. Chemical Mechanical Planarization 169
  A. Chemical Mechanical Polishing/Planarization 169
  B. CMP in Shallow Trench Isolation Processing 171
  C. CMP in Damascene Processing 171
IX. Wafer and Reticle Inspection 173
  A. Basic Principles 173
    1. Unpatterned Wafer Inspection Systems 173
    2. Patterned Wafers 174
    3. Reticle Inspection 175
  B. Advanced Wafer and Reticle Inspection Tools 175
    1. Sub-100 nm Unpatterned Wafer Inspection 175
    2. Sub-100 nm Patterned Wafer Inspection 176
    3. Sub-100 nm Reticle Inspection 176

mks

X. Automation of Process Equipment 181

XI. Semiconductor Fab Utilities 187

A. Ultrapure Water 187

B. Gas Storage and Delivery 191

C. Process Exhaust Gas Traps and Scrubbers 194

D. Cleanroom Air Supply System 200

Bibliography 203

Appendix A 217

Electrical Conduction in Solids 217

Section A

Semiconductor Device Physics, Materials and Fabrication

This page intentionally left blank.

I. Semiconductor Physics and Basic Device Structures

A. Semiconductor Physics

1. Electrical Characteristics of Solids

When it comes to electrical characteristics, solid matter falls into one of three categories (Table 1): it is either a conductor, an insulator, or a semiconductor, depending on whether or how well it conducts electricity. Conductors have many "mobile charge carriers," e.g., negative electrons. These are not strongly bound to individual atoms and they can be easily made to flow through the conductor by the application of an electrical potential [1] [2]. Examples of solid conductors include all of the familiar metals, along with semi-metals such as graphite, α-tin, bismuth, antimony and arsenic. Insulators, on the other hand, have their electrons tightly bound to the atoms; it takes a great deal of energy (i.e., a very high electrical potential) to remove them from the atoms so that they can flow through the material. Examples of insulators include glass, ceramics, plastics and wood.

Conductors	• Many free electrons, which can easily be made to flow through material • Examples: all metals, semi-metals such as carbon-graphite, antimony and arsenic
Insulators	• Very few free electrons • Examples: plastic, glass and wood
Semiconductors	• Between the extremes of good conductors and good insulators • Crystalline materials that are insulators when pure, but will conduct when impurity is added and/or in response to light, heat, voltage, etc. • Examples: elements such as silicon Si, germanium Ge, selenium Se; compounds such as gallium arsenide GaAs and indium antimonide InSb

Table 1. Conductors, insulators and semiconductors.

Semiconductors lie between these extremes of electrical conductivity. They are high purity crystalline materials that conduct electricity under relatively low electrical potential (when compared to insulators), without the high electrical conductivity of metals. When certain impurities are added to the crystal matrix of a semiconducting material, they contribute mobile charge carriers (either electrons or holes) that can dramatically increase the electrical conductivity of the semiconductor under the proper conditions. Semiconducting materials include elements such as silicon and germanium, along with compound (i.e. multi-element) semiconductors such as gallium arsenide (GaAs), gallium nitride (GaN), silicon carbide (SiC), indium antimonide (InSb), and indium phosphide (InP). Figure 1 shows a Periodic Table of the Elements. The elemental semiconducting materials are all from Group IVA in this table, while compound semiconducting materials are formed through the combination of elements from Groups IIIA and VA (commonly referred to as III-V semiconductors) or Groups IIB and VIA (II-VI semiconductors).

IA	IIA	IIIB	IVB	VB	VIB	VIIB	VIIIB			IB	IIB	IIIA	IVA	VA	VIA	VIIA	VIIIA
Hydrogen 1 H 1.0079																	Helium 2 He 4.0026
Lithium 3 Li 6.941	Beryllium 4 Be 9.0122											Boron 5 B 10.811	Carbon 6 C 12.011	Nitrogen 7 N 14.007	Oxygen 8 O 15.999	Fluorine 9 F 18.998	Ne 10 Ne 20.18
Sodium 11 Na 22.99	Magnesium 12 Mg 24.305											Aluminum 13 Al 26.982	Silicon 14 Si 28.086	Phosphorus 15 P 30.974	Sulfur 16 S 32.065	Chlorine 17 Cl 35.453	Argon 18 Ar 39.948
Potassium 19 K 39.098	Calcium 20 Ca 40.078	Scandium 21 Sc 44.956	Titanium 22 Ti 47.867	Vanadium 23 V 50.942	Chromium 24 Cr 51.996	Manganese 25 Mn 54.938	Iron 26 Fe 55.845	Cobalt 27 Co 58.933	Nickel 28 Ni 58.693	Copper 29 Cu 63.546	Zinc 30 Zn 65.38	Gallium 31 Ga 69.723	Germanium 32 Ge 72.64	Arsenic 33 As 74.922	Selenium 34 Se 78.96	Bromine 35 Br 79.904	Krypton 36 Kr 83.798
Rubidium 37 Rb 85.468	Strontium 38 Sr 87.62	Yttrium 39 Y 88.906	Zirconium 40 Zr 91.224	Niobium 41 Nb 92.906	Molybdenum 42 Mo 95.9*6	Technetium 43 Tc 98	Ruthenium 44 Ru 101.07	Rhodium 45 Rh 102.91	Palladium 46 Pd 106.42	Silver 47 Ag 107.87	Cadmium 48 Cd 112.41	Indium 49 In 114.82	Tin 50 Sn 118.71	Antimony 51 Sb 121.76	Tellurium 52 Te 127.6	Iodine 53 I 126.9	Xenon 54 Xe 131.2
Cesium 55 Cs 132.91	Barium 56 Ba 137.33		Hafnium 72 Hf 178.49	Tantalum 73 Ta 180.95	Tungsten 74 W 183.84	Rhenium 75 Re 186.21	Osmium 76' Os 190.23	Iridium 77 Ir 192.22	Platinum 78 Pt 195.08	Gold 79 Au 196.97	Mercury 80 Hg 200.59	Thallium 81 Tl 204.38	Lead 82 Pb 207.2	Bismuth 83 Bi 208.98	Polonium 84 Po 209	Astatine 85 At 210	Radon 86 Rn 222
Francium 87 Fr 223	Radium 88 Ra 226		Rutherfordium 104 Rf 261	Dubnium 105 Db 262	Seaborgium 106 Sg 266	Bohrium 107 Bh 264	Hassium 108 Hs 277	Meitnerium 109 Mt 268	Darmstadium 110 Ds 271	Roentgenium 111 Rg 272							

Lanthanum 57 La 138.91	Cerium 58 Ce 140.12	Praseodymium 59 Pr 140.91	Neodymium 60 Nd 144.24	Promethium 61 Pm 145	Samarium 62 Sm 150.36	Europium 63 Eu 151.96	Gadolinium 64 Gd 157.25	Terbium 65 Tb 158.93	Dysprosium 66 Dy 162.5	Holmium 67 Ho 164.93	Erbium 68 Er 167.26	Thulium 69 Tm 168.93	Ytterbium 70 Yb 173.05	Lutetium 71 Lu 174.97
Actinium 89 Ac 227	Thorium 90 Th 232.04	Protactinium 91 Pa 231.04	Uranium 92 U 238.03	Neptunium 93 Np 237	Plutonium 94 Pu 244	Americium 95 Am 243	Curium 96 Cm 247	Berkelium 97 Bk 247	Californium 98 Cf 251	Einsteinium 99 Es 252	Fermium 100 Fm 257	Mendelevium 101 Md 258	Nobelium 102 No 259	Lawrencium 103 Lr 262

Figure 1. Periodic Table of the Elements.

Electrical conduction in solids is best understood in terms of a model that physicists call “Band Theory.” Those interested in understanding more about this theory should review the material contained in Appendix A of this text. The next section provides an abridged description of how electrical conduction occurs in semiconductors; some knowledge of atoms, electrons, chemical bonds and electron orbitals is required. Those unfamiliar or “rusty” on these concepts can find readily available texts as well as entertaining and instructive Internet videos on the basics [3] [4] [5] [6] [7].

2. Electrical Conduction in Semiconductors

Pure semiconductor crystals are not particularly good electrical conductors, although their conductivities are much greater than those of insulators (Table 2). The electrical property that makes semiconducting materials, and especially silicon, so valuable in electronics and other device applications arises from the fact that their electrical conductivity can be continuously varied through the controlled incorporation of dopant atoms into the crystal lattice. This property allows doped and undoped silicon to be used for the control of electrical current in a multitude of electronic devices, including diodes, capacitors and transistors. Transistors, especially, are extremely important in modern technology. They can be likened to a valve that controls the flow of electricity: a Metal-Oxide-Semiconductor Field Effect Transistor (MOSFET) device, Figure 2(a), behaves as an electrical on/off switch, while a Bipolar Junction Transistor (BJT), Figure 2(b) behaves primarily as an amplifier (controlling signals) in its linear range. Such transistors can only be fabricated using semiconductor materials.

Material	Resistivity (Ω-cm)	Conductivity (Ω^{-1}-cm^{-1})
Insulators		
Hard Rubber	1-100 x 10^{13}	1 × 10^{-15} to 1 × 10^{-13}
Glass	1-10000 x 10^{9}	1 × 10^{-13} to 1 × 10^{-9}
Quartz (fused)	7.5 x 10^{17}	1.33 × 10^{-18}
Semiconductors		
Carbon (Graphite)	3-60 x 10^{-5}	1.67 × 10^{3} to 3.33 × 10^{4}
Germanium	1-500 x 10^{-3}	2.0 to 1.00 × 10^{3}
Silicon	0.10- 60	1.67 × 10^{-2} to 10
Metals		
Silver	1.63 x 10^{-8}	6.17 × 10^{7}
Copper (annealed)	1.72 x 10^{-8}	5.95 × 10^{7}
Aluminum	2.65 x 10^{-8}	3.77 × 10^{7}

Table 2. Relative electrical resistivities and conductivities of some metals.

Today silicon is the primary semiconducting material employed for large scale electronic device manufacturing and we will use it to illustrate the electrical properties of semiconductors. Keep in mind that the properties described for silicon can equally describe other semiconducting elements in Table 2, as well as the various compound semiconductors (GaAs, InP, etc.) that are used in more limited applications (albeit with some variation in the finer details).

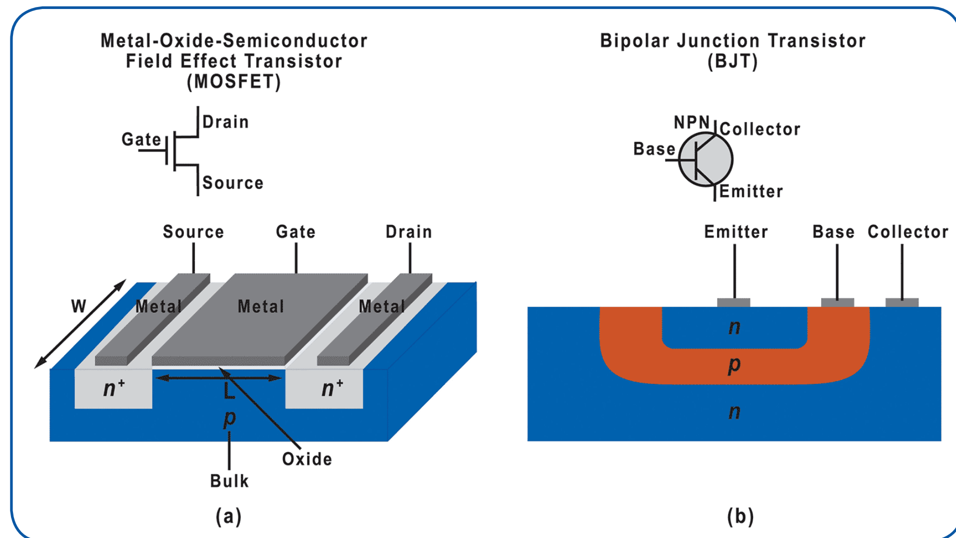


Figure 2. (a) Metal-Oxide-Semiconductor Field Effect Transistor (MOSFET) symbol and structure; (b) Bipolar Junction Transistor (BJT) symbol and structure.

Pure silicon and other elemental semiconductors are commonly referred to as *intrinsic semiconductors* [8] [9]. The term intrinsic means that electrical conductivity is an inherent property of the semiconducting material and independent of the presence of additives. In addition to intrinsic semiconductors, semiconductor device fabrication also makes use of extrinsic semiconductors (see Figure 3) that depend on the presence of dopant impurities to increase the electrical conductivity of the material. Table 2 shows the room-temperature resistivities and conductivities for a number of metals, intrinsic semiconductors and insulators. Elemental silicon has a resistivity of 0.10-60 Ω·cm—significantly higher than metallic resistivities; however, it is many orders of magnitude smaller than the resistivity of a typical insulator.

mks

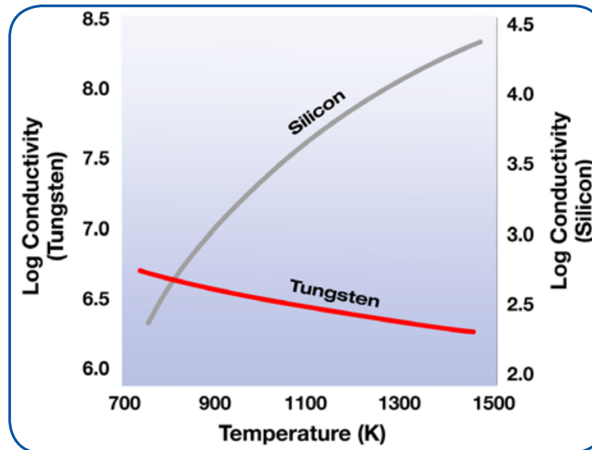


Figure 3. The electrical conductivity characteristics of a semiconductor (silicon) vs. those of a metal (tungsten) [10].

A further difference between the electrical characteristics of semiconductors and those of metals lies in the way their conductivities vary with temperature. At absolute zero (0 K), the electrical conductivity of a semiconductor has a value of zero (i.e., the conductivity is at its minimum) whereas a metal exhibits its maximum electrical conductivity at absolute zero; furthermore, conductivity increases with increasing temperature in a semiconductor, whereas it goes down with increasing temperature in a metal. Figure 3 shows a graph comparing the temperature variation of the electrical conductivity of a semiconductor (silicon) versus a metal (tungsten). The different behaviors of electrical conductivities seen in Figure 3 can be understood in terms of the different mechanisms that govern electrical conductivity in a metal versus a semiconductor.

Metals have many free electrons available to conduct electricity. In a metal, the valence band and the conduction band overlap, so electrons are free to move into many available and vacant energy levels (see Appendix A for a description of valence and conduction bands). When an electrical potential is applied across a piece of metal, these electrons can flow freely from higher to lower potential. The conductivity is limited only by the amount of electron scattering that occurs due to collisions between the flowing electrons and fixed atoms in the metal lattice. At 0 K, the atoms in the metal lattice are at rest and electron scattering is at its lowest possible value. As the temperature of the metal increases, thermal energy causes the atoms in the lattice to vibrate. These vibrations increase the effective diameter which an atom presents to the flow of electrons, increasing each atom's ability to scatter electrons. Thus, as the temperature rises from 0 K, the metal atoms vibrate with continuously increasing amplitudes, increasing the degree of electron scattering and producing the observed reduction in conductivity seen in Figure 3.

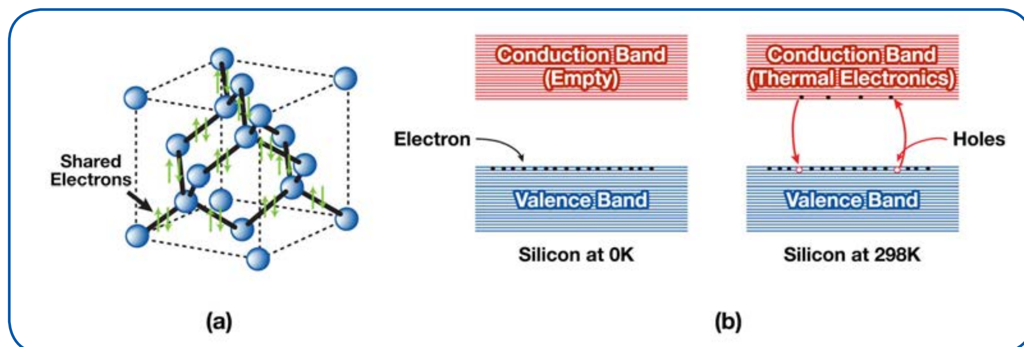


Figure 4. Silicon structure and electrical conduction.

Figure 4 can help in understanding how the mechanism of electrical current flow in a semiconductor differs from current flow in a metal. Figure 4(a) shows silicon atoms in a crystal, along with a crude schematic of how the valence electrons associated with each silicon atom are shared to produce four bonds containing eight electrons in every atom's valence shell. As you will recall from basic chemistry, covalently bound atoms require eight electrons in their valence shell to be stable. As Figure 4(a) shows, every valence electron in a silicon atom is involved in a covalent bond and so, unlike in a metal, none are free for electrical conduction; silicon, at absolute zero, is an insulator.

Figure 4(b) shows a crude schematic of the band structure (see Appendix A) of silicon at 0 K vs. the same structure at 298 K (25°C). At 0 K, the valence band is full and the conduction band empty and there is no thermal energy available to raise a valence electron into the conduction band. As the temperature increases, the silicon absorbs thermal energy. While this energy increases the thermal vibrations of the silicon atoms (producing increased atom-electron scattering), any loss in electrical conductivity due to this phenomenon is swamped by the increase in electrical conductivity due to the thermal promotion of electrons from the valence to the conduction band. This effect is not present in metals since the conduction band overlaps the valence band. Promotion of an electron in the valence band to the conduction band creates an electron-hole pair which constitutes mobile carriers in both the conduction band (electrons) and the valence band (holes). Electrical current will now flow in the material when a potential is applied. From Figure 4(b), we can see that in an intrinsic semiconductor, there is a balance between the number of electrons in the conduction band and the number of holes in the valence band. The concept of holes is important in discussions of electronic devices because electrical current is conventionally depicted as being due to the movement of holes. Since holes represent the absence of negative charges (electrons), it is useful to think of them as positive charges. By convention, electric fields [10] [11] (and therefore electric potential) are depicted as a vector extending outward from areas of positive charge. Therefore, when considering current flow, electrons move in a direction opposite to the applied electric field direction (i.e., toward the positive charge), while holes move in the direction of the electric field.

IIIA	IVA	VA
Boron 5 **B** 10.811	Carbon 6 **C** 12.011	Nitrogen 7 **N** 14.007
Aluminum 13 **Al** 26.982	Silicon 14 **Si** 28.086	Phosphorus 15 **P** 30.974
Gallium 31 **Ga** 69.723	Germanium 32 **Ge** 72.64	Arsenic 33 **As** 74.922
Indium 49 **In** 114.82	Tin 50 **Sn** 118.71	Antimony 51 **Sb** 121.76

P-type Dopants

N-type Dopants

Elemental Semiconductors

Figure 5. Elemental semiconductors and the most common n- and p-type dopants used to create extrinsic semiconductors.

When compared with metals, semiconductor electrical conductivities are not very high, as can be seen from the values in Table 2. However, semiconductor conductivities can be significantly increased (and, in fact, fine-tuned to target values) by adding impurity atoms known as dopants. "Doped" semiconductor materials are commonly referred to as extrinsic semiconductors [8] [12]. Figure 5 shows the area of the Periodic Table that contains the elemental semiconductors and the dopant elements that can be used to create an extrinsic semiconductor. The dopants are those elements on either side of Column IVA in the Periodic Table. A dopant atom is similar in size to the intrinsic semiconductor atom in the same row but has either one fewer or one greater electron in its valence shell. The dopant atoms can easily replace the semiconductor atom in its crystal lattice. Figure 6 shows the insertion of either a Group VA (n-type dopant, phosphorus) or a Group IIIA (p-type dopant, boron) atom into the silicon crystal lattice as well as the effect that the substitution of the silicon atom by a dopant atom has on the band structure of the silicon.

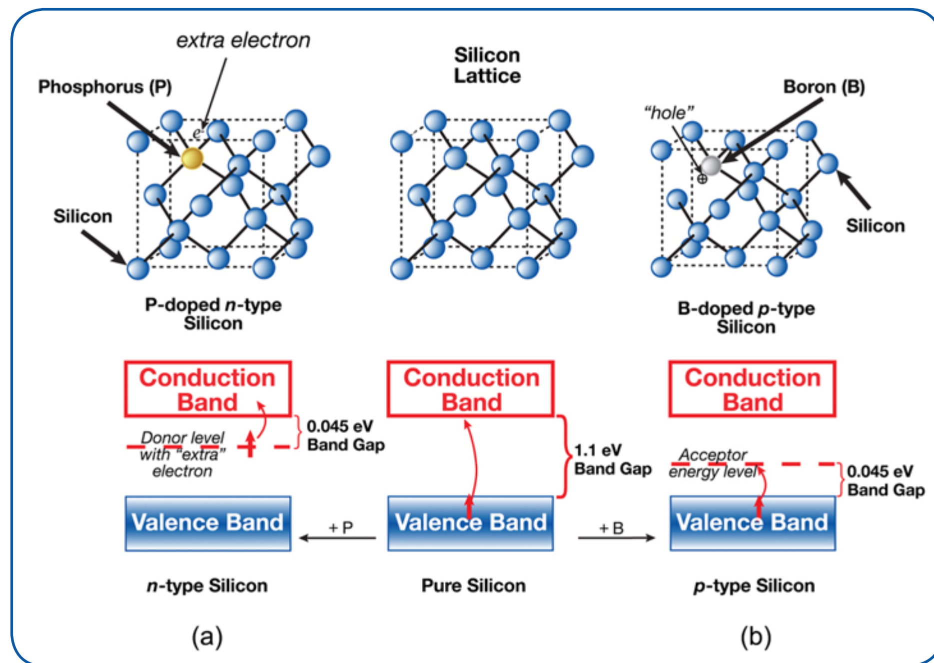


Figure 6. The effect of impurities on the band gap in extrinsic silicon [10].

When phosphorus, an n-type (donor) dopant is substituted for silicon, the extra electron in phosphorus's valence shell is at a higher energy level (the donor level in Figure 6(a)) than the electrons in the filled silicon valence band. This dopant electron can easily jump the small energy gap (0.045 eV for P, much less than the 1.1 eV band gap for pure silicon at room temperature; typical donor band gaps range from .039 to .054 eV) to the empty conduction band and become a free carrier. The concentration of free carriers (and therefore the electrical conductivity) in an n-type extrinsic semiconductor is roughly proportional to the dopant concentration in the material. Since the free electron generated from dopant atoms in an n-type semiconductor does not produce a corresponding hole in the filled valence band, the dominant charge carriers in an n-type semiconductor are electrons. In a similar manner, boron can be easily substituted for silicon in the bulk crystal lattice, as shown in Figure 6(b). Since boron has one less electron than silicon, this creates a hole (an acceptor level) at an energy just greater than the top of the filled silicon valence band. Under the influence of an electric field, it takes little energy (0.045 eV for B; typical acceptor band gaps range from 0.045 to 0.160 eV) for an electron in the valence band to be promoted to this acceptor level, leaving behind a hole in the valence band. Other bound electrons in the valence band can then jump into this hole and the subsequent hole movement carries the electrical current. Thus, for p-type extrinsic semiconductors, holes are the majority charge carriers for current flow. Figure 7 shows the relationship between dopant concentration and resistivity (the inverse of conductivity) for n-and p-type dopants in crystalline silicon.

B. Basic Device Structures

1. P-N Junction

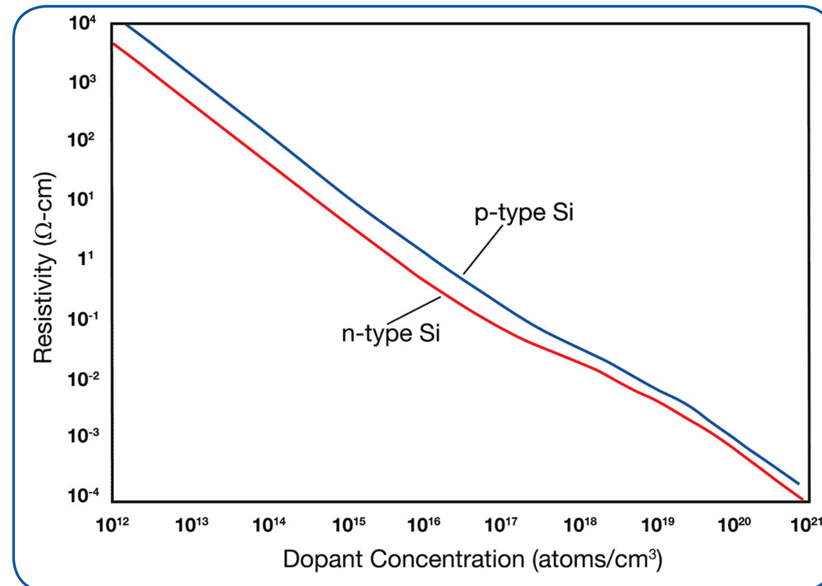


Figure 7. The dependency of the electrical resistivity at 298K on the doping concentration.

The ability to continuously vary a semiconductor's resistivity has only limited value when considered in isolation since a device composed of a single, doped semiconductor is simply a neutral, passive electrical component (i.e., a resistor). Rather, it is the combination of differently doped semiconductors with each other and with other materials that has leveraged the unique physical properties of semiconductors with the properties of more conventional insulators and conducting materials to create the solid state electronic devices that enable much of modern technology. The most basic of these devices is the P-N junction [13] [14].

When p-type semiconductor and n-type semiconductor materials are placed in physical contact, the area around the contact (known as the junction) behaves differently than either of the two source materials. Figure 8 illustrates the unique electrical field and charge conditions that exist at a P-N junction along with the forces at play on the free carriers. In isolation, an n-type semiconductor material has a high concentration of free electrons while a p-type semiconductor has a high concentration of free holes. When the two materials are in contact, nature demands that the concentration of holes and electrons be the same throughout the two materials; this phenomenon is called diffusion [15]. If electrons and holes were uncharged, neutral species, diffusion of holes and electrons would eventually make the local concentration of holes and electrons constant throughout the entire body of the joined p- and n-type materials. However, when electrons migrate from the n-type semiconductor to the p-type semiconductor and holes migrate from the p-type semiconductor to the n-type semiconductor, each type of carrier leaves behind fixed charges; positive in the n-type material (positive Group VA cations in the silicon lattice) and negative in the p-type material (negative Group IIIA anions in the silicon lattice). This sets up an electric field across the junction that eventually prevents further migration. The region over which the electric field extends is known as the depletion region; it is devoid of free carriers and acts as an insulator that prevents further diffusion of electrons and holes. A built-in electrical potential, V_D, is created by the electrical field at the P-N junction; in the case of silicon this potential has a value of about 0.7 V.

2. Diodes

Devices made of a simple P-N junction are solid state diodes (commonly referred to as a *P-N Junction Diode)* [13]. The characteristic response of a solid state diode to an external electrical potential or to an energy source such as light forms the basis for devices such as bipolar junction transistors, solar cells, LEDs, lasers and photodiodes. As discussed above and shown in Figure 8, an isolated P-N junction (i.e., not connected within a powered circuit) is in a state of equilibrium, with carrier generation and recombination and carrier diffusion and drift all balanced in the presence of the electric field across the depletion region. If one were to connect the P and N terminals of the diode shown in Figure 8, zero voltage would be observed across the diode and there would be no current flow. However, when an external potential is applied to the diode terminals, the unique properties of the P-N junction result in one-directional current flow, the unique characteristic of a diode. Figure 9 shows a schematic of a silicon diode under conditions of “zero,” “forward” and “reverse” bias. In the zero bias state Figure 9(a), there is no current flow, as described above.

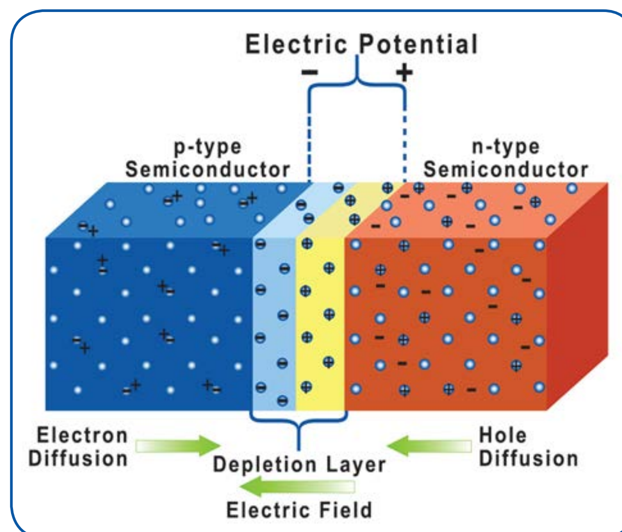


Figure 8. The P-N junction [16].

In the forward bias state Figure 9(b), the positive terminal of a power supply (i.e., a battery) is connected to the p-type material of the diode and the negative terminal of the battery to the n-type material (Figure 9). Under these conditions, the electric field due to the external voltage is oriented in the opposite direction of the electric field due to the depletion region (see Figure 8). When it is connected to a power supply in this manner, no current will flow through the diode until the external voltage exceeds the value of the built-in potential in the depletion region, i.e., 0.7 V in the case of a silicon P-N junction. Additionally, under forward bias conditions, the opposing E-field of the external voltage shrinks the depletion region of the P-N junction, making it much thinner; this has the effect of producing a very low resistance path through the P-N junction that will allow very large currents to flow through the diode with relatively small increases in voltage. This effect can be seen in the voltage/current curve for forward bias conditions that is shown in Figure 9. The effective limits to current flow through the P-N junction diode are so high that resistors are normally required in series with the diode in order to limit the current to values low enough to prevent thermal damage to the diode.

In the reverse bias state Figure 9(c), the positive terminal of a power supply is connected to the n-type material of the P-N junction diode while the negative terminal is connected to the p-type material. Under reverse bias, the positive voltage applied to the n-type material attracts the free electrons towards the positive terminal and the negative voltage applied to the p-type material attracts holes. This has the effect

of widening the depletion region, producing a high resistance to current flow in the diode. The resistance produced in this manner is high enough that essentially no current can flow through the device (with the exception of a small leakage current — see the curve in Figure 9(c)). This state exists under reverse bias conditions until the bias voltage becomes high enough to cause the diode's P-N junction to overheat and fail, producing a short in the electrical circuit.

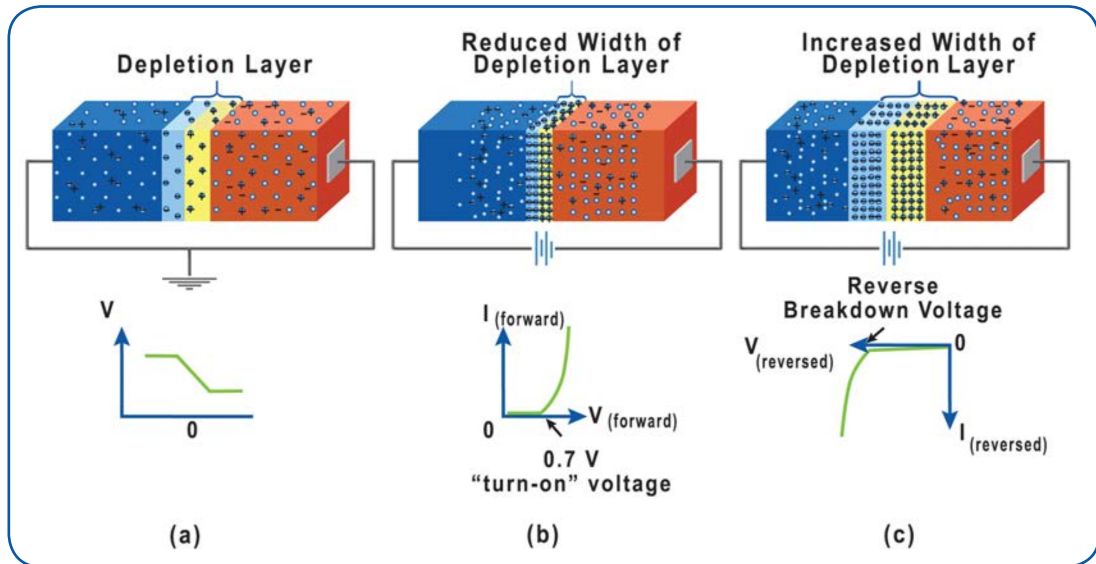


Figure 9. Current flow characteristics of a P-N junction diode [16].

3. Bipolar Junction Transistor

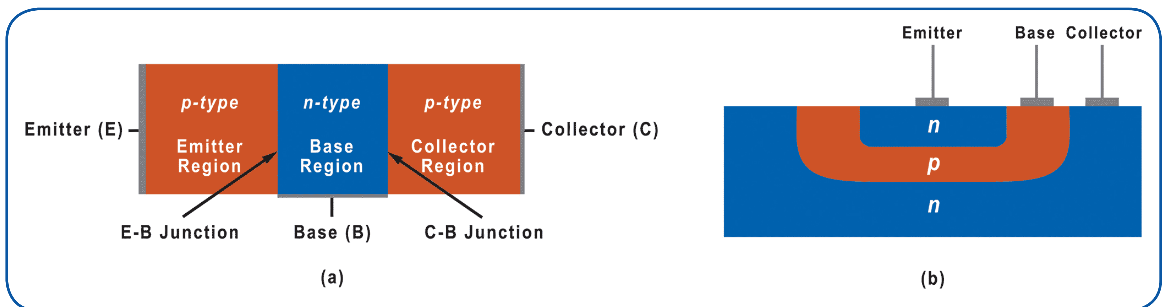


Figure 10. The bipolar junction transistor.

The P-N junction diode is the simplest electronic device that can be built using semiconducting materials (with the exception of a resistor). If we take two P-N junction diodes and physically join them in a back-to-back manner as shown in Figure 10, we can produce the most fundamental of microelectronic devices — a transistor. Transistors such as the one shown in Figure 10 are called bipolar junction transistors [13] [17]. Bipolar junction transistors (BJTs) were invented in late 1947 by William Shockley, Walter Brattain and John Bardeen at AT&T Bell Laboratories. They constituted what is arguably the first step in the microelectronics revolution and Shockley, Brattain and Bardeen were awarded the Nobel Prize in Physics in 1956 for this invention. The BJT is so named because its operation involves both electrons and holes. It can be configured as either a PNP transistor or an NPN transistor which can be constructed as either a three-layer sandwich as shown in Figure 10(a) or as a planar device on a semiconductor wafer

mks

as shown in Figure 10(b). The BJT is fabricated with an *emitter* and *collector* of similar semiconductor type (i.e., either p-type or n-type) at either end with a very thin *base* of opposite polarity in between the emitter and collector. In BJTs a small potential (for silicon devices, 0.7 V, enough to overcome the P-N junction potential) is applied to the base. This overcomes the junction potential between the emitter and the base, producing a small current flow from the emitter to the base. This current flow, in turn, regulates the much larger current flow from the emitter to the collector. For those interested in a more in-depth understanding of how the BJT operates, reference [17] provides an excellent video description. BJTs can be described as "current regulating devices that control the amount of current flowing through them in proportion to the amount of biasing voltage applied to their base terminal, acting like a current-controlled switch." [18]

4. MOSFET

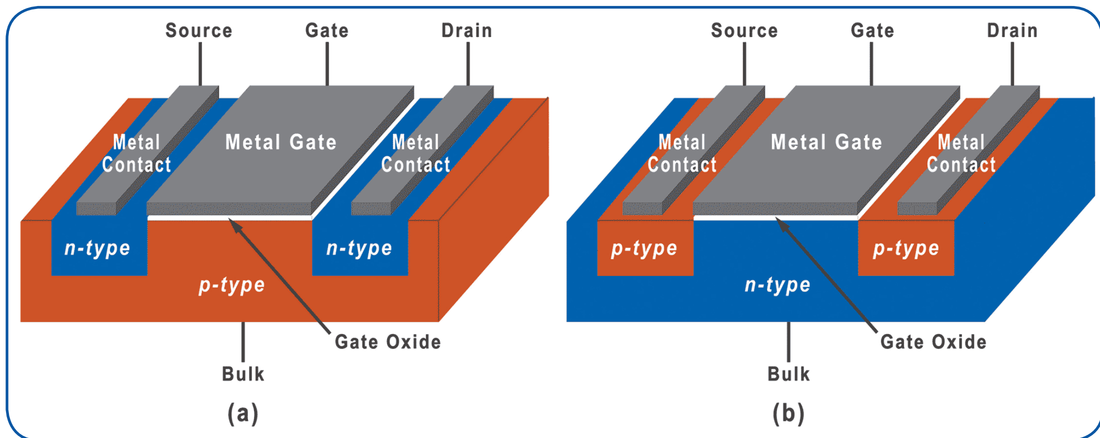


Figure 11. NMOS (a) and PMOS (b) MOSFETs.

Another very common form of transistor is the Metal Oxide Semiconductor Field Effect Transistor (MOSFET) [13] [19]. MOSFETs are planar surface devices that are the most commonly used variant of Field Effect Transistors (FETs); the reader may also encounter Junction Gate Field Effect Transistors (JFETs) and Insulated Gate Field Effect Transistors (IGFETs). MOSFETs are actually a subset of IGFETs. FETs can replace BJTs in most electronic circuits and have advantages for use in microelectronics since they consume and dissipate less power and they can be made much smaller than equivalent BJTs. Indeed, MOSFET devices constitute the ubiquitous "bit" switch that is set to 0 ("Off" state) or 1 ("On" state) in microelectronic logic devices (i.e., computers). FETs differ from BJTs in the way that current flow through the device is controlled; the primary current flow through an FET is controlled by a small voltage applied to one of the terminals rather than by a control current flow through any part of the device. Since MOSFETs are the most widely used FET, we will use these devices to describe the components and operation of this class of transistor. Figure 11 shows the elements of a typical MOSFET where the transistor is built into the surface of a silicon bulk substrate. MOSFETs can be built as either NMOS or PMOS transistors, depending on the polarities of the bulk, source and drain regions as shown in Figure 11. The majority of carriers in NMOS devices are electrons while those in PMOS devices are holes. An NMOS device is built on a p-doped silicon substrate that has had regions of n-type material which are created using ion implantation, as shown in Figure 11(a).These n-type regions are called the source and the drain. This situation is reversed in a PMOS device, as shown in Figure 11(b). The bulk material between the source and drain in a MOSFET is called the channel. A very thin insulating oxide layer covers the channel region; it is commonly referred to as a gate oxide. Finally, a conducting gate material, either a metal or highly doped polysilicon, is deposited on top of the gate oxide, creating the three-terminal device structure shown in Figure 11.

The operational characteristics of an nMOSFET transistor are shown in Figure 12; basic operating principles of a MOSFET device can be explained within the context of an NMOS device as shown in Figure 13. Figure 13(a) shows the device in the "Off" state with the gate, source and drain voltages at zero and the

bulk substrate connected to ground. Two P-N junctions exist between the n-type source/drain regions and the bulk p-type substrate. In operation, the potential between the drain and source (V_{DS}), and that between the gate and source (V_{GS}), are always positive. When a small voltage, (V_{GS}) is applied to the gate, the charge carrying holes in the p-type substrate are repelled away from the substrate surface. When V_{GS} reaches a threshold value (V_{TH}, the minimum gate to source voltage needed to turn the device on; this is less than the 0.7 V required in BJTs, typically 0.2-0.25 V in modern logic processors), the region under the gate becomes completely depleted of charge, producing a region in the substrate called the "depletion zone". Further increases in V_{GS} attract electrons from the electron-rich source (V_{GS}) and drain (V_{GD}) regions into the region under the gate, producing an n+ region known as the "inversion layer", shown in Figure 13(b). This inversion layer is a conducting channel that connects the two n-type regions at the source and drain; it will allow electrons to flow from the source to the drain when there is a positive voltage, $V_{DS,}$ between the source and drain. To assure that the induced inversion channel extends all the way from source to drain, the MOSFET gate structure slightly overlaps the edges of source and drain (the latter is achieved by a method known as a self-aligned process [13] [20]).

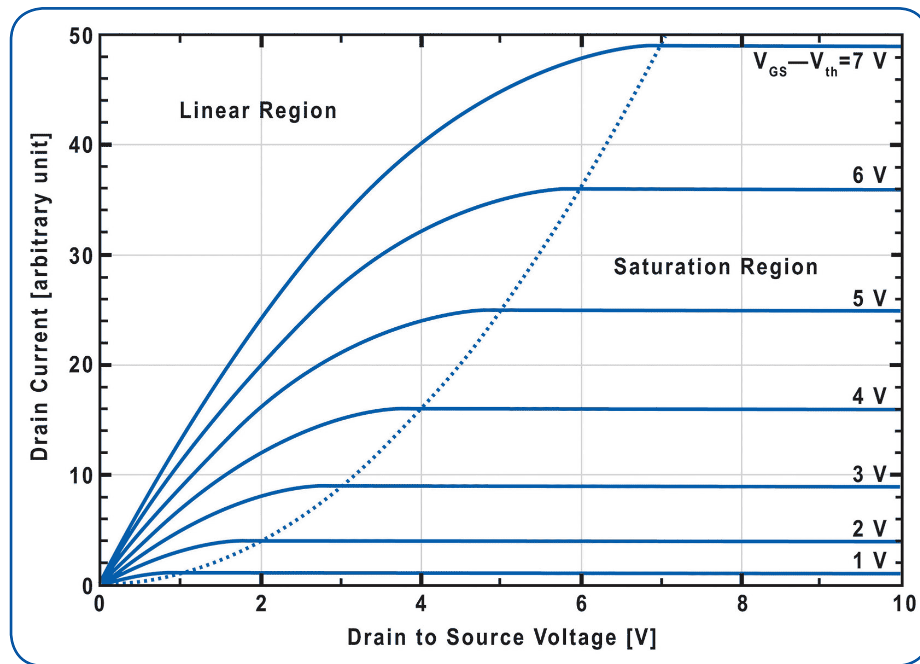


Figure 12. nMOSFET I-V characteristics [21].

When a drain-source bias, V_{DS}, is applied to a NMOS device in the above threshold conducting state, electrons move in the channel inversion layer from source to drain. At relatively small values of V_{DS}, the I/V characteristics of the device are linear with I_D (drain current) increasing with increasing V_D (drain voltage), as shown in Figure 13(b). Any change in the gate-source voltage V_{GS} will alter the electron density in the inversion layer, and, in this way, changes to V_{GS} can also control the device current. For this reason, characteristic I-V curves for NMOS devices typically depict a family of curves at different V_{GS}, as shown in Figure 12. When the drain voltage is increased to a value known as the saturation voltage, V_{SAT}, the charge and current flow characteristics in an NMOS device evolve, as depicted in Figure 13(c). The inversion layer under the gate becomes wedge shaped, wider (or deeper) near the source and essentially disappears (zero thickness) at the drain. This phenomenon is known as "pinch-off" [22] and the point where the inversion layer thickness is reduced to zero is called the "pinch-off point." Pinch-off occurs because, at V_{SAT}, the effective potential between the gate and substrate at the source end of the channel ($V_{eff} = V_{GS}$) is greater than the potential between the gate and the substrate at the drain end of the channel, which is just the

mks

potential needed to form the inversion layer called the threshold voltage ($V_{eff} = V_{GS}-V_{SAT} = V_{TH}$). Any higher voltage on the drain will cause the gate to substrate voltage to be reduced below the threshold voltage and the inversion layer will not be formed, creating the pinch-off point where there are no longer any mobile electron carriers in the channel. When the voltage applied to the drain is increased beyond V_{SAT}, the pinch-off point moves further towards the source, reducing the effective channel length, L_{eff}, as shown in Figure 13(d). Under these conditions, the area between the pinch-off point and the drain is fully depleted with no inversion layer. Since this region has no positive free carriers, there is no possibility for electron-hole recombination if an electron enters the region from the electron-rich source and, if there is an electric field across the depletion zone, the electron can freely transit to the drain. (Interested readers can find a more detailed discussion of this phenomenon in reference [22].) As can be seen from Figure 12, the current through the device becomes controlled solely by the gate voltage under drain saturation conditions.

Figure 13. Operating characteristics of an NMOS field effect transistor.

5. FinFETs

As MOSFET devices have continued to shrink in size, certain limitations have begun to impact their performance. Specifically, at nanometer dimensions, subthreshold current leakage becomes a significant factor (I_{off} becomes unacceptably high) and the potential on the drain begins to dominate the electrostatics in the channel region of the device. This causes a loss of gate control over the device current and an increase in I_{off}; these characteristics are known as short channel effects. Increases in I_{off} have obvious consequences for device performance. The problem especially impacts the devices used in mobile applications where it can lead to an unacceptable drain on battery power. Over the years, this problem has been overcome through the use of ever thinner gate oxides and the deployment of high-k dielectric materials that increase the gate-channel capacitance. These approaches have been effective up to about a decade or so ago when the continued thinning of gate oxides produced irremediable problems with gate leakage and gate-induced drain leakage in nanometer-scale devices. These problems have been resolved

through the use of an approach known as multi-gate FETs (MGFETs), a three-dimensional departure from traditional planar MOSFET designs.

Currently, the dominant non-planar transistor device geometry is the Fin Field Effect Transistor or FinFET [23]. Figure 14 shows a comparison between the structure of a conventional planar MOSFET and that of a FinFET. It can be seen that all of the familiar components of a MOS device are present in the FinFET (drain, source, gate, gate dielectric) and that the device more effectively separates the source and drain from the substrate silicon (FinFETs can be constructed on both bulk silicon and on buried oxide layers) and has a multiple gate configuration. The source and drain regions are now part of a vertical fin structure with the gate dielectric and gate electrode wrapped around it to produce multiple gates, one on either side of the fin as well as one on top of the device. FinFETs can also be extended to create a gate all around the device.

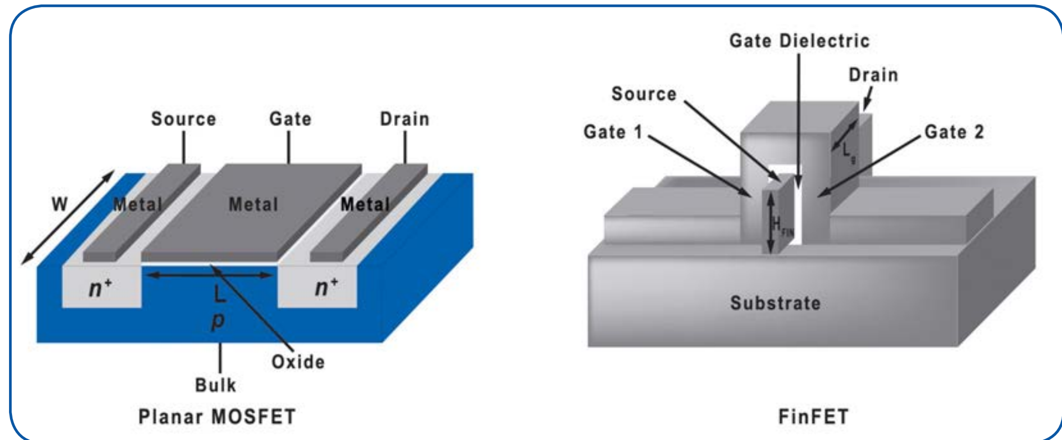


Figure 14. A comparison between the structures of a planar MOSFET transistor and a FinFET transistor.

FinFET designs produce much better control of the gate current by using more than one gate to control a single channel. They have much lower I_{off} leakage characteristics and can conduct significantly more current in the *On* state than comparable MOSFET designs.

6. Other Common Semiconductor Devices

In addition to the semiconductor devices that have been described in detail above, there are many other device types that the reader may encounter in interactions with semiconductor manufacturers. Below are some representative devices with graphic illustrations of their structures.

Flash Transistors

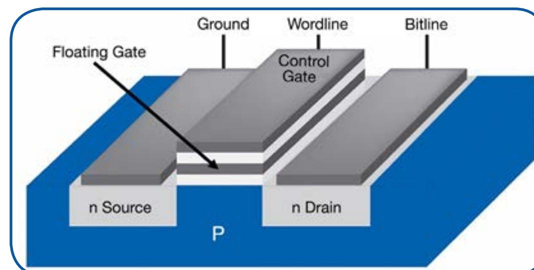


Figure 15. Flash transistor structure.

Flash transistors [24] are used in memory devices. Their design is similar to that of MOSFETs with the main design difference being that they employ two gates rather than one in a planar MOSFET design (Figure 15). When a potential is applied to the control gate of the flash transistor, current flows between the source and drain. It differs from a conventional MOS device in that some electrons tunnel through the gate oxide to produce a permanent charge on the floating gate. The presence of this charge constitutes a logic "1" in the memory device. The charge on the floating gate can be removed by placing a negative potential on the control gate. When the floating gate has no charge, it constitutes a logic "0" in the memory device. Flash memory transistors are thus a fundamental component of most non-volatile memory devices, which are devices that maintain their charge when power is turned off.

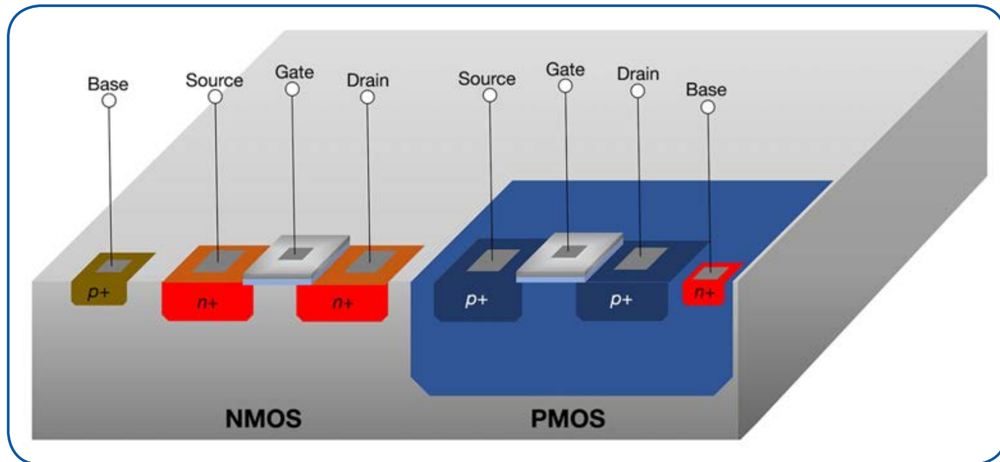


Figure 16. CMOS device structure.

CMOS

The acronym *CMOS* describes a Complementary Metal Oxide Semiconductor device [13] [25]. These composite devices combine NMOS and PMOS switches to achieve unique characteristics that allow a full suite of logic devices to be built [13] [26]. The device structure for a generic CMOS device is shown in Figure 16. CMOS devices have high immunity to noise and low power consumption, thus generating less waste heat, It is the most commonly used technology for very large scale integrated (VLSI) circuits.

Rectifier

Semiconductor rectifiers are *p-n* junction diodes specifically designed to rectify an alternating current — i.e., to give a low resistance to current flow in one direction and a very high resistance in the other direction.

Zener Diode

A Zener diode is a voltage regulator that is a *p-n* junction diode having a precisely tailored impurity distribution that produces a well-defined breakdown voltage. It can be operated in the reverse direction to serve as a constant voltage source, as a reference voltage for a regulated power supply, and as a protective device against voltage and current transients.

Varactor Diode

A varactor (variable reactor) is a device whose reactance, a measure of opposition to a change in current or voltage, can be varied in a controlled manner with a bias voltage. It is a *p-n* junction with a special impurity profile, and its capacitance variation is very sensitive to reverse-biased voltage. Varactors are widely used in parametric amplification, harmonic generation, mixing, detection and voltage-variable tuning applications.

Tunnel Diode

A tunnel diode is a single *p-n* junction in which both the *p* and *n* sides are heavily doped. The depletion layer is extremely narrow and under forward bias electrons can tunnel or pass directly through the junction, producing a negative resistance effect (i.e., the current decreases with increasing voltage). The tunnel diode is used in special low-power microwave applications, such as local oscillators and frequency-locking circuits.

Schottky Diode

A Schottky diode has a metal-semiconductor contact (e.g., an aluminum layer in intimate contact with an *n*-type silicon substrate). It is electrically similar to a *p-n* junction and is used extensively for high-frequency, low-noise mixer and switching circuits. This metal-semiconductor junction has a lower forward bias voltage than a semiconductor - semiconductor junction, and leads to higher switching speeds.

This page intentionally left blank.

II. Semiconductor Materials

A. Bulk Crystalline Silicon and Compound Semiconductors

1. Electronic Grade Polycrystalline Silicon (Polysilicon)

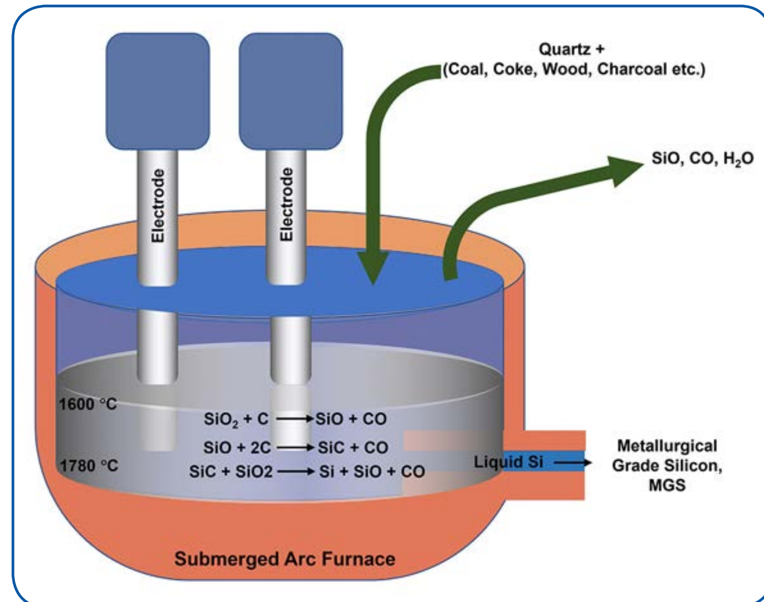
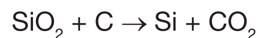


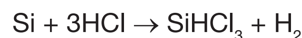
Figure 17. Schematic of a submerged electrode arc furnace used in the production of MG-Si (after [27]).

Silicon [28] [29] is the second most abundant element in the earth's crust (oxygen is the first). It occurs naturally in silicate (Si-O containing) rocks and sands. The elemental silicon used in semiconductor device manufacture is produced from high purity quartz and quartzite sands [30] [31], which contain relatively few impurities. Electronic grade silicon, the name used for the grade of silicon employed in semiconductor device manufacture, is the product of a chain of processes beginning with the conversion of quartz or quartzite sand to "*metallurgical grade silicon*" (MG-Si), in an electric arc furnace (Figure 17) according to the chemical reaction:



$$SiO_2 + C \rightarrow Si + CO_2$$

Silicon prepared in this manner is called "metallurgical grade" since most of the world's production actually goes into steel-making. It is about 98% pure [32]. For more details on the MG-Si process see references [32] and [33]. MG-Si is not pure enough for direct use in electronics manufacturing. A small fraction (5% – 10%) of the worldwide production of MG-Si gets further purified for use in electronics manufacturing. The purification of MG-Si to semiconductor (electronic) grade silicon is a multi-step process, shown schematically in Figure 18. In this process, MG-Si is first ground in a ball-mill to produce very fine (75% < 40 µM) particles which are then fed to a Fluidized Bed Reactor (FBR). There the MG-Si reacts with anhydrous hydrochloric acid gas (HCl), at 575 K (approx. 300ºC) according to the reaction:

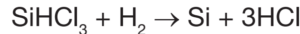


$$Si + 3HCl \rightarrow SiHCl_3 + H_2$$

The hydrochlorination reaction in the FBR makes a gaseous product that is about 90% trichlorosilane ($SiHCl_3$). The remaining 10% of the gas produced in this step is mostly tetrachlorosilane, $SiCl_4$, with some

dichlorosilane, SiH_2Cl_2. This gas mixture is put through a series of fractional distillations that purify the trichlorosilane and collect and re-use the tetrachlorosilane and dichlorosilane by-products. This purification process produces extremely pure trichlorosilane with major impurities in the low parts per billion range.

Purified, solid polycrystalline silicon is produced from high purity trichlorosilane using a method known as “The Siemens Process.” In this process, the trichlorosilane is diluted with hydrogen and fed to a chemical vapor deposition reactor. There, the reaction conditions are adjusted so that polycrystalline silicon is deposited on electrically-heated silicon rods according to the reverse of the trichlorosilane formation reaction:



$$SiHCl_3 + H_2 \rightarrow Si + 3HCl$$

By-products from the deposition reaction (H_2, HCl, $SiHCl_3$, $SiCl_4$ and SiH_2Cl_2) are captured and recycled through the trichlorosilane production and purification process as shown in Figure 18. The chemistry of the production, purification and silicon deposition processes associated with semiconductor grade silicon is more complex than this simple description. There are also a number of alternative chemistries that can be, and are, used for polysilicon production. See references [32] to [34] for more in-depth discussions of this topic. Wacker Chemistry AG has provided an on-line video on silicon production and its use in photovoltaic cells [35].

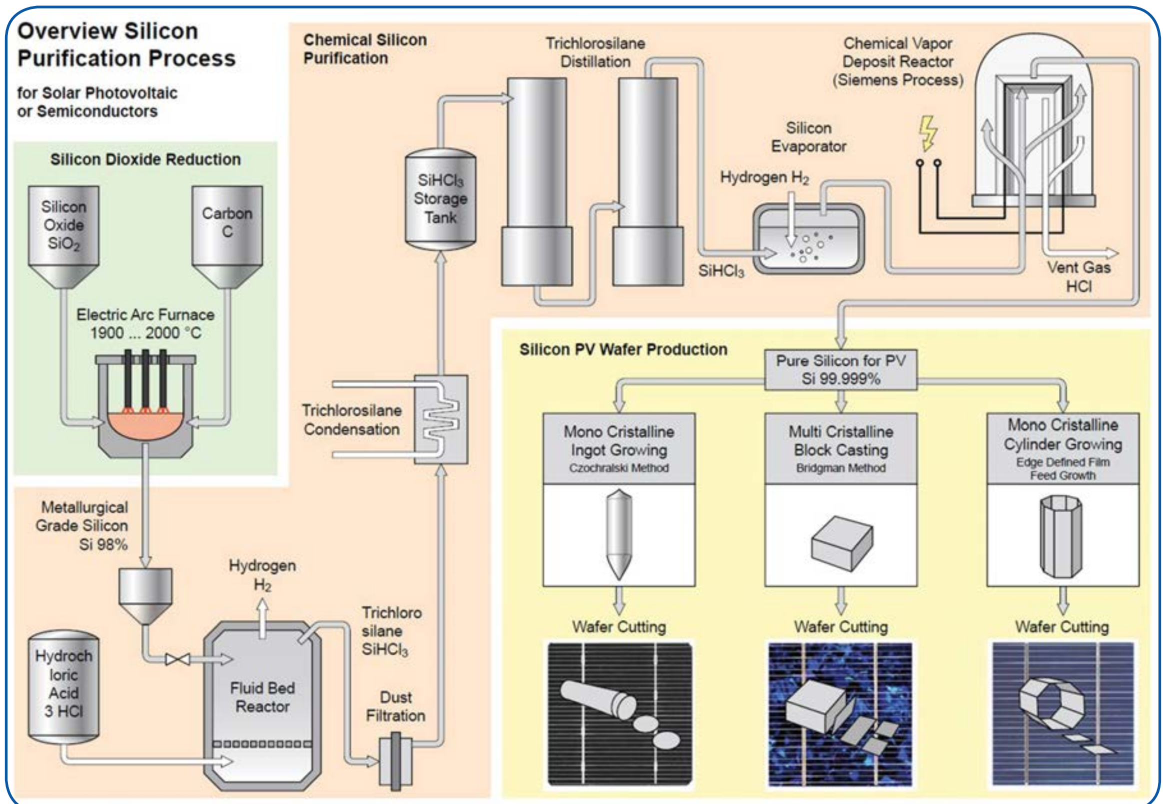


Figure 18. Process flow diagram for the production of semiconductor grade (electronic grade) silicon [36].

2. Single Crystal Silicon Wafer Manufacture

The silicon wafers so familiar to those of us in the semiconductor industry are actually thin slices of a large single crystal of silicon that was grown from melted electronic grade polycrystalline silicon. The process used in growing these single crystals is known as the Czochralski process after its inventor, Jan Czochralski [36]. Figure 19 shows the basic sequence and components involved in the Czochralski process.

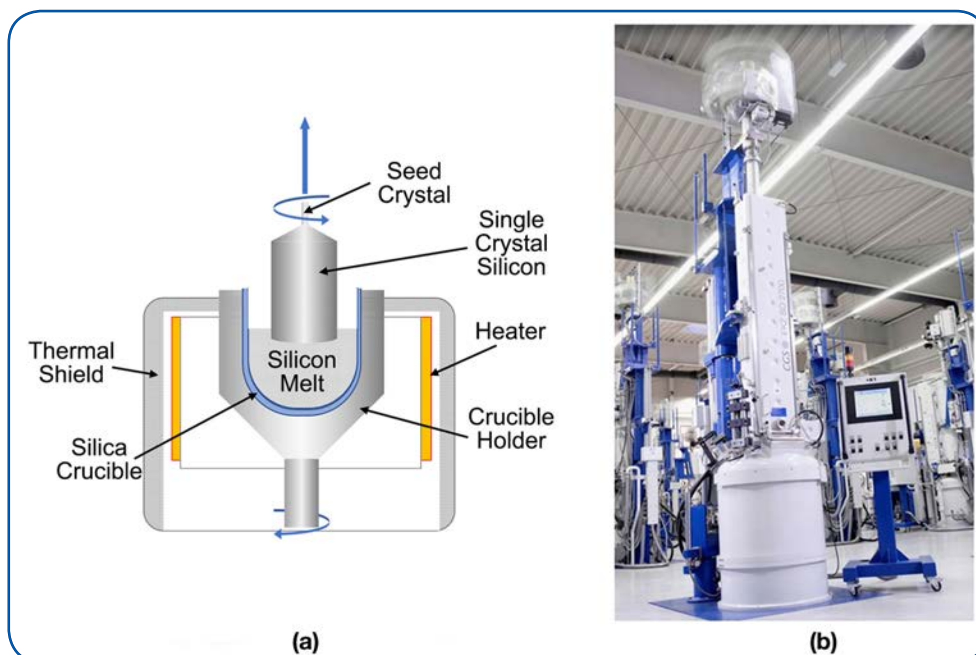


Figure 19. (a) Schematic of Czochralski process (b) Process equipment (reproduced with permission, PVA TePla AG 2017) [37].

The Czochralski process is carried out in an evacuable chamber, commonly referred to as a "crystal puller" that holds a large crucible, usually quartz, and an electric heating element (Figure 19(a)). Semiconductor grade polysilicon is loaded (charged) into the crucible along with precise amounts of any dopants such as phosphorus or boron that may be needed to give the product wafers specified P or N characteristics. Evacuation removes any air from the chamber to avoid oxidation of the heated silicon during the growth process. The charged crucible is electrically heated to a temperature sufficient to melt the polysilicon (greater than 1421°C). Once the silicon charge is fully melted, a small seed crystal, mounted on a rod, is lowered into the molten silicon. The seed crystal is typically about 5 mm in diameter and up to 300 mm long. It acts as a "starter" for the growth of the larger silicon crystal from the melt. The seed crystal is mounted on the rod with a known crystal facet vertically oriented in the melt (crystal facets are defined by "Miller Indices"). In the case of seed crystals, facets having Miller indices of <100>, <110> or <111> are typically chosen. References [36] and [32] provide brief textual descriptions of Miller Indices and their relation to crystal structures, while reference [38] provides an excellent video discussion of Miller Indices and crystal facet orientations. The crystal growth from the melt will conform to this initial orientation, giving the final large single crystal a known crystal orientation. Following immersion in the melt, the seed crystal is slowly (a few cm/hour) pulled from the melt as the larger crystal grows. The pull speed determines the final diameter of the large crystal. Both the crystal and the crucible are rotated during a crystal pull to improve the homogeneity of the crystal and dopant distribution. The final large crystal is cylindrical in shape; it is called a "boule."

Czochralski growth is the most economical method for the production of silicon crystal boules suitable for producing silicon wafers for general semiconductor device fabrication (known as CZ wafers). The method can form boules large enough to produce silicon wafers up to 450 mm in diameter. However, the method has certain limitations. Since the boule is grown in a quartz (SiO_2) crucible, some oxygen contamination is always present in the silicon (typically 10^{18} atoms cm^{-3} or 20 ppm). Graphite crucibles have been used to avoid this contamination, however they produce carbon impurities in the silicon, albeit at an order of magnitude lower in concentration. Both oxygen and carbon impurities lower the minority carrier diffusion length in the final silicon wafer. Dopant homogeneity in the axial and radial directions is also limited in Czochralski silicon, making it difficult to obtain wafers with resistivities greater than 100 ohm-cm.

Higher purity silicon can be produced by a method known as Float Zone (FZ) refining [36] [39]. In this method, a polycrystalline silicon ingot is mounted vertically in the growth chamber, either under vacuum or inert atmosphere. The ingot is not in contact with any of the chamber components except for the ambient gas and a seed crystal of known orientation at its base (Figure 20). The ingot is heated using non-contact radio-frequency (RF) coils that establish a zone of melted material in the ingot, typically about 2 cm thick. In the FZ process, the rod moves vertically downward, allowing the molten zone to move up the length of the ingot, pushing impurities ahead of the melt and leaving behind highly purified single crystal silicon. FZ silicon wafers have resistivities as high as 10,000 ohm-cm.

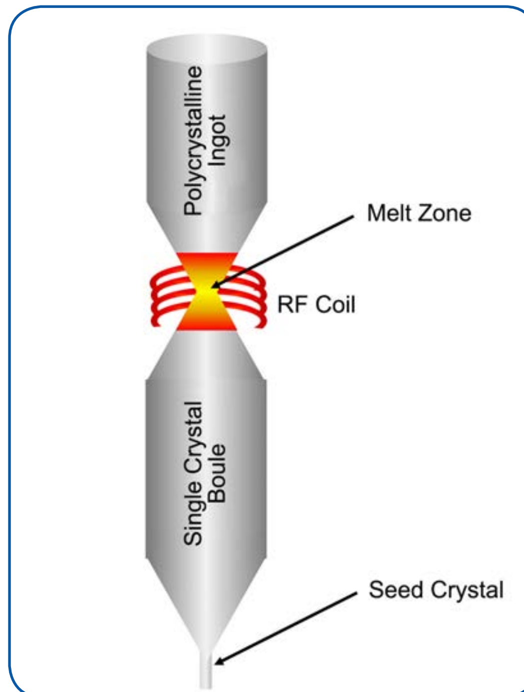


Figure 20. Float zone crystal growth configuration [40].

Once the silicon boule has been created, it is cut into manageable lengths and each length ground to the desired diameter. Orientation flats that indicate the silicon doping and orientation for wafers of less than 200 mm diameter are also ground into the boule at this stage. For wafers with diameters less than 200 mm, the primary (largest) flat is oriented perpendicular to a specified crystal axis such as <111> or <100> (see Figure 21). Secondary (smaller) flats indicate whether a wafer is either p-type or n-type. 200 mm (8-inch) and 300 mm (12-inch) wafers use a single notch oriented to the specified crystal axis to indicate wafer orientation with no indicator for doping type. Figure 21 shows the relationship between wafer type and the placement of flats on the wafer edge.

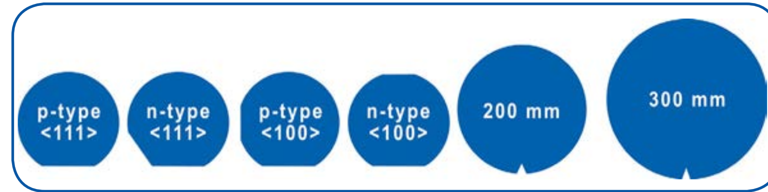


Figure 21. Wafer flat designators for different wafer orientation and doping.

After the boule has been ground to the desired diameter and the flats have been created, it is cut into thin slices using either a diamond encrusted blade or a steel wire. The edges of the silicon slices are usually rounded at this stage. Laser markings designating silicon type, resistivity, manufacturer, etc. are also added near the primary flat at this time. Both surfaces of the unfinished slice are ground and lapped [41] to bring all of the slices to within a specified thickness and flatness tolerance. Grinding brings the slice into a rough thickness and flatness tolerance after which the lapping process removes the last bit of unwanted material from the slice faces, leaving a smooth, flat, unpolished surface. Lapping typically achieves tolerances of less than 2.5 μm uniformity in wafer surface flatness.

The final stage in silicon wafer manufacture involves chemically etching away any surface layers that may have accumulated crystal damage and contamination during sawing, grinding and lapping; followed by chemical mechanical polishing (CMP) to produce a highly reflective, scratch and damage free surface on one side of the wafer. The chemical etch is accomplished using an etchant solution of hydrofluoric acid (HF) mixed with nitric and acetic acids that can dissolve silicon. In CMP, silicon slices are mounted onto a carrier and placed in a CMP machine [42] where they undergo combined chemical and mechanical polishing. Typically, CMP employs a hard polyurethane polishing pad combined with a slurry of finely dispersed alumina or silica abrasive particles in an alkaline solution. The finished product of the CMP process is the silicon wafer that we, as users, are familiar with. It has a highly reflective, scratch and damage free surface on one side on which semiconductor devices can be fabricated.

3. Compound Semiconductor Wafer Production

Compound semiconductors are important materials in many military and other specialty electronics devices such as lasers, high-frequency electronic devices, LEDs, optical receivers, opto-electronic integrated circuits, etc. [43]. GaN has been commonly used in many different commercial LED applications since the 1990's.

Table 3 provides a list of the elemental and binary (two element) compound semiconductors along with the nature of their band gap and its magnitude. In addition to the binary compound semiconductors, ternary (three element) compound semiconductors are also known and used in device fabrication. Ternary compound semiconductors include materials such as aluminum gallium arsenide, AlGaAs, indium gallium arsenide, InGaAs and indium aluminum arsenide, InAlAs. Quarternary (four element) compound semiconductors are also known and used in modern microelectronics.

The unique light-emitting ability of compound semiconductors is due to the fact that they are direct band gap semiconductors [44]. Table 3 denotes which semiconductors possess this property. The wavelength of the light emitted by devices built from direct band gap semiconductors depends on the band gap energy. By skillfully engineering the band gap structure of composite devices built from different compound semiconductors with direct band gaps, engineers have been able to produce solid state light emitting devices that range from the lasers used in fiber optic communications to high efficiency LED light bulbs. A detailed discussion of the implications of direct versus indirect band gaps in semiconductor materials is beyond the scope of this work. Refer to the material contained in references [44], [45] and [46].

Simple, binary compound semiconductors can be prepared in bulk, and single crystal wafers are produced by processes similar to those used in silicon wafer manufacturing. GaAs, InP and other compound semiconductor ingots can be grown using either the Czochralski or Bridgman-Stockbarger [47]

[48] [49] method with wafers prepared in a manner similar to silicon wafer production. Surface conditioning of compound semiconductor wafers, (i.e., making them reflective and flat) is complicated by the fact that at least two elements are present and these elements can react with etchants and abrasives in different fashions.

Compound Semiconductors				
Material System	**Name**	**Formula**	**Energy Gap (eV)**	**Band Type (I = indirect; D = direct)**
IV	Diamond	C	5.47	I
	Silicon	Si	1.124	I
	Germanium	Ge	0.66	I
	Grey Tin	Sn	0.08	D
IV-IV	Silicon Carbide	SiC	2.996	I
	Silicon-Germanium	Si_xGe_{1-x}	Var.	I
IV-VI	Lead Sulfide	PbS	0.41	D
	Lead Selenide	PbSe	0.27	D
	Lead Telluride	PbTe	0.31	D
III-V	Aluminum Nitride	AlN	6.2	I
	Aluminum Phosphide	AlP	2.43	I
	Aluminum Arsenide	AlAs	2.17	I
	Aluminum Antimonide	AlSb	1.58	I
	Gallium Nitride	GaN	3.36	D
	Gallium Phosphide	GaP	2.26	I
	Gallium Arsenide	GaAs	1.42	D
	Gallium Antimonide	GaSb	0.72	D
	Indium Nitride	InN	0.7	D
	Indium Phosphide	InP	1.35	D
	Indium Arsenide	InAs	0.36	D
	Indium Antimonide	InSb	0.17	D
II-VI	Zinc Sulfide	ZnS	3.68	D
	Zinc Selenide	ZnSe	2.71	D
	Zinc Telluride	ZnTe	2.26	D
	Cadmium Sulfide	CdS	2.42	D
	Cadmium Selenide	CdSe	1.70	D
	Cadmium Telluride	CdTe	1.56	D

Table 3. The elemental semiconductors and the binary compound semiconductors (taken from [47]).

B. Thin Films for Device Fabrication

1. Polycrystalline Silicon Thin Films

a. Overview

Thin films of polycrystalline silicon, commonly called polysilicon, have many applications in integrated circuit structures [32] [50]. Heavily doped polysilicon and poly-silicide films have been used as gate electrodes and interconnects in MOS devices (albeit largely replaced by metal in the modern devices). Heavily doped polysilicon has also been employed for the emitter in bipolar junction transistors. More lightly doped polysilicon has been used to fabricate polysilicon diodes for use as high value resistors and as trench refill in dielectric isolation technologies. Polysilicon films have also found use in trench capacitor fabrication. Outside of the area of integrated circuit manufacturing, polysilicon films are critical to the commercial photovoltaic (PV) industry. Most of the available commercial PV panels are based on polysilicon technology.

Polysilicon thin films are produced using a process known as chemical vapor deposition or CVD. We will discuss this methodology in more detail in Section B, Chapter V, but for now, it is sufficient to know the basics of the process. In a CVD process, the substrate upon which a thin film is to be deposited (i.e., an in-process silicon wafer) is placed in a sealed deposition chamber and heated to a specified temperature. Depending on the thin film process (i.e., polysilicon, silicon dioxide, silicon nitride, etc.) and the nature of the CVD process (i.e., thermal low pressure, high pressure CVD, or plasma enhanced CVD) the deposition temperature can range between 300°C and 800°C. The process is most often carried out under vacuum conditions, but atmospheric pressure CVD process can also be employed. Once the substrate is at temperature and a clean ambient (either vacuum or inert atmosphere) has been established, a continuous and controlled flow of a gaseous precursor containing the elements needed to form the thin film is introduced to the deposition chamber. These gas-phase compounds adsorb on the hot surface of the substrate where they undergo decomposition to produce the desired solid thin film and a gaseous by-product that is pumped or purged out of the chamber. In this way amorphous (no crystal ordering) and polycrystalline thin films can be formed in the device manufacturing process.

b. Material and Electrical Properties

Polycrystalline silicon films are composed of domains of small (typically about 0.1 µm) microcrystals of silicon, commonly called grains, separated by grain boundaries. Figure 22 shows a transmission electron micrograph (TEM) cross-section of a polysilicon film on a silicon substrate. In TEM images, crystalline material with different crystal planes show up as having different levels of contrast in the image (i.e., some orientations will appear light, some dark). This is because different crystal planes have different atomic densities which directly impacts the intensity of the TEM image. The TEM in Figure 22 clearly shows the kind of contrast differences that indicate polycrystalline material. The many areas of differing contrast show that the film is composed of many individual crystalline domains with their crystal facets oriented more or less randomly. Note that there is some tendency towards less-than-random grain orientation in most polysilicon films since substrate characteristics can influence the crystal growth direction of individual grains.

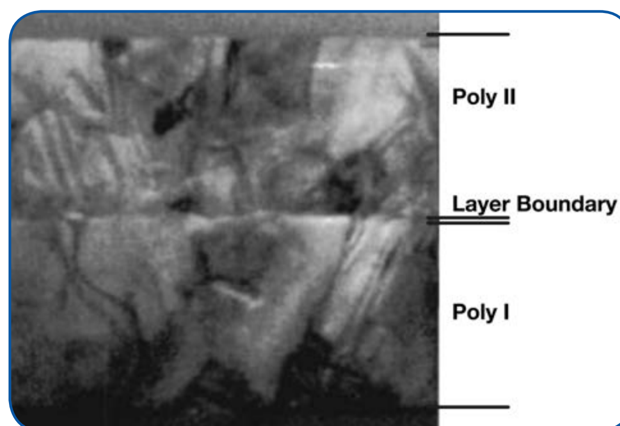


Figure 22. Cross-sectional TEM of polycrystalline silicon thin film [51] (Reproduced by permission of The Electrochemical Society).

mks

Since polycrystalline silicon is actually made up of small domains containing crystalline silicon, it is perhaps unsurprising that many of the materials properties of polysilicon thin films approximate those of bulk crystalline silicon. Table 4 provides a comparison of a few representative properties. It can be seen that properties that depend primarily on the average value of that property in either the bulk or the microcrystals within the thin film, i.e., density and thermal coefficient of expansion, have similar values in both bulk crystalline silicon and in polycrystalline silicon films. However, properties such as the temperature coefficient of electrical resistance that are influenced by the polycrystalline nature of polysilicon thin films and by the properties of the grain boundaries are significantly different. Additionally, thin films deposited on different substrates possess unique properties such as intrinsic stress that are not relevant to bulk crystalline silicon.

Property	Bulk Crystalline Silicon	Polycrystalline Si Thin Film
Density	2.329 g/cm^3	2.3 g/cm^3
Thermal Coefficient of Expansion	2.6 µm/(m·K) (at 25°C)	2 µm/(m·K) (at 25°C)
Temperature Coefficient of Resistance	-70 x 10^{-3} /°C	1 x 10^{-3} /°C
Intrinsic Stress	N/A	Compressive; 1-5 x 10^9 dynes/cm^3

Table 4. A comparison of some selected properties for bulk single crystal silicon and polycrystalline silicon thin films [32].

The primary origins for the differences between the properties of bulk crystal silicon and polycrystalline silicon thin films are the influences of intentional doping and of the grain boundaries present in polysilicon thin films. The influence of doping on semiconductor electrical properties (and hence on the electrical properties of individual grains in a polysilicon film) was discussed in the earlier section on semiconductor physics and does not need to be repeated here. In this section we will focus on describing the impact that grain boundaries have on polysilicon thin film properties. Grain boundary regions are composed of disordered, amorphous silicon that contains many structural defects and partially bonded silicon atoms. These crystalline and chemical defects in the grain boundaries affect the resistance to electrical current flow between the silicon microcrystals and can also influence dopant distribution within the individual microcrystals [52]. These influences have a strong impact on the electrical properties of the polysilicon film, changing their values relative to those observed for bulk, crystalline silicon.

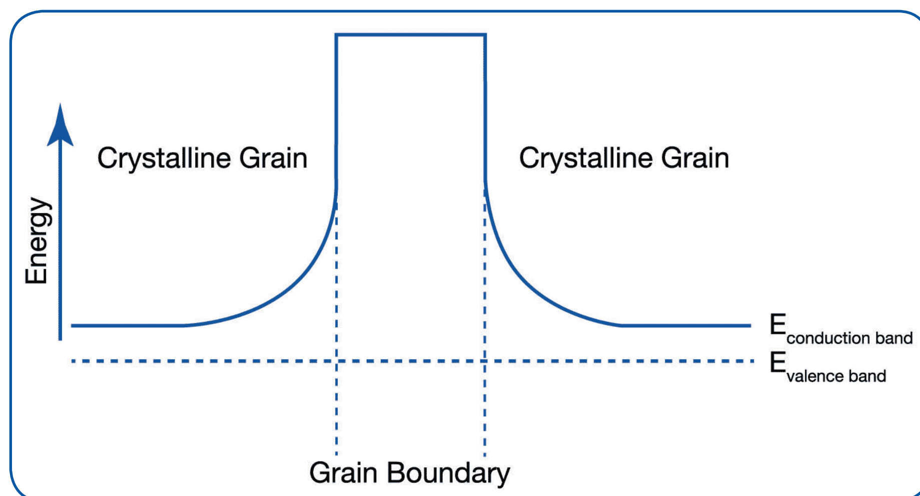


Figure 23. The energy band diagram for a polysilicon thin film near a grain boundary (zero applied voltage) [52].

Doped polysilicon films exhibit higher electrical resistivities than bulk monocrystalline silicon that has been doped to a similar level. This is due to the combined effect of high resistivity in the grain boundaries and thermally induced dopant segregation out of the microcrystals and into the grain boudaries. Resistivity within a grain boundary is relatively high due to the fact that grain boundaries, being amorphous, have high concentrations of chemical defects. Incomplete chemical bonds in a grain boundary act as traps for charge carriers, reducing their average concentration over the breadth of the thin film and therefore increasing the effective resistivity of the film as a whole. Figure 23 shows a representation of the energy band characteristics at a grain boundary (after references [32] and [52]) that portrays the energy band characteristics experienced by free charge carriers transiting a grain boundary. Grain boundaries are typically 5 to 10×10^{-8} cm wide and can be modeled as separate, amorphous silicon regions with an increased band gap.

The diffusion constants for dopant atoms (i.e., how fast a dopant atom will move through a material at a given temperature) are significantly higher in the grain boundaries than in the crystalline regions and this means that, at elevated temperatures, dopant atoms move more quickly through grain boundaries than they do through crystalline regions. Therefore, dopant atoms within a microcrystal that move into the grain boundary are transported away from that boundary faster than dopant atoms within the crystal can move to maintain a uniform concentration. This sets up a driving force for dopant migration out of the crystal that can produce relatively low dopant concentrations at the edge of the microcrystals. Phosphorus and arsenic, especially, are prone to migrate out of the crystalline grains to the grain boundaries. This phenomenon increases the average resistivity of the polysilicon film since dopant atoms that lie within the grain boundary are not effective in reducing electrical resistivity. This characteristic can present significant problems when trying to control dopant concentrations in polysilicon films at levels below saturation.

Consideration of the influence of grain boundaries on resistivity as described above can help us to understand the relationship between polysilicon film structure and electrical resistivity. Films with large grain sizes should exhibit lower electrical resistivity than films with small grain sizes, at comparable doping levels. This is due to the fact that grain boundary densities in thin film are reduced when the grain size is larger, and this leads to a relative reduction in the impact of dopant segregation and charge carrier trapping effects due to grain boundaries. In addition, grain size and dopant concentrations interact to determine the degree of dopant depletion arising from grain boundary effects. Under the influence of heat treatment, small grains with lower transit distances to nearby grain boundaries become depleted of dopants more quickly than do larger grains. This effect can be counteracted to some degree by higher dopant concentrations, since this results in narrower depletion regions that make full depletion of the grains more difficult.

c. Deposition of Polycrystalline Silicon Thin Films

CVD polysilicon process requires high purity, gaseous source materials that contain the silicon to be deposited [53], [54]. Typically, this source material is the chemical compound silane, SiH_4, but some polysilicon processes have been developed that use a related silicon hydride, disilane, Si_2H_6.

It is instructive to consider the purity requirements for CVD source gases. You may often hear the term “five nines purity” or something similar in discussions of the source gases used for CVD processes. This means that the gas must be 99.999% pure, a purity typical of electronic grade silane. Put another way, the maximum total impurity concentration permitted for CVD gases is one part in 100,000 or 10 ppm. The magnitude of 1 ppm can be understood using the following real-world comparison:

1 ppm =

- 1 oz. of salt in 62,500 lb. of sugar
- 1 oz. of sand in 3 tons of concrete
- 1 inch in 16 miles
- 1 minute in 1.9 year

A review of some of the physical properties of silicon hydrides is provided in Table 5.

	SiH_4	Si_2H_6	Si_3H_8	n-Si_4H_{10}
Melting Point (ºC)	-185	-132.5	-117.4	-85
Boiling Point (ºC)	-111.9	-14.5	52.9	107
Density at M.P. (g/cc)	0.68	0.69	0.725	0.825
Vapor Pressure (mm @ 0 ºC)			95	25 (20°C)
(mm @-118 ºC)	530	1		
Heat of Formation (kJ/mol)	+34.31	+75	+121	

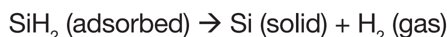
Table 5. Selected physical properties of some silicon hydrides (taken from [55] [56]).

Of the silicon hydrides shown in Table 5, only silane and disilane are indefinitely stable at 25ºC and they are the primary, commercially available silicon hydrides used for CVD polysilicon processes (tri- and tetrasilane, Si_3H_8 and Si_4H_{10}, respectively, are available as specialty chemicals). Silane and disilane are supplied as high pressure compressed gases and it is important to be aware of safe handling procedures for high pressure compressed gases. There are a number of videos available on the web that can provide basic instructions for the safe handling of compressed gas cylinders – example, references [57] and [58]. All silicon hydrides are pyrophoric [59] – meaning they will spontaneously burst into flames (or explode) if they are exposed to air, which makes working with these materials particularly hazardous. Silane and disilane are thus doubly dangerous and special training is needed for anyone who regularly deals with these materials [60], [61], [62].

Polysilicon thin films are most often deposited using surface pyrolysis of silane (pyrolysis: chemical decomposition produced by heating) according to the chemical reaction sequence [32]:

$$SiH_4 \text{ (gas)} + \text{surface site} \rightarrow SiH_4 \text{ (adsorbed)}$$

$$SiH_4 \text{ (adsorbed)} \rightarrow SiH_2 \text{ (adsorbed)} + H_2 \text{ (gas)}$$



$$SiH_2 \text{ (adsorbed)} \rightarrow Si \text{ (solid)} + H_2 \text{ (gas)}$$

The overall reaction can be written as:

$$SiH_4 \text{ (gas)} \rightarrow Si \text{ (solid)} + 2H_2 \text{ (gas)}$$

This reaction is most often performed under vacuum conditions (in a process known as low pressure chemical vapor deposition – LPCVD) at temperatures ranging between 580ºC and 650ºC.

Silane-based LPCVD polysilicon process characteristics have been studied extensively over the past three decades. The deposition rate increases with both temperature and pressure, as shown in Figure 24 [63]. Practically, silane-based LPCVD polysilicon deposition temperatures are limited to values below 650ºC since silane reacts preferentially in the gas phase rather than on the substrate surface above this temperature. Too much gas phase reaction in the LPCVD process produces rough, poorly adhering films and particles that are device killers. Silane-based polysilicon processes performed at temperatures below about 580ºC exhibit deposition rates that are too low to be of any practical use in device fabrication. However, disilane-based polysilicon processes have higher deposition rates at lower temperatures that allow polysilicon processes to operate down to about 500°C. It is important to understand that a mixed amorphous/polycrystalline silicon film is produced in any LPCVD polysilicon process that is carried out at temperatures below about 580ºC and that full polycrystallinity is only achieved when the films are annealed well above this temperature. LPCVD polysilicon films are typically annealed at temperatures ranging between 800ºC and 1000ºC. The relative proportion of amorphous material in a deposited thin film is also dependent on the partial pressure of silicon-containing reactant during the deposition process. Films deposited at higher silane pressures, for example, exhibit significant amorphous silicon content up to temperatures as high as about 650ºC.

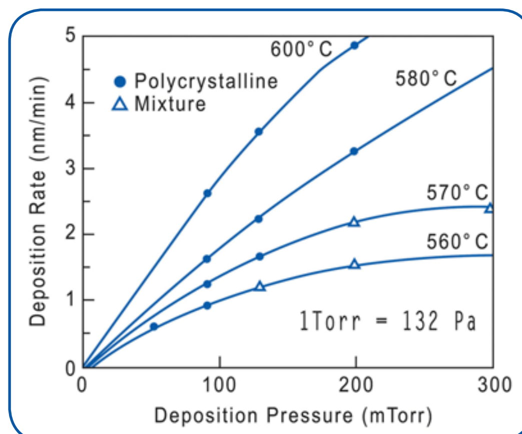


Figure 24. Deposition rate and crystallinity for silane-based LPCVD polysilicon thin films [63].

Quite complex but reasonably well-understood relationships exist between polysilicon deposition process parameters and the physical properties of the deposited polysilicon thin film. It is beyond the scope of this overview to go into great detail in reviewing these relationships, however, it is possible to give the reader an appreciation for how some of these interactions can play out. Figure 25 shows TEMs of silane-based polysilicon films in the as-deposited form [64]. Each film was deposited at a different temperature, varying between 560ºC and 640ºC. The TEM analysis indicates that the films deposited at 560ºC (upper left image, Figure 25(a)) are almost completely amorphous while those deposited at 640ºC (lower right in Figure 25(a)) are nearly fully polycrystalline. TEM images were obtained for the same films after they had been annealed in hydrogen for 30 minutes and these are shown in Figure 25(b). While we won't go into the underlying science here, it is easy to see that the temperature of deposition has a profound effect on the crystalline characteristics of the annealed films. After annealing, the films deposited at 560ºC and 580ºC have much larger grain size than do the annealed film that had been deposited at 620ºC and 640ºC. Similar relationships are found between other aspects of the physical structure of polysilicon thin films and the deposition process pressure, the process gas flows and reaction chamber configuration. Excellent discussions of these relationships and other aspects of polysilicon processes and thin films can be found in textbooks by Stanley Wolf and Richard Tauber [32] and by Ted Kamins [50].

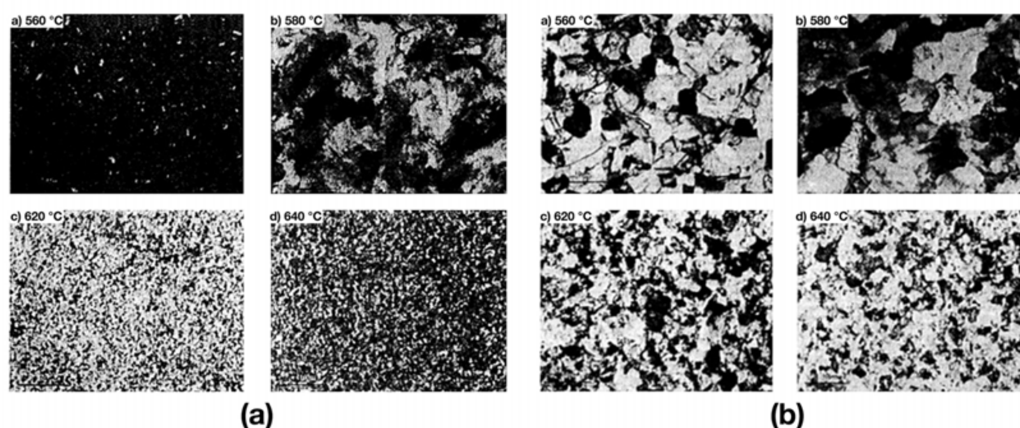


Figure 25. Transmission electron micrographs of as deposited and annealed polysilicon thin films; (a) as-deposited films, deposition temperature clockwise from the upper left: - 560°C, 580°C, 620°C and 640°C; (b) the same films after annealing in H_2 at 980°C for 30 minutes (reproduced by permission of The Electrochemical Society [64]).

mks

Polysilicon thin film deposition processes can be performed in batch mode, using a horizontal or vertical hot-wall LPCVD reactor, by treating one substrate at a time using a single wafer processor (SWP) or as a continuous process in an atmospheric pressure (APCVD) reactor. Basic schematics for some of approaches to these different reactor configurations are shown in Figure 26 [65]. Each of these reactor configurations has its own unique characteristics for the deposition of polysilicon and other thin films and these characteristics will be further discussed in chapters dealing with CVD equipment. In the particular case of polysilicon deposition processes, the key impacts of the process are the film uniformity over the substrate surface and the number of particles that are deposited on the growing film during the process. The film uniformity is determined by geometric and thermal factors unique to each CVD method. Batch mode LPCVD systems must be designed to balance the expected interplay between process gas flows, process pressure, wafer spacing and localized surface-to-volume ratios within the reactions zone. All of these parameters are varied in the development of proprietary processes aimed at achieving the ideal thin film properties for a given user's polysilicon application. Single wafer deposition tools are somewhat more limited in that local surface-to-volume ratios are fixed in these systems and wafer spacing is not a variable parameter. However, the reduced volume and more even temperature control possible in single wafer systems makes it inherently easier to develop polysilicon deposition processes with good film uniformity and tightly controlled physical properties. Single wafer systems can be designed to operate either at low pressure or at atmospheric pressure. The primary advantage offered by continuous APCVD systems rests in the fact that the process can be made continuous, with the concomitant increase in throughput that this allows. Polysilicon deposition rates in SWP and APCVD systems are necessarily higher than those in LPCVD systems and this can produce differences in annealing behavior for films produced in the different systems. SWP systems have gained dominance for most semiconductor CVD process requirements over the past couple of decades owing to the better control and greater production flexibility that these systems can offer. They are almost exclusively used to perform the processing steps required in advanced device fabrication.

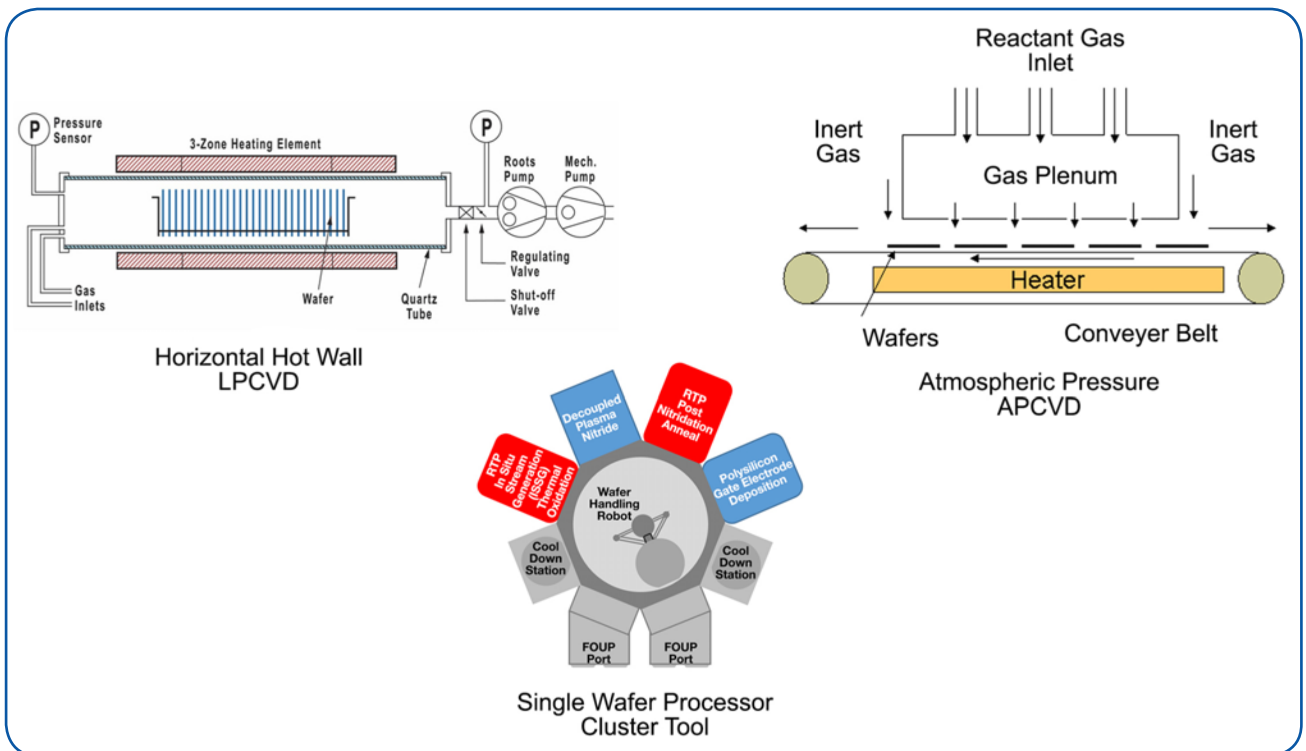


Figure 26. Some of the different reactor configuration used for the deposition of polysilicon thin films [65].

2. Epitaxial Thin Films

a. Silicon Based Epitaxial Thin Films

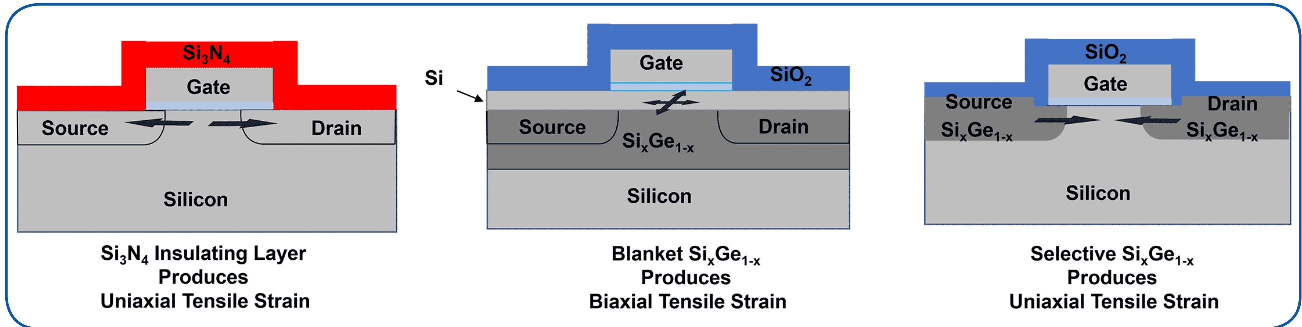


Figure 27. Biaxial and uniaxial strain techniques [66].

Epitaxy is defined as the “regularly oriented growth of one crystalline substance on another” [67]. Both homoepitaxy (growth of an epitaxial layer of the same material as the substrate) and heteroepitaxy (growth of an epitaxial layer composed of material that is different than the substrate) processes are used in semiconductor device fabrication. Heteroepitaxial films can be formed with either strained or relaxed lattice structures, depending on interfacial conditions and lattice parameters (Figure 27, Ref. [68]). The strain in the film can be either biaxial or uniaxial, depending on how the different materials are arranged in a device (Figure 28). Epitaxial films can be formed using a number of different methodologies, including evaporation, sputtering, molecular beam epitaxy, liquid phase epitaxy, and chemical vapor epitaxy.

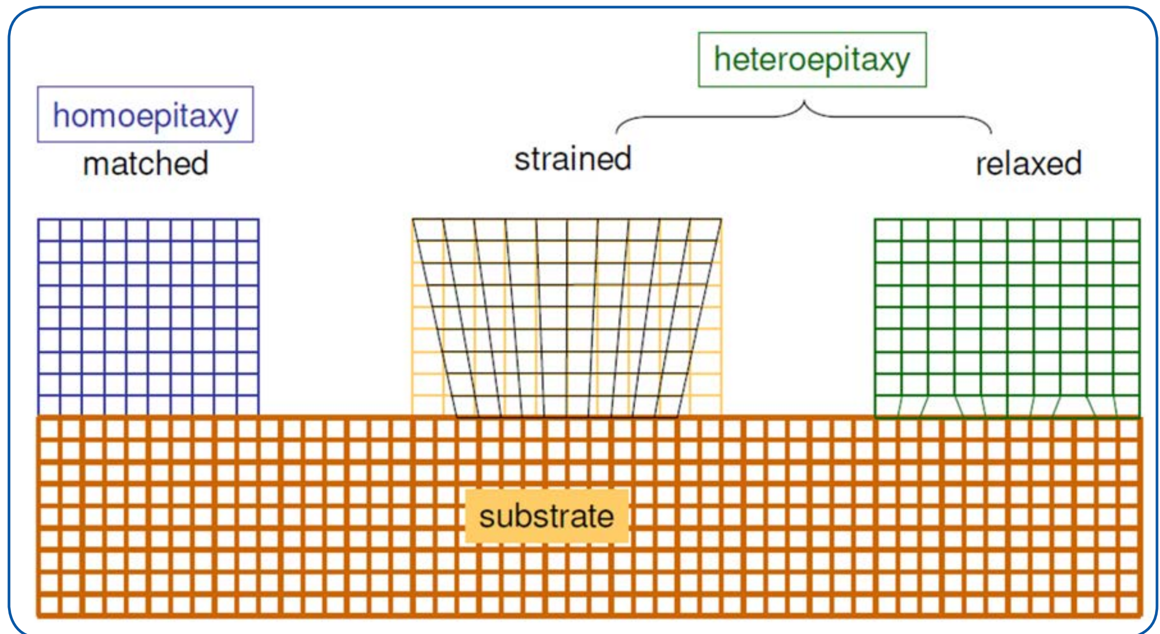


Figure 28. Epitaxial film growth modes.

Epitaxial materials are inherently freer of oxygen and carbon impurities than are bulk crystalline silicon wafers. They provide two unique advantages for device fabrication that cannot be achieved by other means: 1) they offer a direct means for creating controlled and abrupt changes in dopant profiles; and 2) they allow the creation of strained crystalline layers that give the device designer some ability to control electron mobility by a technique known as band gap engineering [69]. Leveraging these advantages has led to the use of silicon epitaxial layers for a variety of technological requirements in semiconductor devices.

Controlled dopant profiles and abrupt dopant profile changes offer unique advantages for device design. In bipolar devices the use of epitaxial layers to control dopant profiles enables higher switching speeds, improves high voltage operation and linearity, lowers base resistance and simplifies isolation schemes. In MOS devices, the use of epitaxial layers can lower diffused-line capacitance and improve diffused-line charge retention, improve resistance to alpha particle and static charge damage and improve dynamic Random Access Memory (RAM) performance. In addition, the use of epitaxial silicon layers in CMOS devices can improve latch-up protection. Hammond [67] notes that Vapor Phase Epitaxy (VPE) silicon makes possible many different doping profiles that are used in modern device fabrication. These include:

- *n*-type silicon over *p*-type silicon
- *p*-type silicon over *n*-type silicon
- Lightly doped over heavily doped layers – either type
- Lightly doped over heavily doped buried layer patterns
- Conducting silicon on insulator
- Silicon layers with controlled profiles

CMOS designs that employ a lightly doped layer of epitaxial silicon over a heavily doped substrate achieve higher breakdown voltages while simultaneously keeping collector resistance low. CMOS transistor performance has also been significantly improved through the exploitation of strain-induced band-structure modification [70] that directly impacts carrier mobility in targeted device layers. This band engineering approach leverages the enhanced electron mobility that is observed in epitaxial Si layers that are under biaxial tensile strain [71] as well as the enhanced hole mobility that is produced in epitaxial $Si_{1-x}Ge_x$ layers under biaxial compressive strain [72]. A more intuitive way to view the influence of lattice stress on electrical conduction is to consider the impact of stress on electrons and holes. Electrons travel through the space that divides the atoms in the crystal lattice in pseudo ballistic trajectories until they encounter an electric field. Therefore, as they move through a crystal lattice, electrons are slowed by interactions with the positive electric fields surrounding each silicon atom in the lattice. When the lattice is subjected to tensile stress, the distance between atoms in the lattice is increased and electrons have longer ballistic trajectories between interactions with lattice atoms and, hence, higher mobility. Hole mobility has the opposite relationship with stress in the crystal lattice. Hole movement in the lattice is caused by an electron moving from the outer shell of a silicon atom to the outer shell of a neighboring silicon atom that has a positive charge (the hole). Basically, the hole is filled in creating a hole in the atom that donated the electron. Compressive stress reduces the distance between neighboring silicon atoms making this transfer of electron to hole easier and, in that way, increasing hole mobility. Strained epitaxial layers that are created through the use of juxtaposed Si and $Si_{1-x}Ge_x$ epitaxial layers are used in the source and drain regions in advanced CMOS transistors [73]. Channel layers in these devices also use epitaxial strained layers.

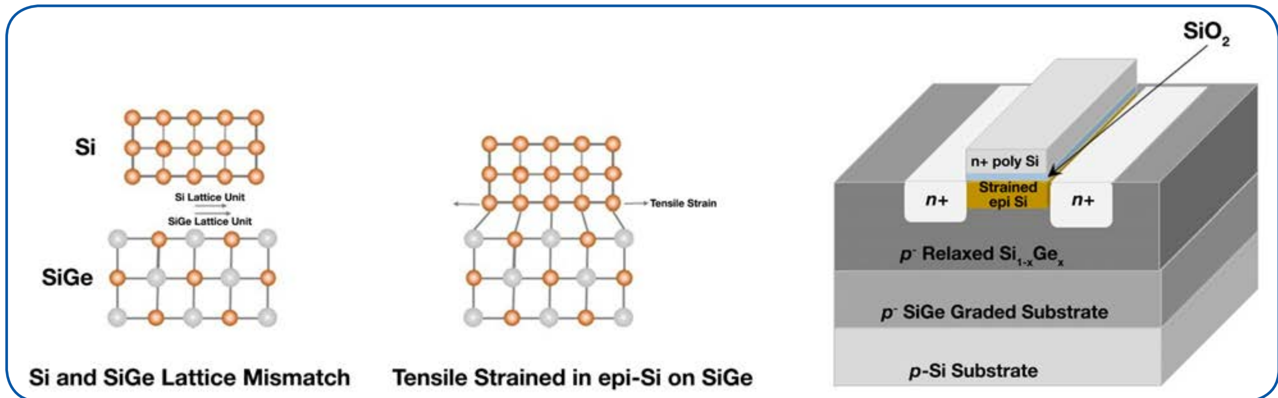


Figure 29. The origin of tensile strain in epitaxial Si/SiGe and its use in devices [65].

Intel was the first to use strained layer technology in high volume manufacturing of their 90 nm CMOS technology (see Figure 27). Figure 29 shows how epitaxial and other layers can be used to achieve uniaxial and biaxial strain characteristics in Intel's 90 nm CMOS transistor structure [66]. A number of recent reviews are available on the characterization and use of strained epitaxial layers in advanced device structures [74] [75] [76]. Strained silicon epitaxial layers may also find use as enabling technology in advanced photonics applications [77]. Table 6 provides some specifications for silicon epitaxial layers in selected device applications (from Hammond).

CVD epitaxy, also known as VPE, is the technology of choice for both homoepitaxial and heteroepitaxial processes for semiconductor device fabrication. Figure 30 shows a schematic representation of a typical VPE deposition process (for an As-doped n-epitaxial Si layer) on a wafer surface [77]. Typically, silicon epitaxial processes are carried out at atmospheric pressure using a high flow of hydrogen carrier gas. $Si_{1-x}Ge_x$ epitaxial processes differ from silicon epitaxy in that they are much more sensitive to ambient contaminants such as oxygen. Special care must be taken in these processes to employ point-of-use ultra-purification of the feed gases and/or to process under ultra-high vacuum (UHV) conditions. In addition, $Si_{1-x}Ge_x$ processes are performed at significantly lower temperatures than silicon epitaxy. Recently, silicon carbide epitaxial layers have become of interest for device makers. These are higher temperature processes, running 300°C – 400°C hotter than silicon epitaxy processes.

Device Type	Thickness (microns)	Resistivity (ohm-cm)
Bipolar Discrete Devices		
High Frequency	0.5-3.0	0.15-1.5
Power	5-100+	0.5-100+
Bipolar Integrated Circuits		
Digital Memory	0.5-5	0.3-1.5
Microprocessor	0.5-5	0.3-1.5
Linear	3-15	2-20
MOS-on-epitaxy ICs		
P/P+	4-20	10-40
N/N+	0.5-7.0	1-10
Bipolar/MOS	0.5-3.0	0.5-3.0

Table 6. Silicon Epitaxial layers, applications and specifications [67].

The chemical reactions used for the production of undoped and doped epitaxial silicon and $Si_{1-x}Ge_x$ include:

- $SiCl_4 + 2H_2 \rightarrow Si + 4HCl$ (~1200°C)
- $SiHCl_3 + H_2 \rightarrow Si + 3HCl$ (~1150°C)
- $SiH_2Cl_2 \rightarrow Si + 2HCl$ (~1100°C)
- $SiH_4 \rightarrow Si + 2H_2$ (~1050°C)
- $GeH_4 \rightarrow Ge + 2H_2$
- $B_2H_6 \rightarrow 2B + 3H_2$
- $PH_3 \rightarrow P + 3/2H_2$
- $AsH_3 \rightarrow As + 3/2\ H_2$

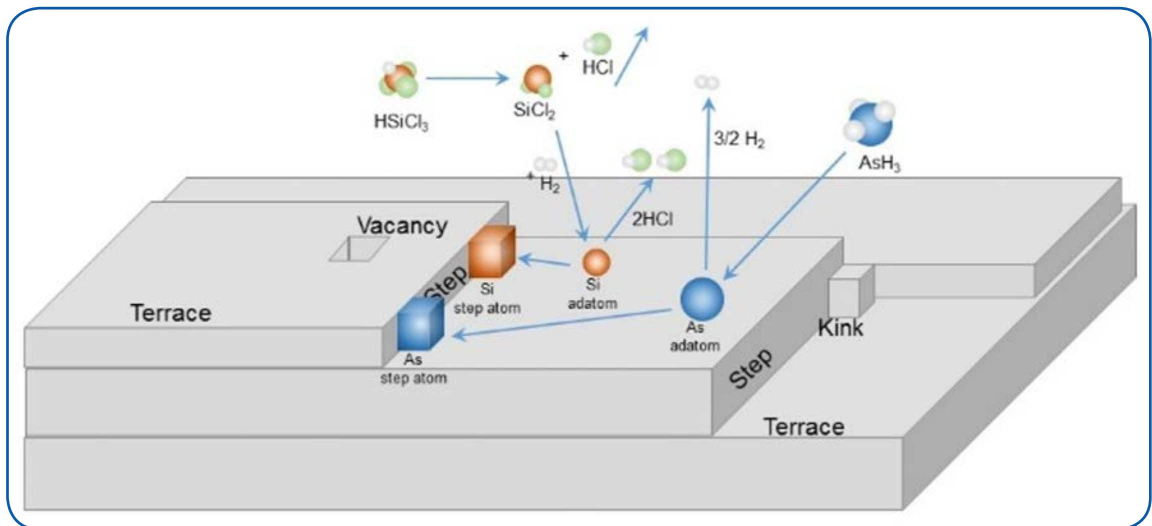


Figure 30. Schematic representations of the mechanism of epitaxial silicon layer formation [77].

Conventional VPE homoepitaxial processes are roughly divided into three categories: high temperature epi (1150°C – 1200°C); low temperature epi (1050°C – 1100°C); and selective epi. The heteroepitaxy processes most often encountered in device fabrication are $Si_{1-x}Ge_x$ on Si, and III-V epitaxy. Reactors may operate at atmospheric pressure or under reduced pressures down to UHV. In addition, reactors may be batch-type systems (horizontal, vertical pancake and barrel reactors) that are almost exclusively used for high temperature epitaxial processes; rapid thermal processing reactors that can be employed for both high temperature (HT) and low temperature (LT) epitaxial processes; and UHV-CVD systems (developed by IBM, especially for $Si_{1-x}Ge_x$). In addition to the industry standard CVD system, physical vapor deposition is employed in molecular beam epitaxy (MBE) systems [78]. It requires an ultrahigh vacuum environment (~10^{-10} Torr) and tightly controlled beams of evaporated atoms which impinge on the substrate to form the epitaxial layer. Semiconductor and dopant sources are arrayed around the substrate and shutters control the exposure of the substrate to each elemental species. Sources can be solids or gases. MBE methods are the most versatile and precise methods available for the formation of epitaxial layers. However, the exacting process requirements and the difficulty of achieving significant substrate throughput in the process have meant that this method has been largely confined to research applications and a few very advanced optoelectronic applications.

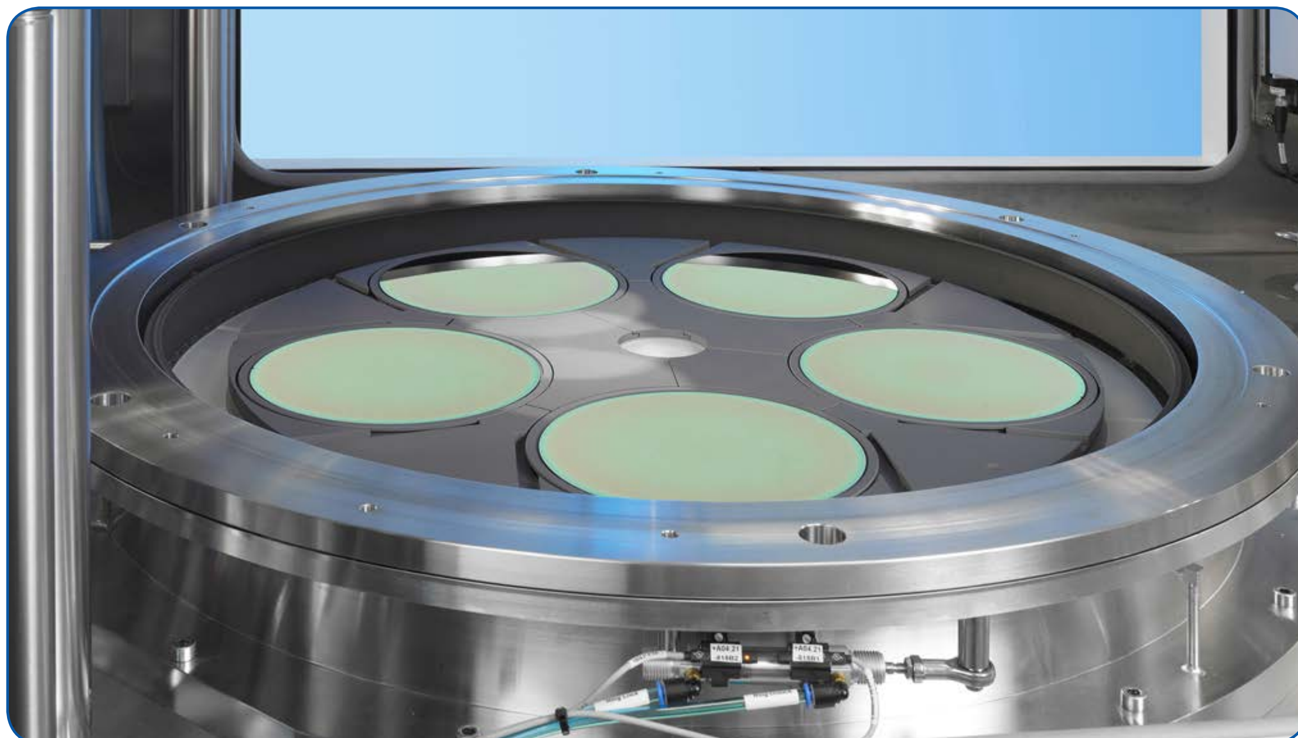


Figure 31. AIXTRON's G5+ Epitaxial Reactor [65].

Most modern epi processes are performed in deposition tools similar to the one shown in Figure 31. Such process tools incorporate both clean and deposit steps and allow the user to deposit sequential homo- and/or heteroepitaxial layers with customized surface preparation and minimal risk of interfacial contamination. The condition of the surface of a substrate strongly influences the perfection and hence the electrical properties of an epitaxial film. Thus the ability to control the interfacial chemical and structural properties is a necessary attribute of epitaxial process tools. Epitaxial processes require a starting substrate free of native oxide to ensure crystallinity in the epitaxial layer and to maintain high surface mobility of the adsorbed reactant species. Substrates must be pre-cleaned to remove contaminants and any native oxide on the surface and be maintained in this cleaned state until the epitaxial process has been completed. Historically, silicon homoepitaxy employed an HCl/H_2 mixture as an in situ clean of the silicon substrate surface prior to the epitaxy process. Modern homo- and heteroepitaxy processes that employ chemical/mechanical polished substrates no longer require this step, and, instead, use a high temperature H_2 bake to remove any native oxide from the substrate surface.

A high co-flow of hydrogen carrier gas during the epitaxial layer deposition process helps to keep the concentration of elemental reactants on the substrate surface low enough to promote ordered crystal growth; it also helps to avoid particle-producing gas phase reactions. Silicon homoepitaxial processes that employ chlorosilane precursors and temperatures in excess of 1000°C have growth rates that range between 0.2 and 2.0 μm/min, depending on the process conditions. Process temperatures in both homo- and heteroepitaxial processes are high enough to promote surface mobility of the adsorbed silicon, germanium and dopant atoms so that they move freely on the substrate surface until they become locked into an ordered crystal lattice site, as shown in Figure 30. In addition, high temperatures in epitaxial processes promote desorption of the reaction by-products from the substrate surface (i.e., HCl, H_2) which, in turn, maintains high deposition rates and surface mobility.

The mechanistics of both homo- and heteroepitaxial process can be roughly broken down into the following steps, illustrated in Figure 30:

- transport of the reactants to the deposition region
- transport of the reactants by diffusion from the main gas stream, through the boundary layer next to the substrate surface
- adsorption of the reactants on the substrate surface
- surface processes: migration, decomposition, reaction and site incorporation
- desorption of by-products from the surface
- transport of by-products through the boundary layer
- transport of by-products out of the deposition region

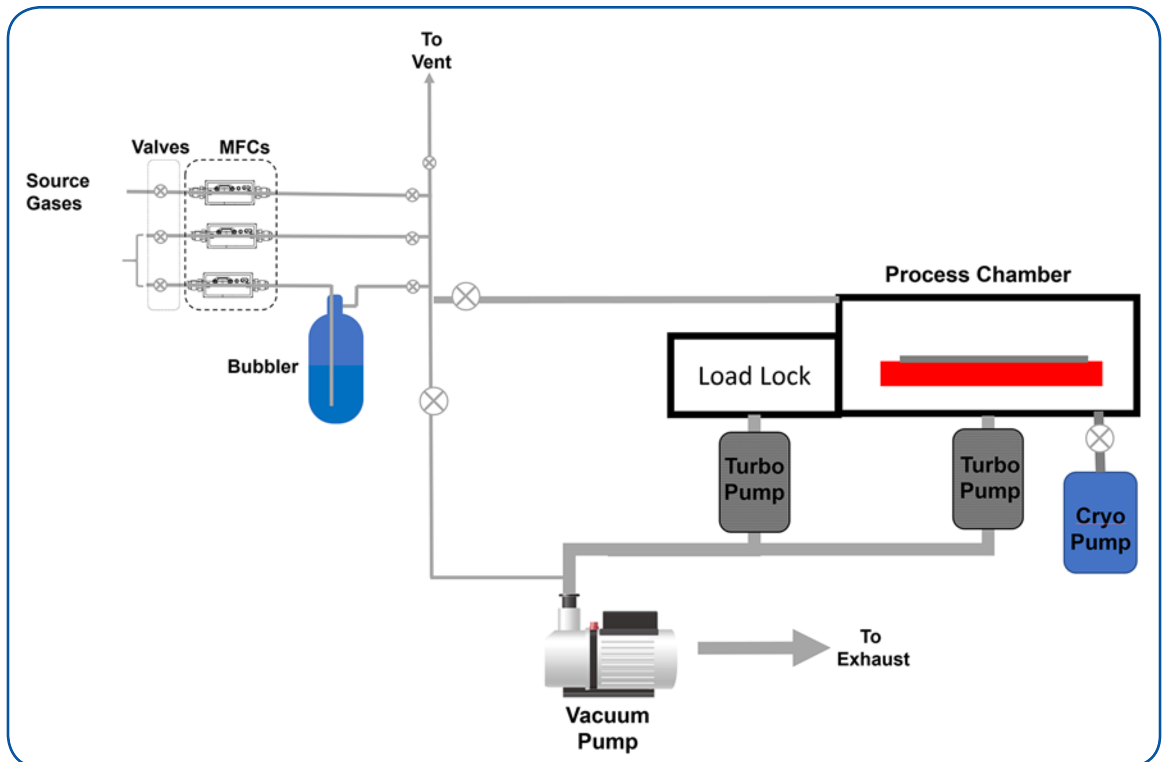


Figure 32. Schematic of a UHV-CVD Epitaxial Reactor [79].

The development of epitaxial processes for $Si/Si_{1-x}Ge_x$ hetero- and quantum well structures requires low temperature processing to avoid diffusion and intermixing at the Si- $Si_{1-x}Ge_x$ interfaces. The silicon and germanium sources used in $Si_{1-x}Ge_x$ epitaxy are normally silane, SiH_4, and germane, GeH_4; compounds that begin to decompose at process temperatures of 500°C (some $Si_{1-x}Ge_x$ processes use chlorosilanes and require somewhat higher temperatures). The assumption that extremely low vapor pressures of water and oxygen in the gas phase are required in Si CVD epitaxy processes [80] [81] stimulated the development of the UHV-CVD epi processes in the late 1980s. It was found that the minimum temperature, in general, for epitaxial processes was lower for UHV-CVD [81] with films deposited at temperatures less than 500°C.

However, the relatively high cost associated with the technique tends to restrict its application to those areas with no other option. $Si_{1-x}Ge_x$ processes in UHV-CVD reactors can be configured to operate as batch processes and Figure 32 shows a schematic view of such a UHV-CVD system [79]. A turbomolecular pump combined with a roots blower and a mechanical pump in series are used to achieve UHV conditions, typically with base pressures of $<10^{-9}$ Torr. As with conventional epitaxial processes, hydrogen is used as a carrier gas, silane and germane are the reactive gases and B_2H_6 and PH_3 are mixed into the gas flow for *p*- and *n*-type doping of the epitaxial layers. UHV-CVD epitaxial $Si_{1-x}Ge_x$ films are grown at deposition temperatures ranging from 450°C – 520°C. At higher deposition temperatures the hydrogen passivation of the substrate desorbs and a clean wafer surface cannot be maintained. High temperature bakes such as those in conventional silicon epitaxial processes are not possible in UHV systems due to the limits of the system components. With wet chemical cleaning, including an HF dip for surface H-passivation, and rapid transfer to the deposition chamber, the $Si/Si_{1-x}Ge_x$ interface typically shows oxygen and carbon contamination in the range of 10^{10} to 10^{12} atoms/cm^2. The silicon surface atom density is 7×10^{14} atoms/cm^2, which means that less than one atom in a thousand at the interface is a contaminant.

b. Compound Semiconductor Epitaxial Thin Films

IIB	IIIA	IVA	VA	VIA
	Aluminum 13 **Al** 26.982	Silicon 14 **Si** 28.086	Phosphorus 15 **P** 30.974	Sulfur 16 **S** 32.065
Zinc 30 **Zn** 65.38	Gallium 31 **Ga** 69.723	Germanium 32 **Ge** 72.64	Arsenic 33 **As** 74.922	Selenium 34 **Se** 78.96
Cadmium 48 **Cd** 112.41	Indium 49 **In** 114.82	Tin 50 **Sn** 118.71	Antimony 51 **Sb** 121.76	Tellurium 52 **Te** 127.6
Mercury 80 **Hg** 200.59	Thallium 81 **Tl** 204.38	Lead 82 **Pb** 207.2	Bismuth 83 **Bi** 208.98	Polonium 84 **Po** 209

Figure 33. Source elements for compound semiconductors - e.g. II-VI and III-V.

Compound semiconductors [82] are materials such as GaAs, AlGaAs, GaN, CdSe, InP, InSb, etc. These materials combine the elements from equidistant columns on either side of Column IVA of the Periodic Table, to produce materials having the general formula $A^N B^{8-N}$, where the superscripts refer to the periodic group containing the element (i.e., II-VI, III-V, etc.). III-V materials, especially GaN and GaAs, along with silicon carbide, SiC, are the dominant commercial compound semiconductors and have the longest history of device usage. GaAs-based microelectronic devices, for instance, exhibit much lower signal noise than do microelectronic devices employing other types of semiconductors. This is a very important attribute in weak signal amplification. In addition, they are much more resistant to radiation-induced damage than are silicon devices. GaAs has a direct optical band gap of around 1.4 eV [83] which allows it to be used for thin film solar cells as well as for light emitting diodes. The term “direct band gap” implies that a material emits a photon of light through a single, direct electronic transition between the valence and conduction bands of the material (Figure 34). “Indirect” band gap materials such as silicon require that an electron pass through an intermediate state during such transitions, prohibiting direct photon emission.

Device technology employing GaAs and InP substrates is reasonably well developed and these substrates are currently used in ways analogous to silicon for the fabrication of advanced electronic devices. Microelectronic device fabrication on GaAs substrates has a long history in high speed and

radiation hardened military and space applications and, more recently, in commercial opto-electronic devices. InP substrates have been employed for the fabrication of optical network components such as Dense Wavelength Division Multiplexing (DWDM) lasers [84], Vertical Cavity Surface Emitting Lasers (VCSELs) [85], pump lasers [86], etc. as well as for high speed electronics circuits. GaN substrates have become available and are finding use in optical device manufacture due to the fact that GaN alone offers light emission in the blue region of the spectrum (needed for white light LEDs). GaN substrates are also being used for the manufacture of power devices. Finally, silicon carbide, SiC, is being used for substrates for more efficient and compact power electronics in motor controls, power supplies and inverters [87].

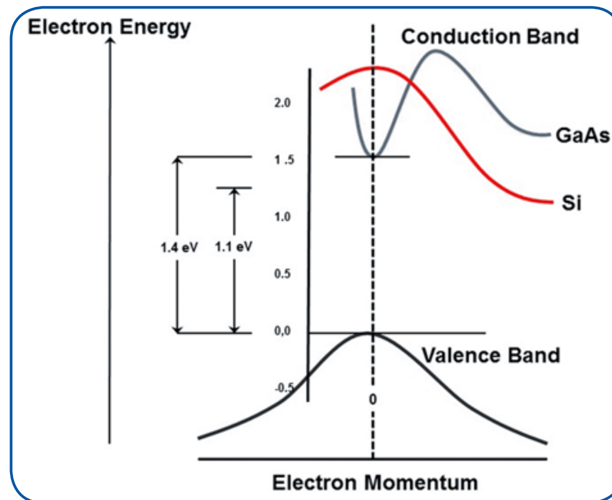


Figure 34. Energy band structure of Si vs GaAs [88].

Compound semiconductor substrates suffer from certain material properties when used in device fabrication. Heat dissipation in compound semiconductors is hampered by their much lower thermal conductivity as compared with silicon. This is a significant problem for the use of these substrates in applications with even normal power dissipation requirements. Another drawback of compound semiconductors, when compared with silicon, is that it is impossible to grow a well-behaved, strongly insulating thermal oxide on these binary and ternary compounds. Thermal oxidation of GaAs or InP (and others) results in preferential oxidation of one of the elements over the other and in the production of an oxide with poor material and electrical properties. This issue has probably been the single most important factor in limiting the use of compound semiconductor substrates since the early days of integrated circuit manufacture. Indeed, in the latter part of the 20th Century there was a standing joke within the industry that GaAs was and would always be the “material of the future” for just this reason. This being said, the advent of smaller feature size and advances in deposition methods such as Atomic Layer Deposition (ALD) has made the issues of native oxide removal and insulating oxide layer formation more easily addressable in compound semiconductor device fabrication, opening many opportunities for this technology. Epitaxial film deposition is a critical technology for leveraging the advantageous electrical properties of compound semiconductors in advanced devices and significant effort has been applied to the development of processes and equipment for this purpose.

Epitaxial methods for the deposition of thin films of compound semiconductor materials have made possible the synthesis of artificial crystal structures such as quantum wells and superlattices that have opened myriad possibilities for the use of compound semiconductors in microelectronic devices. In the past two decades, these new structures have found many applications that exploit their superior electrical and optical properties in advanced electronics, photovoltaic, optical and optoelectronic devices. Modern multiple quantum well lasers, for example, contain graded index separate confinement heterostructures (GRINSCH) [89] or VCSEL’s [85] that emit light at very low threshold currents at room temperature under continuous wave operation.

The quality of epitaxial layers of compound semiconductors, like that of silicon epitaxial films, is critically dependent on the crystallographic characteristics of the substrate on which the layer is grown and on interfacial relationships between the substrate and epitaxial layer. As with silicon epitaxy, substrate surfaces must be clean and free of all contaminants, including any native oxide. Surface cleaning is a much more complex problem with compound semiconductors than it is with silicon, due primarily to the binary, ternary or quaternary chemistry of these surfaces. Many different approaches have been, and continue to be, evaluated for cleaning. To date, there is no universal approach for such cleans, with successful cleaning procedures being very much tied to the chemical make-up of the particular compound semiconductor and to the interfacial characteristics required in a given application. Bromine/methanol, ammonium hydroxide and hydrochloric acid based wet chemistry methods, analogous to the ammonium hydroxide and hydrofluoric acid chemistries used in silicon processing, have been used to clean GaAs surfaces for some applications [82]. Thermal annealing methods have also proven successful in removing native oxides from compound semiconductor surfaces [82]. Hydrogen passivated surfaces have been produced on GaAs and other compound semiconductors by methods such as atomic hydrogen cleaning utilizing remote-plasma-generated atomic hydrogen [90]. The crystal lattice quality and orientation of the substrate also plays a critical role in the properties of epitaxial compound semiconductors since lattice mismatching between substrate and film leads to strain in both crystal lattices. Those readers seeking in-depth discussion of the issues surrounding interfacial effects in compound semiconductor devices are referred to reference [82].

A number of approaches have been (and continue to be) developed for producing epitaxial layers of compound semiconductors. Methods such as Liquid Phase Epitaxy (LPE), UHV-CVD, VPE, and MBE have all been successfully employed for compound semiconductor epitaxy [91], however, the most successful approach for industrial applications has proven to be Metal Organic Chemical Vapor Deposition, MOCVD. MOCVD is being used broadly for epitaxial layer deposition in both conventional, atmospheric pressure CVD reactors and in advanced ALD reactor configurations. The interested reader is encouraged to obtain and review the available comprehensive treatments on the chemistry and deposition technology of MOCVD for epitaxial compound semiconductor films and other applications available in the literature [92] [93].

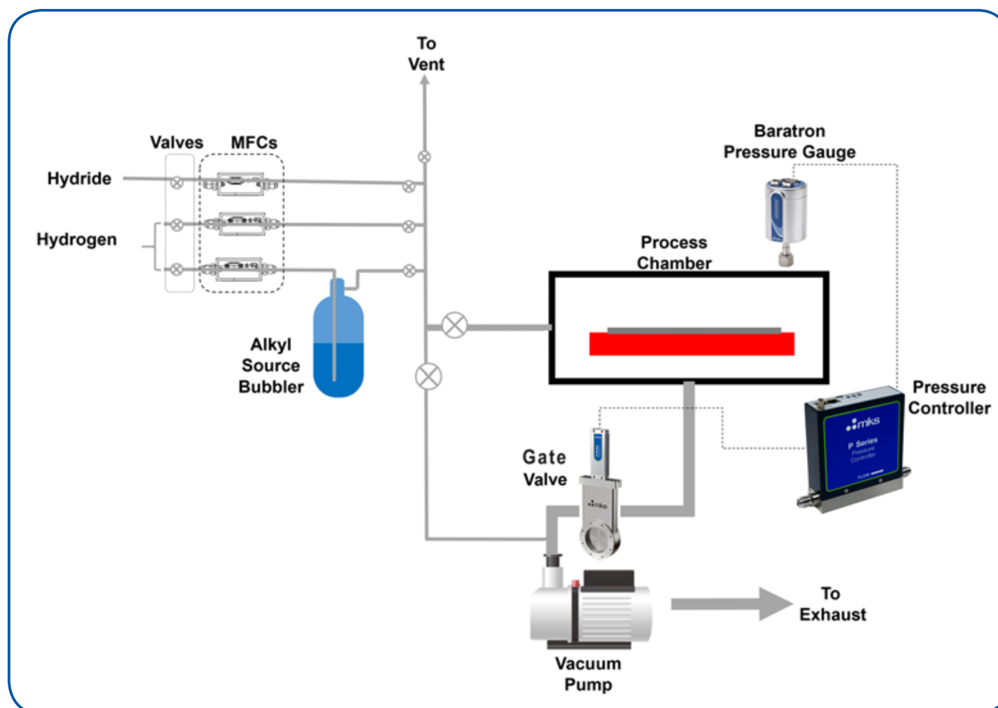
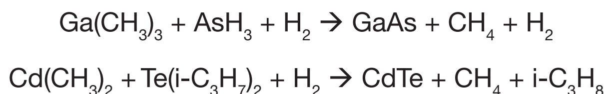


Figure 35. Schematic of a simple MOCVD reactor system [94].

The primary advantage of MOCVD for compound semiconductor epitaxy is the fact that most of the metal constituents in these materials are available as readily transportable, commercially available, high purity organometallic compounds. Organometallics are chemical compounds that incorporate both a central metal atom and attached organic groups. They are all relatively volatile and thus they can be used to carry relevant metals, especially those in Groups II and III in the periodic table, into a reaction chamber where they decompose on the hot substrate surface, depositing the desired epitaxial film. Volatile hydrides of these metals, analogous to the common precursors for silicon (SiH_4) germanium (GeH_4) and dopants (B_2H_6, PH_3) are either unknown or chemically unstable and thus unavailable for use in CVD epi processes. With the exception of gallium chloride, Ga_2Cl_6, halides of these metals (analogous to the chlorosilanes) are unusable in MOCVD, primarily due to the fact that they are non-volatile solids. As a result, metal organic compounds have proven the best precursors for epitaxial processes and they include compounds such as:

- trimethylgallium (TMG – $Ga(CH_3)_3$)
- triethylgallium (TEG – $Ga(C_2H_5)_3$)
- trimethylindium (TMI – $In(CH_3)_3$)
- trimethylaluminum (TMA – $Al(CH_3)_3$)
- trimethylantimony (TMSb – $Sb(CH_3)_3$)
- diethylselenium (DES – $Se(C_2H_5)_2$)
- dimethylcadmium (DMCd – $Cd(CH_3)_2$)

Examples of the process chemistry employed for MOCVD compound semiconductor epitaxial processes include:



$$Ga(CH_3)_3 + AsH_3 + H_2 \rightarrow GaAs + CH_4 + H_2$$

$$Cd(CH_3)_2 + Te(i\text{-}C_3H_7)_2 + H_2 \rightarrow CdTe + CH_4 + i\text{-}C_3H_8$$

As in Si and $Si_{1-x}Ge_x$ epitaxial processes, MOCVD compound semiconductor epitaxial processes are carried out using a high flow of hydrogen carrier gas to ensure good crystallinity in the epitaxial layer and, in some cases, to participate in the deposition chemistry. Reactor system for compound semiconductor epitaxy by MOCVD come in a variety of configurations. The basic components of a MOCVD reactor system are shown in Figure 35. The system is composed of subsystems that include reactant storage, a gas handling and control manifold, the reaction chamber and a pump exhaust system. Hydrogen carrier gas and gaseous metal hydride co-reactants feed systems for the MOCVD reaction chamber and are similar to those in conventional APCVD and LPCVD systems. Since most organometallic compounds are volatile liquids, bubbler systems that use hydrogen to entrain the organometallic compound in the vapor phase are required (“alkyl source” in Figure 35). Modern MOCVD epitaxy systems incorporate similar subsystems to those shown in Figure 35 and incorporate advanced designs for the deposition chamber that ensure high quality epitaxial layer growth and high throughput. AIXTRON, for example, offers batch systems that employ either a planetary susceptor (Figure 36(a)) or a close-coupled showerhead design as shown in Figure 36(b).

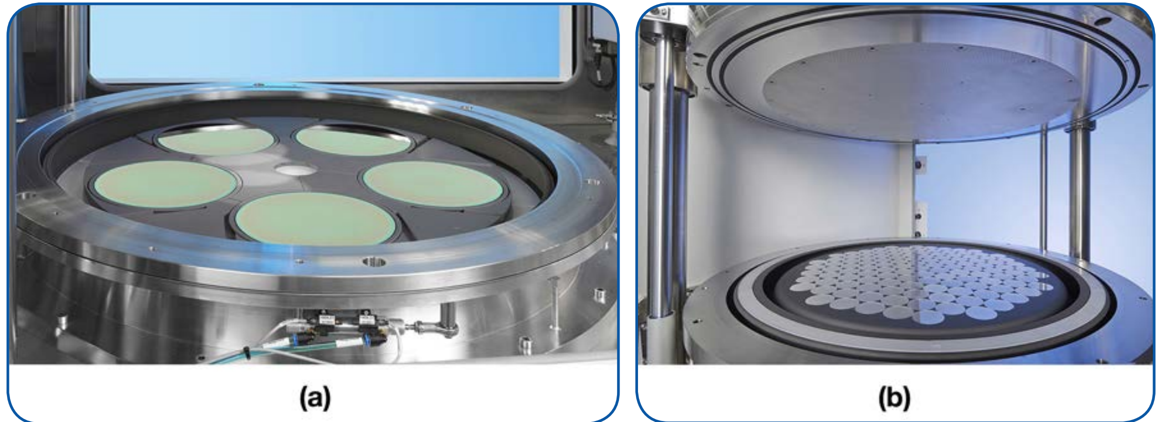


Figure 36. AIXTRON MOCVD epitaxial reactor systems, (a) planetary system; (b) close-coupled showerhead system [95].

3. Dielectric Thin Films

a. Silicon Dioxide

Insulating thin films of silicon dioxide SiO_2 are uniquely critical materials within silicon semiconductor device technology since they are major components in all MOS devices, as well as in other metal–insulator–semiconductor (MIS) and tunneling structures [96]. For instance: thermal grown oxides are employed as sacrificial protective layers and masks for substrate doping; thick thermal oxide films are used as isolating field oxides in devices with relatively large design rules (thermal oxide use in this application has been largely supplanted by deposited oxides and trench isolation in nanometer scale devices); and thin oxide films are used as gate oxides or gate oxide components in more advanced devices, albeit supplemented by high-k ALD deposited materials.

Thermal SiO_2 films for device components can be formed in a variety of ways; the selection of method is dependent on such factors as substrate composition/topography and the thermal tolerances of other device features on the substrate. Methods for producing SiO_2 films include thermal oxidation of silicon, PECVD, LPCVD, APCVD, MOCVD, and PVD. The chemistries range from simple oxygen-based reactions to the complex organometallic chemistry used in MOCVD. Of these methods, the thermal oxidation of silicon produces the highest quality insulators.

The availability of an adherent, impermeable, high dielectric strength thermal oxide is, arguably, the reason that silicon technology has dominated semiconductor device design and microelectronics manufacturing for the past 50 years. Other materials such as germanium and GaAs have superior electrical properties when compared with silicon; indeed, germanium was selected as the substrate in the first transistor design. However, neither of these materials nor any other likely candidates except silicon has a well-behaved insulating oxide that can be grown on the substrate by simple thermal oxidation in large batch processes. It is this single attribute that has allowed silicon technology to dominate the microelectronics world for so long.

DC Resistivity (Ω-cm, 25°C)	> 10^{20}	Melting Point (°C)	~1700
Density (g/cm³)	2.27	Molecular Weight	60.08
Dielectric Constant	3.8 – 3.9	Molecules/cm³	2.3 x 10^{22}
Dielectric Strength (V/cm)	5 – 10 x 10^6	Refractive Index	1.46
Energy Gap (eV)	~9	Specific Heat (J/g°C)	1.0
Etch Rate (BOE, Å/min)	1000	Stress in film on Si (dyne/cm²)	2 – 4 x 10^9
Linear Expansion Coefficient (cm/cm °C)	5.0 x 10^{-7}	Thermal Conductivity (W/cm °C)	0.014

Table 7. Selected physical constants of thermally grown silicon dioxide. Taken from "Silicon Processing for the VLSI Era", Stanley Wolf and Richard N. Tauber, Lattice Press, 1986 [32].

Table 7 is a table taken from Wolf and Tauber [32]. It contains selected physical properties of silicon dioxide that are relevant to its use in semiconductor device manufacture. It can be seen that thermal oxide has excellent electrical properties for use as an insulator in microelectronic devices. It exhibits very high resistivity, but, more importantly, it has a high breakdown electric field strength. The large energy gap provides good protection from tunneling, at least down to nanometer-scale thicknesses. Physically, thermally grown oxides are stable with reproducible Si/SiO_2 interfaces, an important factor in device fabrication at low nanometer dimensions. Finally, the oxide can be grown in a conformal manner over varied topography on the silicon surface. This means that insulating films produced by thermal oxidation will not exhibit thin spots at sharp edges and corners in the underlying topology, obviously important for good electrical insulation in a device. Silicon dioxide thin films are an amorphous form of silica consisting of SiO_4 tetrahedra linked together through all four oxygen atoms. The term amorphous means that the films do not have a well-defined crystal structure, having only short range order in their molecular components.

Silicon can be thermally oxidized using either dry oxygen or steam. Dry oxidation is typically used for films having thicknesses of less than 100 nm since growth occurs more slowly (and thus more controllably) in the dry process (typically 14 – 25 nm/hour). Wet oxidation uses steam as the oxidant and exhibits significantly higher growth rates (100 – 200 nm/hour). Wet oxidation is preferred for the formation of thicker films, up to 1000 nm. Dry oxidation, however, produces a higher quality oxide than wet oxidation since films formed using the latter process have significant concentrations of hydroxyl groups, OH^-, both in the bulk film and at the Si/SiO_2 interface. This degrades their electrical properties as compared to those of "dry" oxides. Note that advanced methods such as in situ steam generation (ISSG) at low pressures (< 20 Torr) resolve some of the purity issues of wet oxides and these processes have found use for ultra-thin gate oxidation and shallow trench isolation liners. Figure 37(a) provides a schematic that identifies the key physical and chemical aspects of the thermal oxidation of silicon while Figure 37(b) shows a typical equipment arrangement for the process (in this case ISSG).

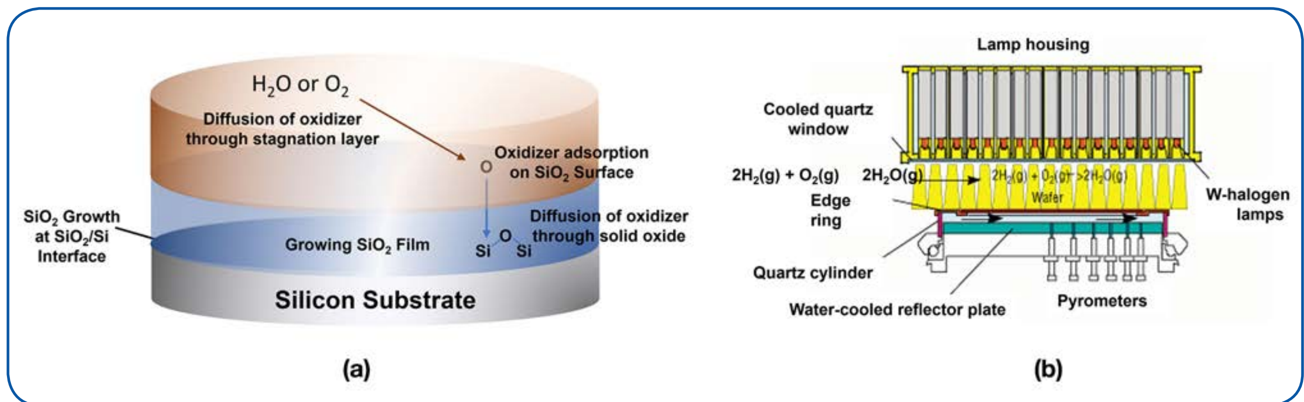
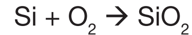


Figure 37. Thermal silicon oxidation: a) Process [97]; b) Equipment configuration (Reproduced with permission of Solid State Technology) [98].

Since the simplest oxidation chemistry is that associated with dry oxidation and since, in general, it produces the highest quality thermal oxide, we will use this process as the basis for a discussion of how silicon is oxidized and the characteristics of the film produced by the oxidation. The growth process of silicon oxidation has been extensively studied over the past 50 years. While the chemistry can be simply expressed:



$$Si + O_2 \rightarrow SiO_2$$

the kinetics of the process and the structure of the film and interface [99], [100], [101], [102] are far from simple. The oxidation process occurs at the Si/SiO_2 interface and the rate of the oxidation is determined by a complex relationship between the concentration of oxygen in the gas and on the surface of the growing film, the rate of diffusion of oxygen through the growing oxide film, and the chemical reaction between oxygen and silicon at the interface. Since the reaction consumes silicon, the interface continually moves into the silicon substrate during the oxidation process, with Si-Si bonds being broken and new Si-O bonds formed. This complicates the kinetics, since it causes the oxygen diffusion times and concentration of oxygen at the interface to vary with time. Also, the transition from the tightly ordered crystal lattice of silicon to the more open amorphous structure of silicon dioxide produces a volume expansion of the substrate/oxide film system as the oxide film grows. In the final structure, about 46% of the silicon dioxide that is grown is found within the bounds of the original silicon while 54% of the oxide is new volume. An understanding of all of these factors is critical in the device fabrication process since they must be quantitatively known or estimated for precision in gate oxide processing at the nanometer scale.

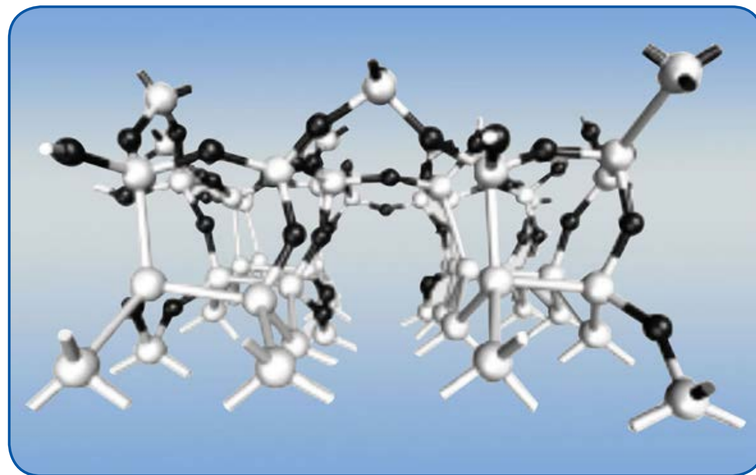


Figure 38. A model of the Si/SiO2 interface; grey balls represent silicon atoms, Black Balls represent oxyten atoms [103].

The nature of the interface between the silicon substrate and the silicon dioxide film (Figure 38) in thermally grown gate oxides is also a critical parameter since it strongly influences the operational characteristics of transistors at nanometer dimensions. The interface is not perfect at the atomic level, but rather a randomized transition from Si – Si to Si – O bonding. Interfacial zones that are as close to perfect at the atomic level are sought in the manufacturing process in order to achieve optimal device performance since the greater the perfection of the interface, the more predictable and repeatable are the electrical properties in devices. Many studies have been performed over the past few decades in an effort to understand and control the character of the interface and to limit the impact of imperfections on device performance ([100] [101] [102] are representative). Much work has gone into finding process and equipment approaches to improve this interface.

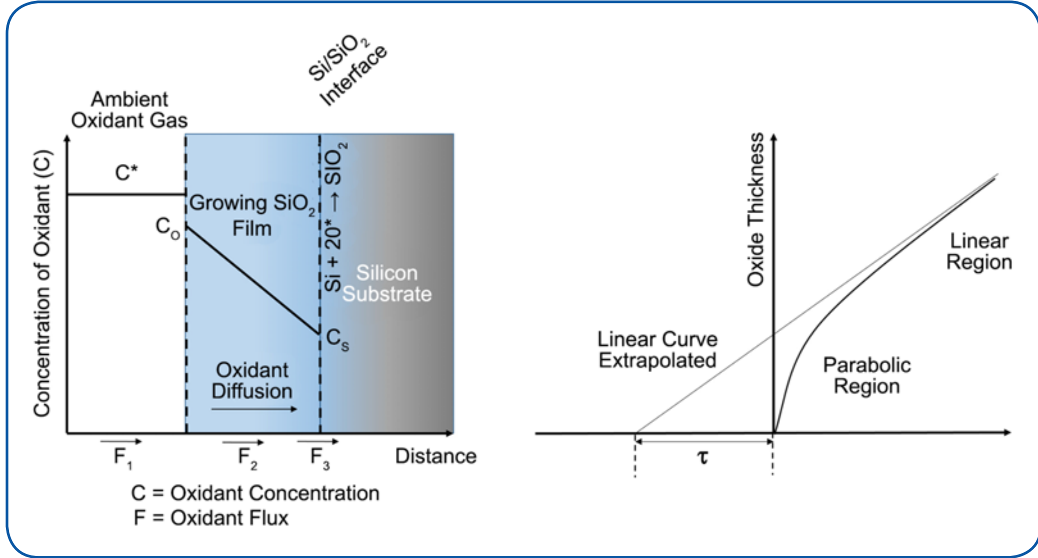


Figure 39. One-dimensional Deal-Grove model for the oxidation of silicon [104].

The growth process of thermal SiO_2 films has been extensively modeled and experimentally evaluated. Almost all models of the process begin with the model developed by Bruce Deal and Andy Grove in 1965 [105]. They proposed that the oxidation rate was determined by the combination of two processes. The first involved the actual chemical reaction of oxygen with silicon at the oxide/substrate interface, while the second was the diffusion of oxygen through the previously formed oxide film. The combination of these processes resulted in the formulation of the classical "linear-parabolic" rate law of Deal and Grove (Figure 39):

$$dx/_{dt} = F/_{N_1} = \left(kC^*/_{N_1}\right)\Big/\left(1 + k/_h + kh/_{D_{eff}}\right)$$

where x is the oxide thickness, F the total flux of oxidant molecules through the oxide, k the first-order rate constant for the oxidation reaction, C^* the concentration of oxidant at the oxide surface, N_1 the number of oxidant molecules incorporated into a unit volume of the oxide layer, h the gas phase transport coefficient of oxygen, and D_{eff} the effective diffusion coefficient of oxygen in silicon dioxide. When the differential equation is solved, the above equation can be rewritten as

$$x^2 + Ax = B(t + \tau)$$

where

$$A = 2D_{eff}\left(1/_k + 1/_h\right)$$

$$B = 2D_{eff}C^*/_{N_1}$$

and

$$\tau = (x_0^2 + Ax_0)/B.$$

The Deal–Grove model requires the parameter τ (a shift in the time coordinate) in order to account for the presence of an initial oxide layer on the silicon surface. This approach was found to provide an excellent fit to experimental data for oxidation processes utilizing H_2O/O_2 mixtures and for dry oxidation

processes at thicknesses in excess of approximately 40 nm [106]. Deal–Grove estimates fail, however, in predicting oxidation rates in dry oxygen within the thin regime (<20 nm) and in accounting for the observed pressure dependence of both thin and thick oxidations. Within the thin regime, observed oxidation rates in dry processes are significantly greater than Deal–Grove predictions and the reaction exhibits a power law dependence on the oxygen pressure with rate proportional to P^m where m = 0.6–0.8 [107] [108]. It is beyond the scope of this overview to detail the refinements to the Deal-Grove model that have occurred over the past few decades. In 1995, one author wrote in a publication on thermal oxidation kinetics "*An examination of the literature and available experimental evidence does not permit a definitive choice of model for the oxidation of silicon*" [109]. A review of the literature since that time suggests that, while our understanding has improved, the thermal oxidation kinetics remains a valid field of study. The reader is referred to some of the reports and reviews that are available in the literature, and the references therein [110] [111] [112] [113] [114] for more detailed information on this topic.

Deposited Silicon Dioxide Thin Films

Silicon dioxide thin films can be deposited under a variety of conditions and using a variety of chemistries. Over the past half century, silicon dioxide films have been deposited using reaction conditions that varied from atmospheric pressure through sub-atmospheric pressures (SACVD) to low, vacuum pressures (LPCVD and PECVD). The reactions that produced the oxide film have been both thermal and plasma-enhanced. The chemistries employed in the deposition processes have varied. In this section, we will describe a few of these routes to deposited silicon dioxide films and the properties of the films produced.

Silane-Based Oxides (Low Temperature Oxide, LTO)

For much of the past 50 years, the most widely used deposited silicon dioxide film processes have used the reaction of silane with oxygen or a milder oxidant such as nitrous oxide, N_2O. The primary advantage of hydride-based deposited silicon dioxide films is the low deposition temperature, compatible with aluminum metallization. The chemistry for the deposition processes are relatively simple:

$$SiH_4 + 2O_2 \rightarrow SiO_2 + 2H_2O$$

$$SiH_4 + 2N_2O \rightarrow SiO_2 + 2N_2 + 2H_2$$

The first of these processes is usually thermally driven while the second is normally plasma enhanced; both were originally batch processes, depositing oxide films on up to a hundred wafers at a time, however they have been adapted to single wafer processing environments. Process temperatures are in the 200 – 450°C range with PECVD reactions normally performed at the lowest temperatures. These processes have noteworthy safety risks since the precursor, silane, is a high pressure, pyrophoric gas. This means that, should the gas leak from its high pressure cylinder, it will spontaneously burst into flame or, under certain conditions, explode. There have been a number of lethal accidents in the semiconductor industry associated with silane and safety precautions are critical in its handling.

While the chemistry of hydride-based silicon dioxide deposition can be summarized by relatively simple reactions, the detailed chemistry that occurs in these systems is actually quite complex. Under the pressure and temperature conditions employed in all of the hydride-based oxide deposition processes (SACVD, LPCVD and PECVD) unstable reactive intermediate species formed in the gas phase make a significant contribution to the overall deposition process. Reactive intermediates are formed by reactions such as:

$$SiH_4 + O_2 \rightarrow SiH_2(OH)_2$$

$$SiH_4 + H_2O \rightarrow SiH_3OH + H_2$$

$$SiH_4 + N_2O \rightarrow SiH_3OH + N_2$$

mks

The formation of silanols, $SiH_x(OH)_{4-x}$, is chemically favored ($\Delta H_{f(SiH_2(OH)_2)}$ ~ -100 kcal/mol) in this system. They are highly reactive chemical species that further react to produce SiO_2 once adsorbed on the substrate surface. The presence of reactive silanol intermediates in these deposition processes has major consequences for both the equipment requirements and the materials properties of the oxide films produced. The equipment used in the deposition processes, especially for the LPCVD process, must be configured with precise surface area/volume relationships between the surface of the substrate, the spacing of substrates in batch processes, and the surfaces surrounding substrate edges. If these relationships are not maintained, the uniformity of the thin oxide film over the substrate surface can be seriously degraded. Additionally, silanols in the gas phase readily react with other silanols and with silane to produce gas phase particulates that are extremely damaging to device yields. As with film uniformity issues, control of gas phase nucleation in the reactions is strongly dependent on reactor geometry and surface-to-volume rations in the reaction chamber.

Deposition	Thermal	LPCVD/APCVD LTO (Densified)	LPCVD/APCVD TEOS (as deposited)	HTO	PECVD
Source	O_2	$SiH_4 + O_2$	TEOS + O_2	$SiCl_2H_2 + N_2O$	$SiH_4 + N_2O$
Temperature (°C)	900-1000	400-450	700	900	200-400
Composition	SiO_2	SiO_2(H)	SiO_2	SiO_2(Cl)	SiO_2(H)
Step Coverage	Conformal	Nonconformal	Conformal	Conformal	Nonconformal
Thermal Stability	Stable	Densifies	Stable	Loses Cl	Loses H
Dielectric Constant	3.9	4.1	4.3	-	6-9
Dielectric Strength (MV/cm)	10	6-7	5-6	-	7
Density (g/cm^3)	2.2	2.1	2.2	2.2	2.3
Stress (10^9 Dynes/cm^2)	3 Compressive	3 Tensile	1 Compressive	3 Compressive	3 Compressive to 3 Tensile

Table 8. Comparison of different silicon dioxide processes and properties.

The as-deposited materials properties of LTO (silane-based) silicon dioxide films deposited in these processes are distinctly inferior to those of thermal silicon dioxide films. Films of these materials have to be densified after deposition and, even following densification, significant quantities of Si-OH and Si-H remain in the film. The result of this contamination is a dielectric strength for deposited films that is significantly weaker than that of thermal oxides, typically 6-7 MV/cm (vs. approximately 10 MV/cm for thermal oxide). These films cannot, therefore, be used in applications requiring high electrical integrity such as gate oxides. Therefore, their material property limitations have historically restricted their use in device fabrication to the formation of insulating layers between the active layer on the substrate and the different levels of metal in the device. In nanometer-scale device structures, however, intermetal dielectric films require almost perfect conformality in the deposited films, a property that LTO films lack. The term conformality refers to the uniformity of coverage over sharp edges and deep trenches in the substrate topography. Figure 40 illustrates good vs. poor conformality characteristics. The poor conformality of hydride-based silicon dioxide processes has led to their being largely supplanted by TetraEthOxySilane (TEOS, sometimes also called TetraEthylOrthoSilicate) and TEOS/O_3 processes. Table 8 provides a comparison of some of the critical properties of LTO films compared with other silicon dioxide thin film sources.

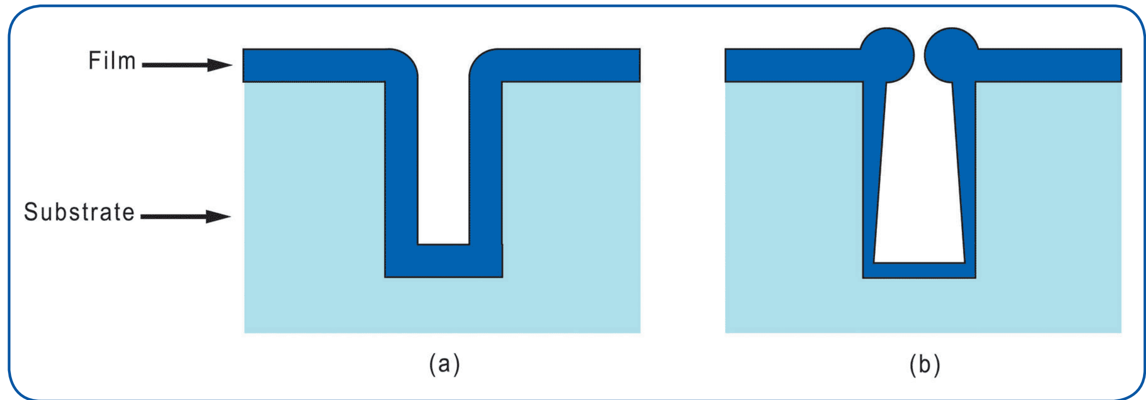


Figure 40. Deposited film conformality: a) conformal film; b) nonconformal film.

High Temperature Oxide (HTO)

High temperature silicon dioxide (HTO) films are formed by the reaction of dichlorosilane and nitrous oxide:

$$SiH_2Cl_2 + 2N_2O \rightarrow SiO_2 + 2N_2 + 2HCl.$$

HTO processes are carried out at high temperatures, typically around 900°C and under low pressure conditions, typically around 200 mTorr process pressure. The high process temperature results in an oxide that is very close to thermal oxide in many of its properties (Table 8). However, the high process temperature also limits the utility of HTO, since dopant migration and loss of metal line integrity become problems. In addition, the presence of chlorine as a contaminant in HTO can be a problem. For these reasons, HTO films do not have an extensive list of applications in device fabrication. However, the use of rapid thermal processing techniques with very short time at temperature has raised interest in the use of HTO as a replacement for Oxide-Nitride-Oxide (ONO) films as interpoly dielectrics in flash cells and gate dielectrics.

TEOS-Based Oxides

TEOS-based oxides have largely supplanted LTO as passivation and intermetal dielectric films in device manufacture [115] [116]. The TEOS molecule has the chemical formulation $Si(OC_2H_5)_4$. Formally, TEOS is an organometallic compound and TEOS-based deposition processes can be classed as MOCVD, although this is rarely noted as such.

Early use of TEOS-oxides employed LPCVD batch processing and simple thermal decomposition of the TEOS molecule [117]:

$$Si(OC_2H_5)_4 \xrightarrow{700°C} SiO_2 + 4C_2H_4 + 2H_2O.$$

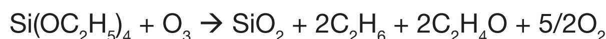
In practice, the process usually required a small amount of co-reactant oxygen for good process characteristics such as film uniformity and low particle counts. The LPCVD TEOS process produced a very stable silicon dioxide film with properties approaching those of thermal oxide (Table 8). More importantly, the films exhibited excellent conformality, able to uniformly cover relatively severe surface topographies to produce an excellent interpoly dielectric.

The process temperature requirements of LPCVD TEOS-oxide processes limited their use as intermetal dielectrics, so PECVD-based TEOS oxide processes were quickly developed to fill this need. PECVD TEOS-oxide processes exhibit high deposition rates at low deposition temperatures while simultaneously producing oxide films having excellent conformality. The low deposition temperatures

mks

(200°C-400°C) allow these processes to be used for the deposition of intermetal dielectrics. Unfortunately, the low deposition temperature can be problematic for film quality, yielding oxide films that are hygroscopic and containing significant structural and electrical defects and some process modifications and/or post-deposition processing may be required to raise the film quality to acceptable levels [118] [119] [120].

More recently, TEOS/O_3 processes have been developed that produce excellent silicon dioxide films for intermetal dielectric applications [121]. Originally, TEOS/O_3 was developed as LPCVD [122] and APCVD [123] processes; the most successful variant on TEOS/O_3 has been the subatmospheric pressure CVD (SACVD) TEOS/O_3 process [124]. SACVD TEOS/O_3 processes deposit highly conformal, high quality silicon dioxide films at process pressures ranging from 100 Torr to 600 Torr and temperatures as low as 300°C, allowing their use in all interlevel insulation applications. The process chemistry can be simply written as:

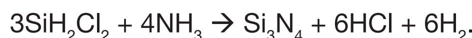


$$Si(OC_2H_5)_4 + O_3 \rightarrow SiO_2 + 2C_2H_6 + 2C_2H_4O + 5/2O_2$$

however, it is actually much more complex [125] [126].

b. Silicon Nitride

Silicon nitride, Si_3N_4, films can be deposited using either LPCVD or PECVD techniques. The process chemistry is simply written as:



$$3SiH_2Cl_2 + 4NH_3 \rightarrow Si_3N_4 + 6HCl + 6H_2.$$

In practice, HCl is undesirable in the process effluent since it can cause severe corrosion in the downstream components and therefore, a sufficient excess of ammonia is added to the process to ensure that the reaction chemistry is:

$$3SiH_2Cl_2 + 10NH_3 \rightarrow Si_3N_4 + 6NH_4Cl + 6H_2.$$

The presence of NH_4Cl in the process effluent can also be detrimental to the maintenance requirements of the deposition system (although not as bad as HCl). LPCVD silicon nitride processes require high temperatures, typically in excess of 700°C, while PECVD nitride can be deposited at temperatures below 400°C. The usual trade-off between process temperature and film quality exists for these two processes, with the LPCVD process producing a high quality, stoichiometric film and the PECVD process producing a non-stoichiometric film containing varying concentrations of silicon-hydrogen bonds.

Silicon nitride films are amorphous, hard materials that find two primary uses in device manufacture: masking layers and final passivation. Masking layers are typically deposited using LPCVD since this produces the most impervious film. Silicon nitride masking layers are especially useful for thermal oxidation processes since oxygen does not easily diffuse through the nitride. Silicon nitride has a number of desirable qualities as a passivation layer. PECVD methods allow it to be deposited at temperatures compatible with the underlying device structures. The film is nearly impervious to critical ambient contaminants such as moisture and sodium. Finally, by adjusting the PECVD process conditions the inherent stress in the film can also be adjusted to eliminate any risk of film delamination or cracking. As will be discussed later, silicon nitride finds use as a stop for chemical mechanical polishing processes. Table 9 provides a list of representative properties of LPCVD and PECVD silicon nitride thin films.

Property	LPCVD Si_3N_4	PECVD Si_3N_4
Composition	Si_3N_4	$Si_xN_yH_z$
Density (g/cm^3)	2.8 – 3.1	2.5 – 2.8
Dielectric Constant	6 – 7	6 – 9
Dielectric Strength (MV/cm)	10	0.6
Bulk Resistivity (Ω-cm)	10^{15} - 10^{17}	10^{15}
Stress ($dyne/cm^2$)	1.8×10^{10} (tensile)	8×10^9 (compressive)

Table 9. Representative properties of LPCVD vs. PECVD silicon nitride [32].

c. High-k and Low-k Dielectrics

The dielectric constant (k) of a material is a measure of the polarizability of that material. In a dielectric, polarization can have three origins. An external electric field can induce electronic polarization in which the electrons in neutral atoms are displaced from a symmetrical arrangement around the nucleus. It can also induce atomic or ionic polarization in which adjacent positive and negative ions are distorted from their equilibrium positions in a crystal lattice. Finally, it may cause dipolar or orientational polarization in which permanent dipoles in a material align with the external electric field. Low dielectric constant (low-k) materials exhibit weak polarization when subjected to an applied electric field; high-k materials are strongly polarized in the presence of an external electric field.

Over the past two decades, materials having dielectric constants (k-values) that are both higher and lower than those of conventional silicon dioxide and silicon nitride have become increasingly important in device manufacturing. Low-k materials are important for the reduction of Resistance-Capacitance (RC) delays in interconnects. High-k materials are important for use in storage capacitors and nonvolatile memory devices, and have become a critical component that allows continued scaling in the ultra-thin film stacks employed as gate insulators in nanometer scale devices.

Low-k materials, when used as insulators between the multilevel interconnections in a device, have the ability to improve device speed through reductions in RC delays. RC effects occur due to the unintended formation of capacitor structures when two metal interconnects are separated by an insulating layer. When this occurs, there is a delay in the transmission time of interconnect signals that is proportional to the time it takes to charge the incidental (parasitic) capacitor. This time, denoted by the symbol τ, is equal to the resistance of the interconnect times the capacitance of the incidental capacitor. This capacitance is, in turn, directly proportional to the dielectric constant of the insulating material and the area of the incidental capacitor and inversely proportional to the distance between the metal lines. Thus, as the distance between metal lines in multilevel metal structures is reduced, the delay in signal transmission increases unless the dielectric constant of the material separating the metal lines can be reduced.

Historically, the insulator used to separate metal lines in a device (the inter-level dielectric, ILD) has been SiO_2 which has a dielectric constant of about 3.9. By way of comparison, the lowest possible dielectric constant is 1.0 (vacuum or air). Materials for ILD applications have been specifically designed for the reduction of their dielectric constant. In general, the design principals focus on the reduction of any chemical characteristics that can result in moisture absorption and in designing the molecular chemistry to reduce polarization strength and density. This latter characteristic can be accomplished by reducing the Si-O bond density in the material and through the incorporation of fluorine (SiOF dielectric constant = 3.7) and organic species such as CH_3 (SiOC(H) dielectric constant = 2.8).

Dielectric Material	Dielectric Constant
Silicon Dioxide SiO_2	3.9
Silicon Nitride Si_3N_4	7-8
Aluminum Oxide Al_2O_3	8-10
Hafnium Oxide HfO_2	25
Titanium Oxide TiO_2	30-80
Zirconium Oxide ZrO_2	25
Tantalum Pentoxide Ta_2O_5	25-50
Barium-strontium-titanate BST	100-800
Strontium-titanate-oxide STO	230+
Lead-zirconium-titanate PZT	400-1500

Table 10. High dielectric constant materials [127].

Many materials with dielectric constants lying between the limits of air and SiO_2 have been evaluated as ILDs over the past few decades. CVD methods have been employed to deposit carbon-doped silicon oxide and organosilicate glass, OSG. These methods have produced ILDs with dielectric constants as low as 2.5, albeit with decreasing mechanical strength in the films. Spin-on hydrocarbon and fluorocarbon polymers have been evaluated, including polyimides, polyarylethers (PAE), polynorbornene and polytetrafluoroethylene (PTFE). These materials typically exhibit dielectric constants around 2.5. However, they tend to have weak mechanical strength, poor thermal stability and poor adhesion. Mixed organic/inorganic silicon based polymers having very open lattice structures, collectively known as polysilsequioxanes (SSQ), have also been evaluated. Most recently, the very low dielectric constant of air (1.0) has been leveraged for ILD dielectric constant reduction through both direct use [128] and the use of highly porous materials for ILDs. Air gaps have also been used in both logic interconnects and DRAM devices. The latter materials include porous silicate glass, xerogel/aerogel, porous organosilicate glass, porous SSQ, porous organics and porous SiLK (a trademarked porous organic polymer). These materials are normally applied using a spin-on sol-gel process followed by heat treatment to remove the solvent.

High-k gate dielectric materials, mainly transmission metal oxides, have proven crucial for successful device operation at device feature sizes of 65 nm and below. At these dimensions, the gate oxide in Field Effect Transistors, if fabricated from SiO_2, must have a thickness of 0.7 nm. This is much less than the direct electron tunneling thickness for SiO_2 films of 3.0 nm. This means that electrons can "tunnel" through such thin oxides to produce gate leakage currents of about 100 A/cm^2 at 1V. Obviously, such large leakage currents effectively destroy the transistor functionality. The use of high-k materials as part of the transistor gate structure avoids this problem since the high degree of polarizability of these materials allows low electrical potentials on the gate to more strongly influence the underlying channel. Table 10 provides dielectric constant values for some high-k materials of interest to the semiconductor community.

d. Other Dielectrics

Silicon Oxynitride, SiO_xN_y

Silicon oxynitride was one of the first high-k materials to be explored in the quest to avoid excessive leakage through ultra-thin gate insulators. SiO_xN_y has the additional benefit of improved barrier performance for boron penetration in surface-channel P-MOSFETs that have P^+ gate electrode doping. Silicon oxynitride is a compromise in that silicon nitride provides better material characteristics; however, the deposition process for silicon nitride produces a high density of traps in the film, making it unsuitable for use as a gate dielectric. SiO_xN_y films can exhibit dielectric constants up to around 5.0, as compared with silicon dioxide values of 3.9. Currently, SiO_xN_y materials are of interest as gate insulators for conventional device fabrication [129] [130] and as components having tunable refractive indices for photonics application [131].

Silicon oxynitride films can be formed on the substrate using a number of simple, accessible chemical approaches:

- Thermal nitridation of silicon followed by re-oxidation
- Annealing of SiO_2 films in nitrogen-containing ambient such as nitrous oxide (N_2O), nitric oxide (NO_2) or ammonia (NH_3)
- Growth of SiO_2 on nitrogen implanted substrates
- Low temperature remote plasma nitridation of SiO_2
- ALD nitride deposited on SiO_2
- PECVD, CVD

Using electron cyclotron resonance (ECR) remote plasma reactors, SiO_xN_y can be deposited at low temperatures and with minimal damage to the gate insulator-substrate interface with the substrate. ECR PECVD approaches allow for relatively precise control of oxygen to nitrogen ratios, low S-H concentrations, good uniformity and microstructure control in the SiO_xN_y film. Approach (e) has received recent attention as a preferred route to SiO_xN_y formation.

Aluminum Oxide, Al_2O_3

Aluminum oxide is another material of potential interest as a gate insulator, especially for thin film transistors (TFTs) on glass and flexible substrates, owing to its high dielectric constant (~9), relatively low leakage current and good thermal/electrical stability [132] [133]. To be of use in modern device fabrication, and especially for applications on flexible substrates, Al_2O_3 films must be deposited at quite low temperatures. Conventional PVD and CVD deposition approaches for Al_2O_3 at low temperatures have not produced films with acceptable properties for device applications. The films exhibit high leakage and poor stability due to the grainy growth and rough, polycrystalline surfaces. ALD methods, however, have proven to produce excellent films for TFT gate insulator applications. ALD processes for Al_2O_3 film deposition can be performed at temperatures as low as 120°C, making them ideal for the production of TFTs on flexible substrates and as a blocking dielectric in memory devices. Al_2O_3 is also used in non-volatile memory applications such as 3D-NAND.

4. Metal Thin Films for Contacts and Interconnects

a. Metal Materials Requirements

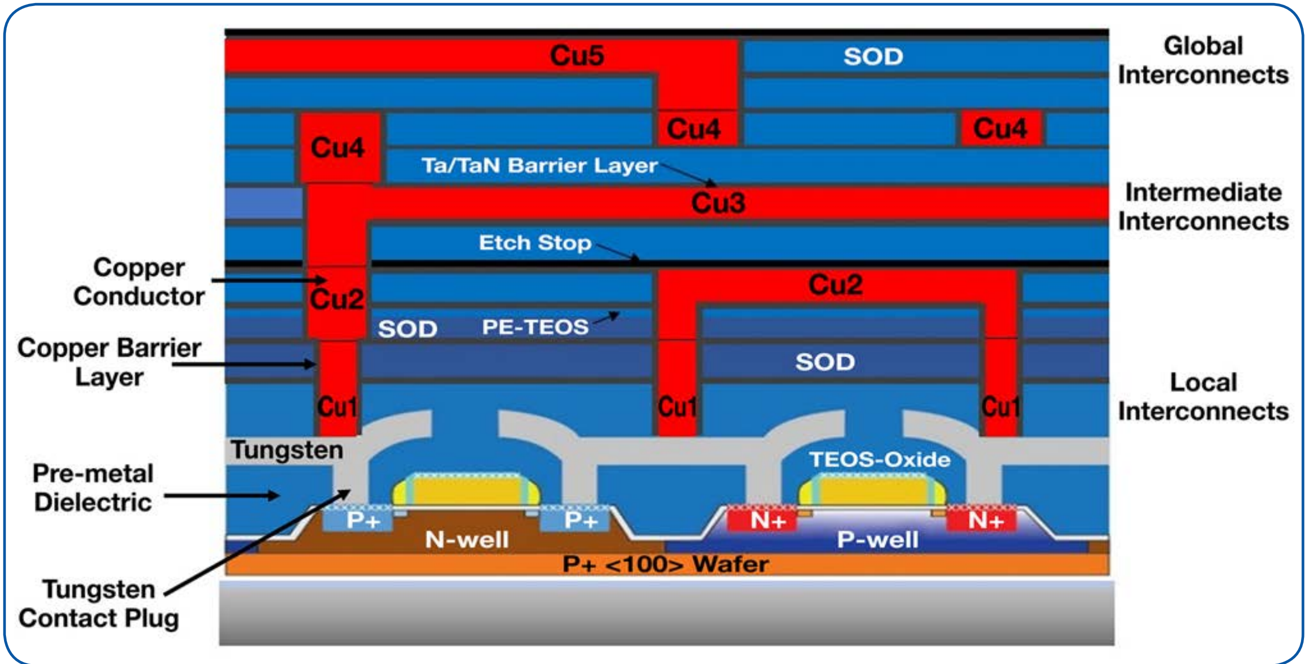


Figure 41. Metal applications within an older planar IC. Multilevel metallization terminology [134].

In semiconductor device fabrication, contacts are the metal components in direct contact with silicon in transistors and other devices, while interconnects are the current-carrying lines that connect discrete devices within an Integrated Circuit (IC). Figure 41 shows a cross-section of a (somewhat older) device depicting a multi-level metallization scheme. It can be seen that, within an IC structure, metals are used as gate and electrode contacts (Tungsten Silicide, light region above the source and drain), contact plugs (Tungsten Plug) and interconnects (Copper 1 through Copper 5). Diffusion limiting barrier layers (i.e. Ta and TaN) are required to prevent metal diffusion into semiconducting and insulating areas of the device.

It was noted in the section on low-k and high-k dielectrics that parasitic capacitances in metal interconnects are a cause of RC delays in signal transmission through the electrical connecting lines of an IC (the C component in RC). Of equal importance is the other parameter in this relationship, the resistance of the interconnect lines, R. For this reason, manufacturers have sought the lowest resistivity materials for interconnects in device fabrication. The electrical resistivities and conductivities of some conducting materials are shown in Table 11. It is important to note that the R component in RC delays is also reduced through the use of shorter interconnect lines. Line length reduction, along with the ability to interconnect a larger number of components per chip (packing density), have been the driving forces behind multilevel metallization in device designs. Figure 41 defines some of the terms common to this area.

Material	Resistivity (Ω-m)	Conductivity (/Ω-m)
Silver, Ag	1.59×10^{-8}	6.29×10^{7}
Copper, Cu	1.68×10^{-8}	5.95×10^{7}
Gold, Au	2.44×10^{-8}	4.10×10^{7}
Aluminum, Al	2.65×10^{-8}	5.81 x 10
Tungsten, W	5.6×10^{-8}	1.79×10^{7}
Iron, Fe	9.71×10^{-8}	1.03×10^{7}
Platinum, Pt	10.6×10^{-8}	0.94×10^{7}
Lead, Pb	22×10^{-8}	0.45×10^{7}
Manganese, Mn	48.2×10^{-8}	0.21×10^{7}
Constantan (Ni, Cu)	49×10^{-8}	0.20×10^{7}
Mercury, Hg	98×10^{-8}	010×10^{7}
Nichrome (Ni, Fe, Cr Alloy)	100×10^{-8}	0.10×10^{7}
Carbon (Graphite)	$3\text{-}60 \times 10^{-5}$	--

Table 11. Electrical resistivity and conductivity of metals at 298K [127].

The different metal functions within an IC have different materials property requirements. Gate and electrode contacts must provide a stable, ohmic interface between the metal conducting lines (or contact plugs) and silicon. These contacts must be able to endure the temperatures and other physical and chemical conditions experienced in post-deposition processing. Contact plugs must have highly conformal deposition processes that produce gap-free metal structures leading to a low resistance path between the electrode contacts and the interconnect network above the transistor device. Interconnects must provide a stable, high speed path for signal transfer between the device components that make up the IC.

Yves Pauleau, in a review some years ago [135], succinctly described the materials properties that are desirable in a metal interconnect material for use in sub-micron device fabrication:

- Low electrical resistivity
- Low resistance, ohmic contact structures
- Stable contact structures
- High electromigration resistance
- Adherence to insulating films (SiO_2)
- Stable with respect to insulating films (SiO_2)
- Low internal stress
- Low surface roughness
- Easily etched using plasma processes
- Fine line patternability
- Compatibility with all other semiconductor processes
- Low cost

Early on, heavily doped polysilicon and heavily doped regions in silicon substrates were used to form some of the conducting lines for integrated circuits. However, as device size shrank and speed requirements increased, the low resistivities required for contact and interconnect structures in these devices proved unachievable using doped silicon. For this reason, silicon interconnects and contacts will not be considered further in our discussion. Rather, we will discuss the elemental metal thin films now commonly employed for contact and interconnect formation. Of these, we will focus on aluminum, tungsten, and copper, along with the barrier layer materials that are normally used in conjunction with these metals. Our discussion of metal films will be conducted within the context of the material requirements listed by Pauleau.

Historically, aluminum was the primary metal employed for contact and interconnect metallization in IC device fabrication. However, aluminum has been supplanted in modern devices by a combination of tungsten and copper with accompanying barrier layers such as titanium nitride and Ta/TiN stacks. Figure 42 compares some of the differences between the aluminum metallization in earlier device designs with current approaches using tungsten contact plugs prior to first level metallization.

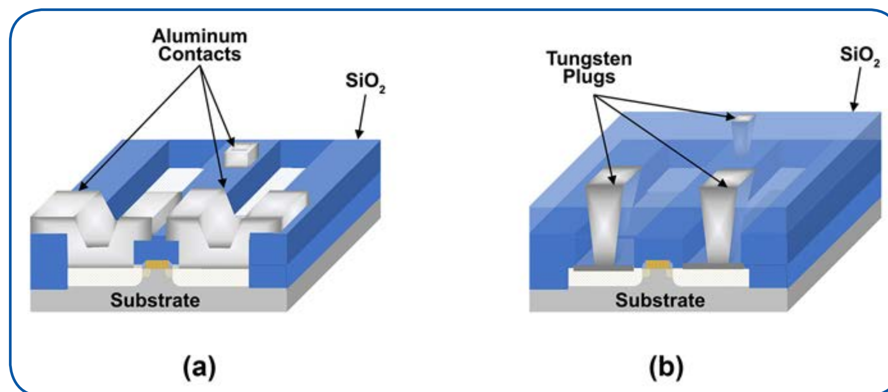


Figure 42. (a) aluminum metallization for device designs > 350 nm;
(b) first level metallization for designs < 350 nm [136].

b. Aluminum

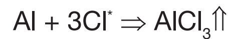
Aluminum has a number of advantageous properties which led to its widespread use in earlier device generations. Aluminum metal has the third lowest resistivity of the candidate metals for metallization (Table 11). It is easily deposited as a uniform, smooth, polycrystalline film over the entire substrate surface using evaporation or sputter deposition methods (PVD) [32]. While aluminum thin films are most often deposited using PVD techniques, it is also possible to deposit aluminum films by chemical vapor deposition using volatile organometallic precursors such as tri-isobutylaluminum, $Al(i\text{-}C_4H_9)_3$:

$$Al(i\text{-}C_4H_9)_3 \rightarrow Al + 3C_4H_8 + H_2$$

While the CVD process has become reasonably well developed over the past few decades, most aluminum thin films for semiconductor device fabrication continue to be deposited by PVD owing to the relative simplicity and low cost of these processes. In particular, magnetron sputtering is the most commonly used process for aluminum thin film deposition due to the high deposition rates (1µm/min) achievable in this process. A key difference between CVD and PVD films has been the fact that, historically, CVD films have produced far better step coverage in high aspect ratio contact structures (CVD: 50% -100% coverage with gap fill; PVD: 15% coverage with poor gap fill). Whether deposited by PVD or CVD techniques, aluminum forms a coherent, stable film that strongly adheres to both the silicon and dielectric layers in an IC.

Once formed, aluminum thin films are readily patterned to produce contact or interconnect structures using standard photolithographic methods and chloride-based Reactive Ion Etch (RIE):

$$CCl_4 + e^- \Leftrightarrow CCl_3^+ + Cl^* + 2e^-$$



$$Al + 3Cl^* \Rightarrow AlCl_3\Uparrow$$

The aluminum trichloride produced in the RIE reaction is volatile, with a boiling point of 177.8°C, and can be easily pumped away under vacuum. The corrosive nature of $AlCl_3$ introduces some unique maintenance problems in the RIE vacuum equipment and these will be discussed in more detail in Section B.

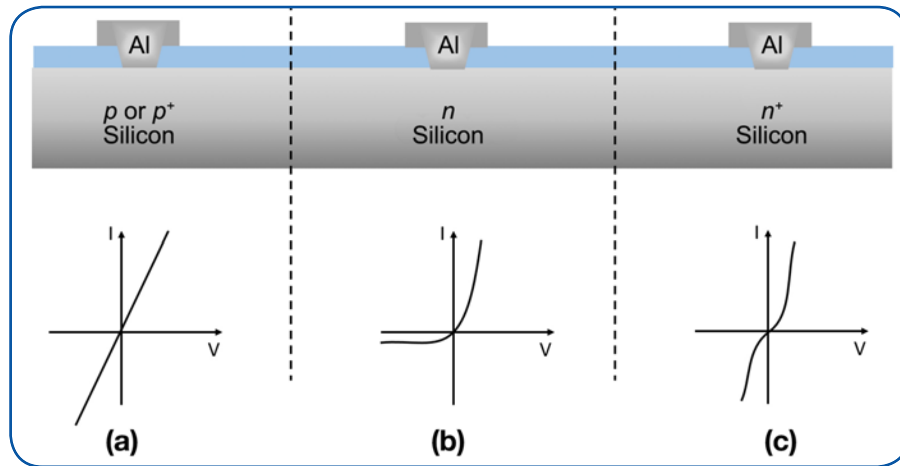


Figure 43. Aluminum contacts to silicon (a) ohmic contact;
(b) rectifying (Schottky) contact; (c) tunneling contact [137].

In early contact applications, aluminum was deposited directly onto silicon [135]. It was found that, depending on the doping of the silicon, such metal contacts exhibited different electrical properties. Figure 43 shows these different electrical characteristics. Aluminum (or any other metal) in direct contact with silicon can produce ohmic, rectifying or tunneling electrical characteristics. Ideally, an aluminum-silicon contact should have ohmic characteristics, however, this only occurs when the contact is to p or p^+ silicon since aluminum constitutes a p-type dopant for silicon. Aluminum contacts to n-type silicon form rectifying diodes, as described in Section A, Chapter I.2.b. For this reason, practical contacts between aluminum (or any metal) and n-type silicon employed a thin, n^+-doped region at the interface that creates a tunneling contact. While such contacts are, in fact, rectifying, when the width of the depletion layer at the metal-semiconductor interface is very thin (3 nm or less), carriers can easily tunnel across, producing the equivalent of an ohmic contact. Such barriers require n^+ doping densities of 10^{19} cm^{-3} or more.

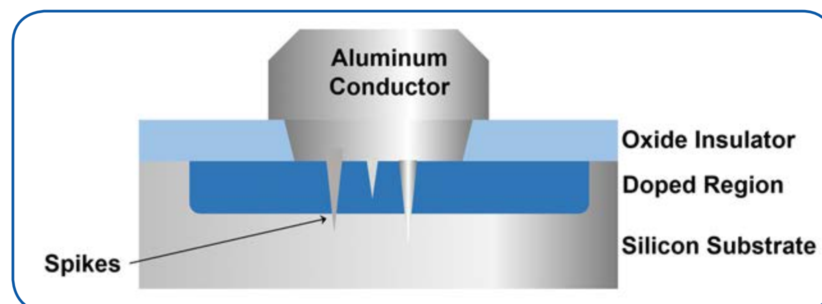


Figure 44. Pitting or spiking of aluminum through the p/n junction [138].

Once formed, aluminum contacts are normally annealed at high temperature (400°C) in a reducing atmosphere (forming gas; 20:1 N_2:H_2). This step significantly improves the contact resistance with silicon. Care must be taken that annealing temperatures remain below the Al/Si eutectic temperature since, at high enough temperatures, silicon can diffuse from the contact interface into aluminum to produce an Al/Si alloy (solubility of Si in Al at 400°C = 0.5%). When significant diffusion of silicon into the aluminum occurs, it creates voids in the silicon which are “backfilled” by the aluminum. This produces a phenomenon known as pitting or junction spiking (Figure 44). When spikes penetrate through the underlying p/n junction they can dramatically increase leakage current and/or reduce breakdown voltage. In most aluminum deposition processes, ~2% Si is added to the aluminum to prevent out-diffusion and spiking. This is another reason why PVD continues to dominate aluminum film deposition; CVD processes cannot effectively incorporate silicon into the growing aluminum film. Modern aluminum contact structures employ shallow junctions and barrier layers to eliminate rectifying behavior and prevent unintentional doping of the substrate and/or junction spiking due to metal/silicon alloy formation. In recent device generations, simple aluminum contact structures have been largely replaced by structures containing metal silicide electrodes, diffusion barrier layers and tungsten contact plugs.

Aluminum was the most common metal employed for interconnect metallization in CMOS processes above about 200 nm, however its use has been supplanted in later generations by more conductive and electromigration resistant copper metallization (see below). In addition to aluminum’s higher resistance relative to other metal candidates, electromigration has been the other major issue prompting a shift away from aluminum interconnect structures. Electromigration [138] [139] [140] is a phenomenon that arises from two factors: 1) the kinetic influence of electron-atom collisions in the interconnect line; 2) the field-induced movement of ionized aluminum atoms. Under the influence of these two factors metal ions move toward the positive end of the conductor while vacancies move toward the negative end. The ultimate effects of electromigration are the reduction of connectivity through the creation of voids in the interconnect lines and the creation of shorts through the production of hillocks and whiskers (Figure 45). While aluminum has been largely supplanted by copper in advanced devices, those remaining applications in which aluminum is used for interconnects typically employ aluminum-copper alloys to help reduce this problem.

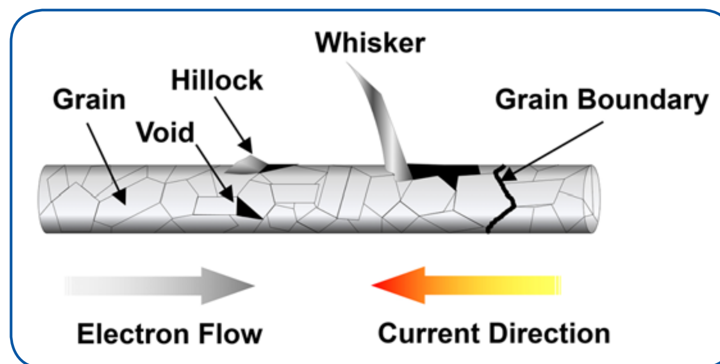


Figure 45. Electromigration, cause and effects [141].

c. Tungsten

CVD tungsten, despite having somewhat higher resistivity, has replaced aluminum in contact plug applications. The use of tungsten rather than aluminum is primarily due to the better gap fill characteristics of tungsten in high aspect ratio structures (Figure 46) making it ideal for 3D-NAND. Figure 47 depicts the evolution of metal contact structures over the past few decades.

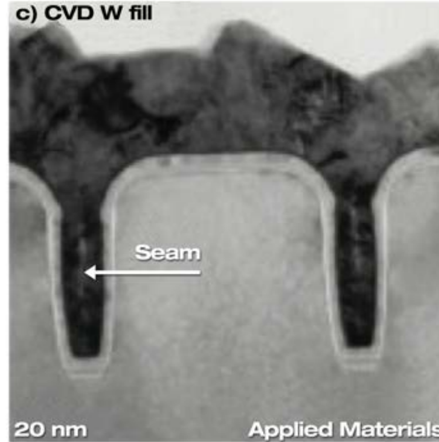
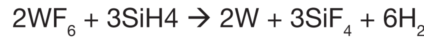
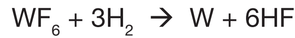


Figure 46. SEM of a tungsten contact plug showing excellent gap fill characteristics [142].

Because of its high resistivity, tungsten is employed only for plugs along with local and sometimes first level interconnects. Blanket tungsten thin films are deposited using one of the following chemical reactions:



$$2WF_6 + 3SiH4 \rightarrow 2W + 3SiF_4 + 6H_2$$



$$WF_6 + 3H_2 \rightarrow W + 6HF$$

$$2WF_6 + 3Si \rightarrow 2W + 3SiF_4.$$

Care must normally be taken to minimize the last reaction in the above series since it can lead to the same junction spiking problems as described earlier for aluminum. For this reason, contact plugs that are fabricated using CVD tungsten processes require the use of a barrier layer, typically TiN (see below), between tungsten and silicon. More advanced device structures with very high aspect ratio contact holes employ a PVD tungsten seed layer to ensure conformal coverage by the tungsten film.

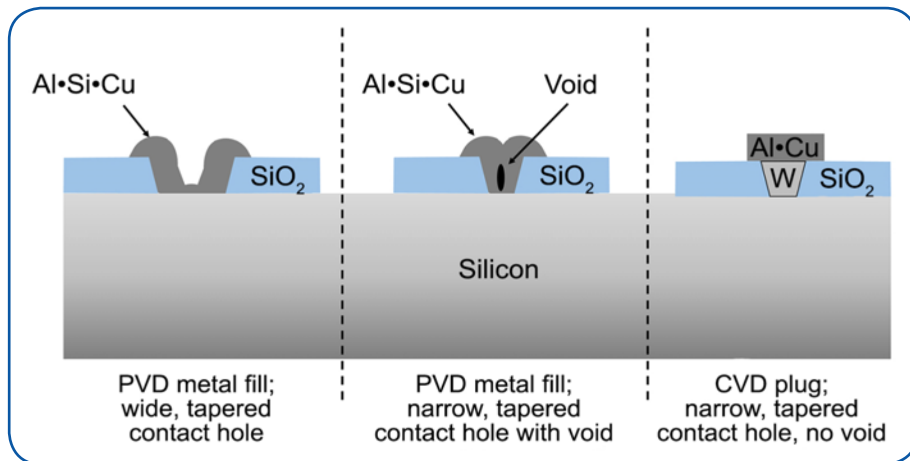


Figure 47. Evolution of metal/silicon contact structures.

mks

d. Copper

Copper has a much lower resistivity than aluminum and it has largely replaced aluminum in interconnect applications in device structures with design rules below about 200 nm [143]. The low electrical resistivity of copper wiring greatly reduces RC delays and lowers power consumption in nanometer scale devices. Low resistivity wiring has become critical as interconnect technology has become the determining factor in signal delays on an IC chip. Additionally, copper is much more resistant to electromigration than is aluminum.

The advent of copper as an interconnect material was initially hampered by the fact that it is a highly mobile metal and thus can produce unintended contamination, both in the fab and on-chip. Furthermore, the conventional approaches to thin film deposition and etch are not possible with copper, making it difficult to create patterned structures on a device surface. Any issues associated with unintended copper contamination within the fab were fairly easily resolved by isolating copper processing areas from other fab areas. On-chip contamination presented a somewhat more difficult problem. Since copper easily migrates into both the dielectric and semiconducting materials that surround it on a device surface, diffusion of the metal could only be prevented by the total encapsulation of copper structures by barrier materials. The need for patterning copper layers on substrate surfaces eventually necessitated the development of a completely unique approach for the formation of copper interconnect structures. Finally, copper, unlike aluminum, lacks a stable, adherent native oxide and this makes it more prone to corrosion and adhesion problems in integrated device schemes. Barrier (capping) layers have been developed that alleviate this problem.

Copper can be deposited using PVD, CVD or electrochemical plating technologies. Either a physical or chemical pre-deposition clean is needed prior to the deposition of a barrier layer (see below) followed by a seed layer for copper electroplating. Practical PVD processes for copper deposition do not produce usable thicknesses for interconnect formation; they are limited to the formation of copper seed layers that promote follow-on copper electrodeposition for interconnects. PVD processes have poor step coverage and gap-fill characteristics and this makes them unsuitable for copper deposition beyond this initial seed layer. CVD processes could conceivably fulfill copper thin film needs. However, the available volatile precursors for copper CVD processes are very limited. No volatile, simple, binary copper compounds similar to silicon halides or hydrides are known. Rather, copper CVD precursors are complex organometallic species that are difficult to manufacture and costly to use. Additionally, CVD copper processes tend to produce relatively impure copper films and blanket CVD copper films do not lend themselves to subtractive patterning. As a consequence of the issues with PVD and CVD copper, the primary method for depositing copper layers in semiconductor device fabrication is a unique-to-copper electroless deposition process [144] [127] [145] [146].

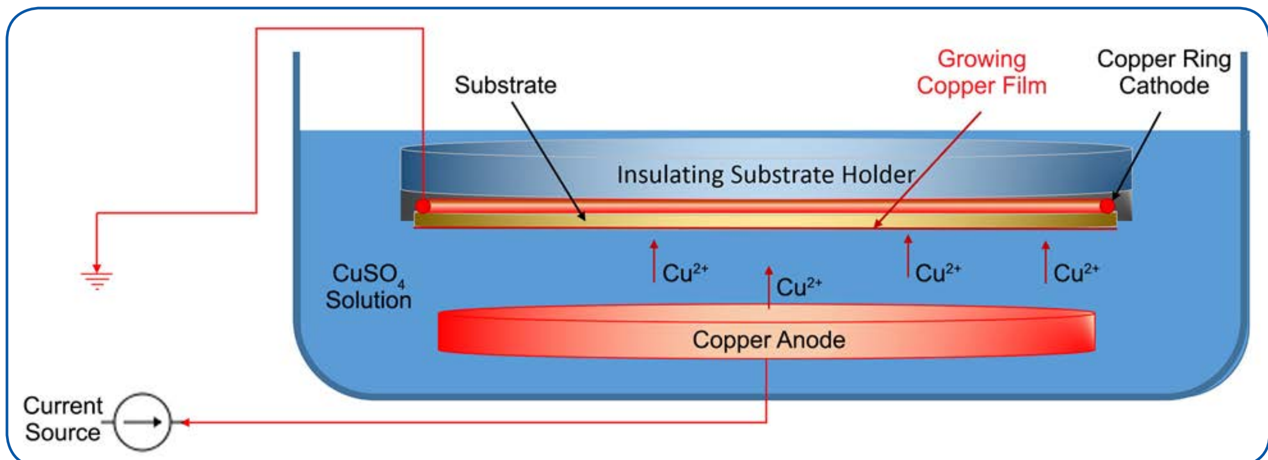


Figure 48. A a schematic of an electroplating process [147].

Figure 48 shows a schematic representation of the electroplating process. Electrochemical processes for copper typically use some variant on solutions containing inorganic copper salts to deposit bulk copper on a substrate surface. In electroplating processes, these solutions typically employ copper sulfate, $CuSO_4$ and a copper anode to drive an electrochemical reduction:

$$Cu^{2+} + 2e^{-} \rightarrow Cu^{0}$$

Electroless plating reactions (i.e. electrochemical processes that do not require external electrodes) are also being explored. These processes reduce copper ions to copper metal without the need for electrochemically driven reactions, for example:

$$Cu^{2+} + 2HCHO + 4OH^{-} \rightarrow Cu^{0} + 2HCOO^{-} + 2\,H_2O + H_2$$

The actual solution chemistries employed by different equipment/process suppliers for use in their copper deposition tools are highly proprietary. These proprietary solutions tend to be complex mixtures of inorganic copper salts with various inorganic and/or organic process promoters.

Coupled with the problems in finding viable deposition processes for copper was the fact that patterning copper films cannot be accomplished using conventional methods. Common fluoride or chloride-based dry etching technology does not work for copper since the fluoride and chloride by-products of the etching process are non-volatile and cannot be easily removed to create a patterned interconnect structure. These difficulties with the deposition and patterning of copper led, in the late 1990's, to the development of a unique copper metallization process known as the Damascene Process, shown in Figure 49 [148]. Damascene processing completely eliminated the need for the use of dry etching in patterning copper lines and vias. Rather, dry etching is performed on the dielectric material where the material is placed. Damascene processes also introduced the use of special barrier layers such as Tantalum, Tantalum Nitride, Titanium Nitride, and Titanium-Tungsten that prevented copper diffusion into other materials in the IC and improved the adhesion of the subsequent copper layer. Figure 49 shows a schematic for the Dual Damascene process. In this process, the interlayer dielectric is first deposited on the substrate using conventional CVD methods, then patterned using standard photolithography and dry etching techniques to produce the via pattern. Next, the trench pattern is created, again using conventional photolithography/dry etch. A barrier layer impermeable to copper diffusion is then deposited, followed by the deposition of a PVD or CVD copper seed layer. Once the seed layer is in place, the substrate undergoes a blanket copper electroplating process, filling both the vias and trenches in a single step to create the copper interconnect. Following the electroplating process, the copper on the surface of the substrate is polished back using chemical mechanical polishing, CMP. This step stops when it contacts the interlevel dielectric (ILD), leaving only the copper in the vias and trenches. The CMP process is described in more detail in Section B, Chapter VIII. Finally, a capping layer of a material impermeable to copper diffusion is deposited, resulting in complete encapsulation of the copper in the vias and trenches. Manufacturable copper process technology was first introduced by IBM [149] and Motorola [150] in 1997 and developments since then have made copper metallization common throughout the industry.

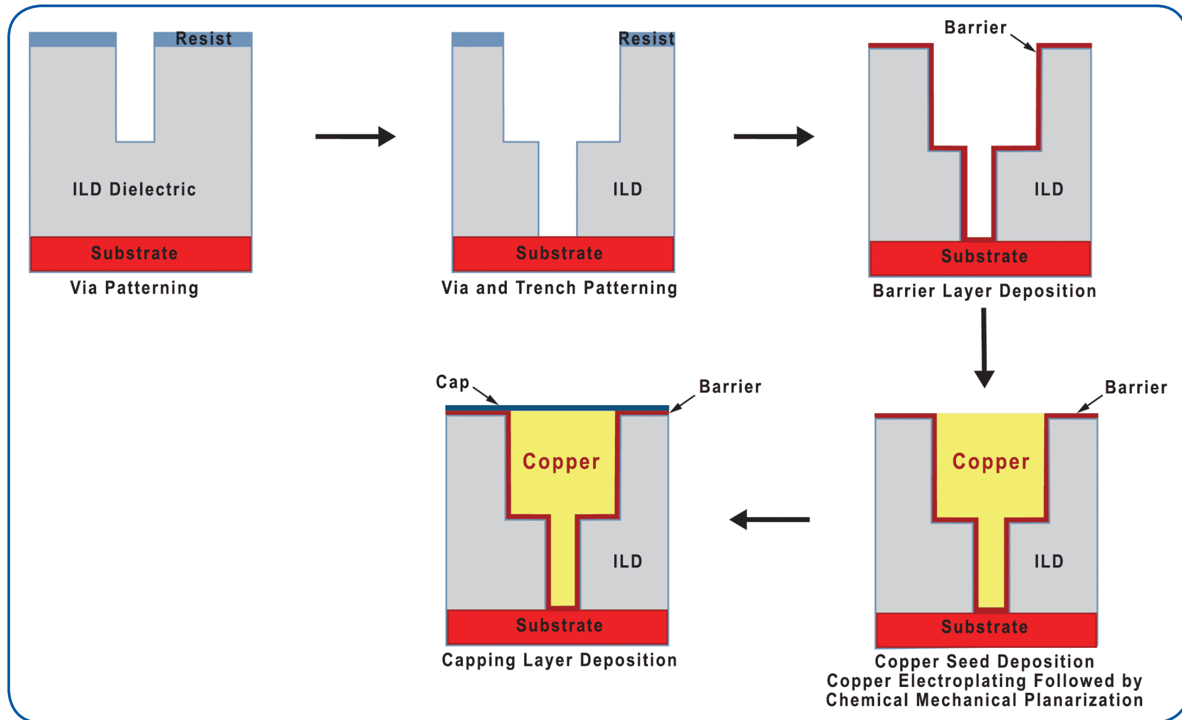


Figure 49. Schematic representation of the Dual Damascene Process.

e. Barrier Layers

Barrier layers are needed as metal diffusion barriers for contact electrodes and copper capping layers. Depending on the materials and process compatibilities between the barrier layer and silicon or a particular metal, different barrier layers are preferred for different applications. However, a number of general materials properties can be identified for barrier layers, including:

- High degree of metal impermeability at sintering temperatures (sintering refers to joining of the materials by thermal means)
- Ohmic and low contact resistance
- Good electrical conductivity
- Good adhesion between the semiconductor and metal
- High electromigration resistance
- Good step coverage
- High deposition rates in high aspect ratio gaps
- Compatible with chemical-mechanical polishing (CMP)

Aluminum is no longer common in contact structures and, so, we will not consider barrier layers for aluminum in any detail. In those remaining multilevel metallization schemes that still employ aluminum, titanium nitride, TiN, is the barrier material of choice since it has a low enough resistivity that it can become the principle conductor should electromigration produce a void in an aluminum line. TiN is usually

deposited by PVD methods (typically by ion plasma PVD). CVD TiN processes have also been developed that use either inorganic halide or organometallic titanium chemistries:

$$3TiCl_4 + 7NH_3 \rightarrow 3TiN + 4NH_4Cl + 5/2H_2$$

$$Ti[N(CH_3)_2]_4 + 4NH_3 \rightarrow TiN + \text{Mixed by-products}$$

The ammonium chloride product of the inorganic ($TiCl_4$) route to TiN can lead to high maintenance requirements in the deposition equipment and this has produced a strong interest in organometallic CVD processes employing TDMAT ($Ti[N(CH_3)_2]_4$) and similar chemical compounds. Titanium nitride is also commonly used as a barrier layer for tungsten contact plug and interconnect applications.

Barriers, liners and capping layers are an inherent part of copper metallization technology. Damascene and Dual Damascene Cu metallization processes normally employ Ta/TaN as the barrier layer for sidewalls and at the bottom of lines to prevent copper diffusion into surrounding device components. Typically, these layers are deposited using advanced PVD methods that can achieve uniform and conformal coverage in high aspect ratio vias and trench structures. Once copper metallization is in place on a device, the metal requires a top (capping) barrier to prevent corrosion of the copper and to promote adhesion between the copper and subsequent dielectric layers. Capping layers that have been developed for copper are primarily relatively high-*k* dielectric materials, typically silicon nitride, silicon carbon hydrogen, or silicon oxide carbide. Reference [151] provides a comprehensive discussion of barrier and capping layers for copper metallization.

This page intentionally left blank.

III. Example Fabrication Process: CMOS Devices

The materials requirements and methodologies employed for the manufacture of a semiconductor device are admirably demonstrated using a discussion of the steps and processes involved in the production of a basic CMOS device structure. As noted earlier, the term CMOS stands for Complementary Metal Oxide Semiconductor with CMOS devices incorporating both NMOS and PMOS transistors. Figure 50, taken from UC Berkeley EE course notes on Digital Integrated Circuits by Jan M. Rabaey and Andrei Vladimirescu, shows the basic operations required for the formation of a planar CMOS device in its most simplified form. This section describes the individual steps involved in a typical CMOS process flow for the planar devices employed in older technologies for the purpose of illustrating device fabrication at a level accessible to readers with a range of technical backgrounds.

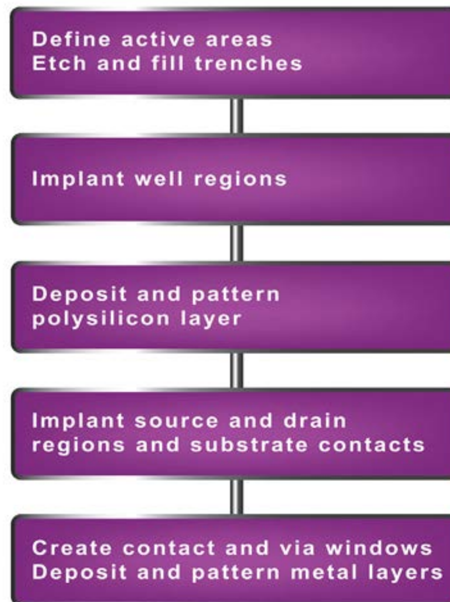
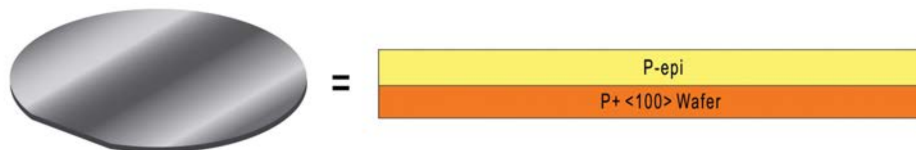


Figure 50. Simplified CMOS process flow [152].

A. Wafer Surface Cleaning



CMOS devices are normally fabricated on substrates having a p-type <100> silicon surface. Modern CMOS devices employ p+ silicon wafers that have a p-type epitaxial layer (sometimes referred to as “epi” wafers) on which the devices are to be formed. The p-type epitaxial silicon on a p+ substrate configuration achieves higher breakdown voltages while simultaneously keeping collector resistance low in the device. The material and process characteristics of epitaxial silicon are described in Section A, Chapter II.B.2. The epitaxial silicon layers used for CMOS device fabrication are typically deposited on silicon substrates using dichloro- or trichlorosilane in a hydrogen carrier gas at temperatures that range between 1100°C and 1200°C.

mks

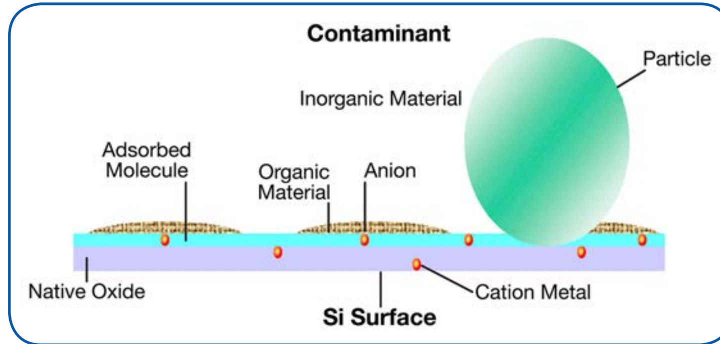


Figure 51. The surface contaminants that may be present on a silicon wafer.

Prior to a wafer's entry into the CMOS fabrication process, its surface must be cleaned to remove any adhering particles and organic/inorganic impurities. Silicon native oxide also needs to be removed. Continually shrinking device design rules have made cleaning technologies ever more important to achieving acceptable product yields. In modern device fabrication, wafer cleaning procedures can make up 30% – 40% of the steps in the total manufacturing process. Wafer cleaning has a long developmental history within the semiconductor industry. See collections and studies of the classic methods from authors such as Werner Kern [47], [153] and Tadahiro Ohmi [154] for more detailed discussion.

Contaminants on wafer surfaces may be present as adsorbed ions and elements, thin films, discrete particles, particulates (clusters of particles), and adsorbed gases [153]. Figure 51 shows a schematic of the kinds of contaminants present on the wafer surface prior to it entering the process flow; Table 12 describes the impact of the different kinds of surface contamination on device performance while Table 13 shows the cleaning solutions employed to remove the different contaminants.

<table>
<tr><th colspan="3">Type of Contamination</th><th>Main Influences on Device Characteristics</th></tr>
<tr><td colspan="3">Particle Contamination</td><td>• Pattern defects
• Ion implantation defects
• Insulating film breakdown defects</td></tr>
<tr><td rowspan="2">Metallic Contamination</td><td colspan="2">Alkali Metals</td><td>• MOS transistor instability
• Gate oxide film breakdown/degradation</td></tr>
<tr><td colspan="2">Heavy Metals</td><td>• Increased PN junction reverse leakage current
• Gate oxide film breakdown defects
• Minority carrier lifetime degradation
• Oxide excitation layer defect generation</td></tr>
<tr><td rowspan="5">Chemical Contamination</td><td colspan="2">Organic Material</td><td>• Gate oxide film breakdown defects
• CVD film variations (incubation times)
• Thermal oxide film thickness variations (accelerated oxidation)
• Haze occurrence (wafer, lens, mirror, mask, reticle)</td></tr>
<tr><td rowspan="3">Inorganic Contamination</td><td>Dopants (B, P)</td><td>• MOS transistor V_{th} shifts
• Si substrate and high resistance poly-silicon sheet resistance variations</td></tr>
<tr><td>Bases (amines, ammonia)</td><td rowspan="2">• Degradation of the resolution of chemically amplified resists
• Occurrence of particle contamination and haze due to salt generation</td></tr>
<tr><td>Acids (SOx)</td></tr>
<tr><td colspan="2">Native and Chemical Oxide Films Due to Moisture, Air</td><td>• Increased contact resistance
• Gate oxide film breakdown/degradation</td></tr>
</table>

Table 12. Wafer contamination and its effects.

Contaminant	Cleaning Procedure Name	Chemical Mixture Description	Chemicals
Particles	Piranha (SPM)	Sulfuric acid/hydrogen peroxide/DI water	$H_2SO_4/H_2O_2/H_2O$ 3-4:1; 90°C
	SC-1 (APM)	Ammonium hydroxide/hydrogen peroxide/DI water	$NH_4OH/H_2O_2/H_2O$ 1:4:20; 80°C
Metals (not copper)	SC-2 (HPM)	Hydrochloric acid/hydrogen peroxide/DI water	$HCl/H_2O_2/H_2O$ 1:1:6; 85°C
	Piranha (SPM)	Sulfuric acid/hydrogen peroxide/DI water	$H_2SO_4/H_2O_2/H_2O$ 3-4:1; 90°C
	DHF	Dilute hydrofluoric acid/DI water (will not remove copper)	HF/H_2O 1:50
Organics	Piranha (SPM)	Sulfuric acid/hydrogen peroxide/DI water	$H_2SO_4/H_2O_2/H_2O$ 3-4:1; 90°C
	SC-1 (APM)	Ammonium hydroxide/hydrogen peroxide/DI water	$NH_4OH/H_2O_2/H_2O$ 1:4:20; 80°C
	DIO3	Ozone in de-ionized water	O_3/H_2O Optimized Mixtures
Native Oxide	DHF	Dilute hydrofluoric acid/DI water	HF/H_2O 1:100
	BHF	Buffered hydrofluoric acid	$NH_4F/HF/H_2O$

Table 13. Cleaning solutions used to prepare substrates for the CMOS process.

1. Contaminant Types and Solution Cleaning Methods

a. Particles

Particle contamination can originate as airborne dust from a variety of sources including fab equipment, process chemicals, the internal surfaces of gas lines, wafer handling, gas phase nucleation in film deposition systems, and fab operators. Even particles of low nanometer dimension have the potential to generate “killer” defects, either through the action of physically occluding the formation of key features in the device (producing patterning, feature and implant defects) or by creating localized electrically weak spots in thin insulating films. Cleaning solutions for particle contamination include piranha cleans for gross particulate (and organic) contamination and SC-1 cleans for small, strongly adhering particles. Piranha solutions are extremely strong acids which oxidize many surface contaminants to produce soluble species that can be removed in solution. SC-1 solutions remove insoluble particles by oxidizing a thin layer of silicon on the surface of the substrate which then dissolves into the solution, carrying adsorbed particles with it. Modern SC-1 cleans employ megasonic (0.8 – 2.0 MHz) vibration to aid in the removal of particles from the surface. SC-1 solutions prevent re-adsorption of the particle by inducing the same zeta potentials a measure of electrostatic repulsion, on the particle and substrate surfaces. All cleaning solutions that contain hydrogen peroxide (piranha, SC-1, SC-2) leave a thin oxide layer on the silicon wafer surface.

b. Metals

Semiconductor devices are particularly sensitive to metallic contaminants since metals are highly mobile in the silicon lattice (especially metals such as gold) and therefore they easily migrate from the surface into the bulk of the silicon wafer. Once in the bulk silicon, even moderate process temperatures cause metals to rapidly diffuse through the crystal lattice until they are immobilized at crystal defect sites. Such “decorated” crystal defects degrade device performance, permitting larger leakage currents and producing lower breakdown voltages [32]. Metal contaminants can be removed from the substrate surface using an acidic clean such as SC-2, piranha or dilute hydrofluoric acid (HF); these cleans react with the metal to produce soluble, ionized metallic salts that can be rinsed away.

c. Chemical Contamination

Chemical contamination of the substrate surface can be broken down into three types: surface adsorption of organic molecular compounds; surface adsorption of inorganic molecular compounds; and an ill-defined, covalently-bound thin (ca. 2 nm) native oxide consisting of the chemical oxide/hydroxides of silicon, $SiO_x(OH)_y$.

Organic Compounds

Surface contamination by organic compounds either through airborne contamination or as residue from organic photoresists (PRs) is omnipresent in cleanrooms due to the presence of volatile organic solvents/cleaners and outgassing from polymer construction materials. Gross contamination by organics, such as occurs with incomplete PR removal, can impact device yields by leaving residues that form carbon during high temperature process steps. These carbon residues can form nuclei which behave as particle contaminants. Small amounts of residual metal resident in the PR compound can be trapped on the surface in these carbon residues [153]. PR residual contamination can be removed using piranha cleans and other high efficiency PR cleanup methods, as described in Section B, Chapter II.B.

Organic contamination due to ubiquitous volatile airborne contaminants also require removal from the wafer surface. The presence of these contaminants can hinder the removal of native oxide by dilute HF solutions (see below), producing poorly defined interfaces between the gate oxide and the substrate and gate electrode. Poor interface characteristics strongly degrade gate oxide integrity [155] [156] [157] [158]. The presence of organic compounds on the surface can affect the initial rates of both thermal oxidation and CVD processes, introducing undesirable and unknown variations in film thickness [159]. The SC-1 clean removes these organic residues through oxidation by peroxide and solvation of the products by NH_4OH. The SC-1 clean slowly removes any native oxide, replacing that layer with a new oxide produced by the oxidizing action of the peroxide. In recent years, ozone dissolved in DI water (DIO_3) is finding increasing use as a replacement for older Pirhana and SC-1 cleans as a “green” and safer alternative for the removal of organic contaminants.

Inorganic Compounds

Chemical compounds containing dopant atoms such as boron and phosphorus can be present on wafer surfaces due to effects such as the outgassing of phosphorus-containing flame retardants or dopant residuals in process tools [160]. If they are not removed from the wafer surface prior to high temperature processing, these elements can migrate into the substrate, modifying the targeted resistivity. Other kinds of volatile inorganic compounds such as basic compounds like amines and ammonia and acidic compounds like sulfur oxides (SO_x) will also produce defects in semiconductor devices if they are present on the substrate surface. Acids and bases can cause unintentional shifts in the basicity or acidity of chemically amplified resists leading to problems in pattern generation and resist removal. These compounds are highly reactive and will readily combine with other volatile ambient chemical species to create particles and haze due to the formation of chemical salts on the substrate surface. Adsorbed acidic and basic species can be removed from the substrate surface by the combined action of SC-1 and SC-2 cleans.

Native Oxide

Silicon, like many elemental solids, naturally forms a thin layer of oxidized material on its surface by reaction with oxygen and moisture in the ambient air [161]. The chemical formulation of this layer is not well-defined, being a more or less random aggregation of Si-O-Si, Si-H and Si-OH species. The presence of this native oxide on the silicon surface causes problems in semiconductor device manufacturing since it can lead to difficulties in controlling the formation of very thin thermal oxide thicknesses. Any native oxide that is present on the substrate during thin gate oxide formation will electrically weaken the gate insulator through the incorporation of hydroxyl groups. Additionally, if native oxide is present on the silicon surface of a contact pad, it will increase the electrical resistance of that contact. Over the past 50 years, our understanding of the nature of silicon native oxide and its impact on device performance has greatly increased [162] [163] [164] [165] [166]. These studies found that very dilute solutions of HF, in de-ionized water, DI, or dilute solutions of ammonium fluoride, NH_4F, HF and DI water (buffered oxide etch, BOE) completely remove silicon native oxide, leaving a hydrogen-terminated clean silicon surface according to Figure 52.

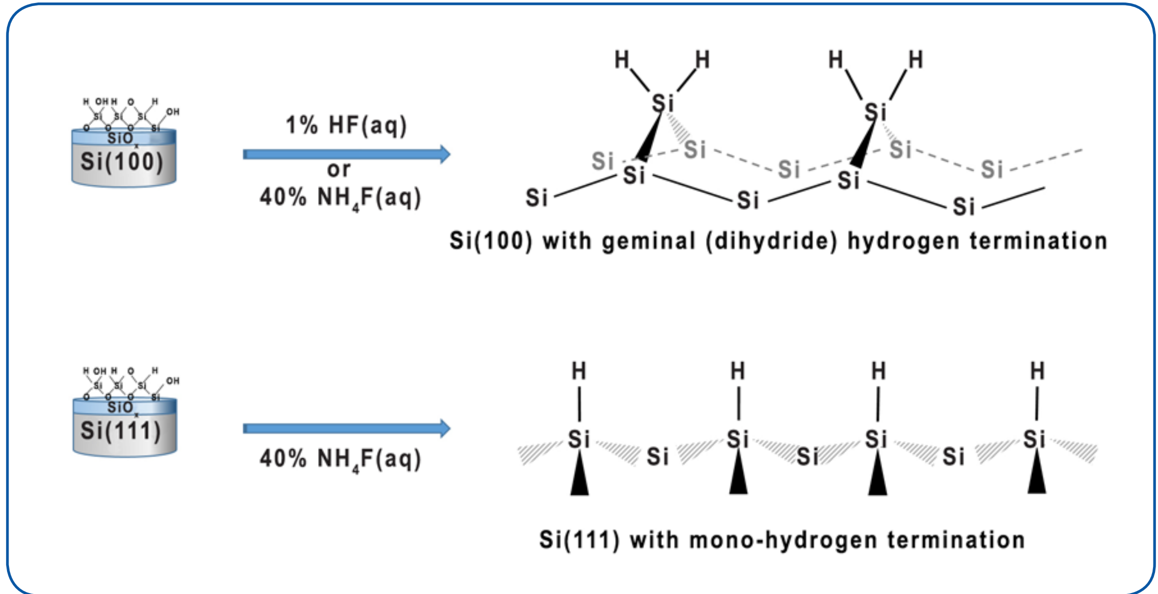


Figure 52. Native oxide removal and H-terminated silicon surface characteristics (after [167]).

RCA Cleaning

The first successful wet-cleaning process for front-end-of-line (FEOL) silicon wafers was developed at RCA by Werner Kern and co-workers and published in 1970. Since then, there have been many developments and successful modifications of the approach and RCA cleaning continues to be the primary FEOL pre-deposition cleaning in the industry today.

RCA cleaning procedures are a combination of the different procedures described above. The process consists of consecutive SC-1 and SC-2 solutions, followed by treatment with a dilute HF solution or buffered oxide etch (BOE). The product is a clean, hydrogen-terminated silicon surface, ready to be used in the CMOS process flow.

B. CMOS Device Structure

The CMOS process flow can be demonstrated using a description of a simple device formed by two inverters. This discussion is adapted from on-line lecture notes that have been made available from University of California Berkeley [152]. Figure 53 shows the circuit schematic for the device.

The device structure employs a twin-tub CMOS process that has a number of advantages. It allows the manufacturer to separately optimize the n- and p-type transistor structures which, in turn, allows for independent tuning of the electrical characteristics of the n- and p-type devices.

The starting material for device fabrication is a heavily doped silicon <100> wafer with a lightly doped epitaxial layer, as described above. This substrate structure protects the device against “latch-up,” a condition that allows excessive current flow through the device.

mks

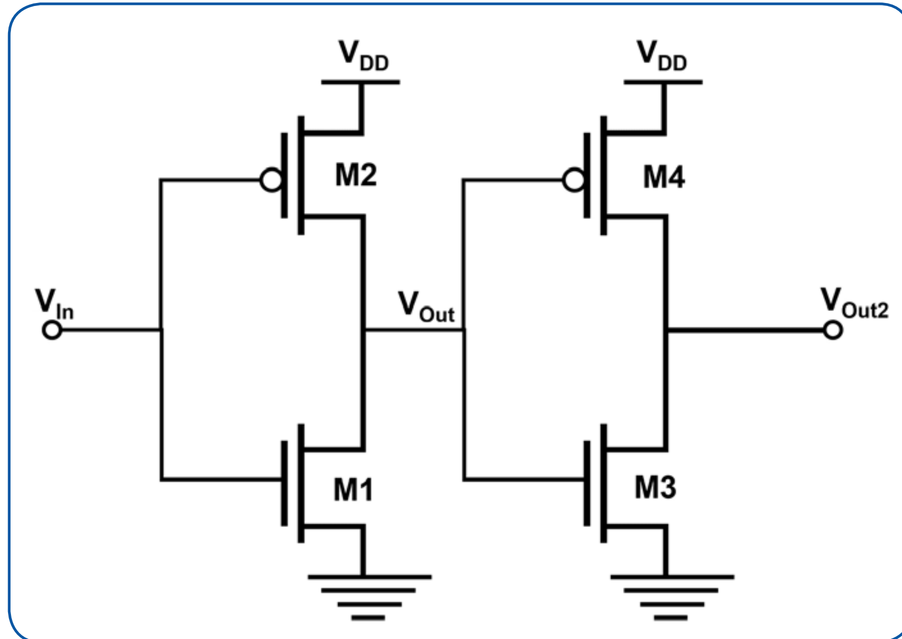


Figure 53. Two-inverter circuit [152].

Figure 54 shows a schematic cross-section of the device structure detailing the different layers and materials used in the CMOS fabrication process for a planar device.

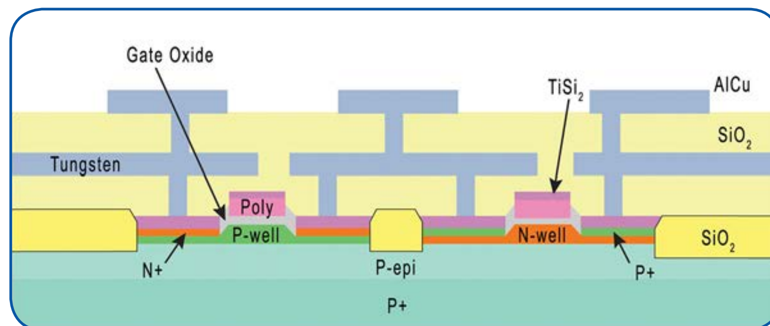
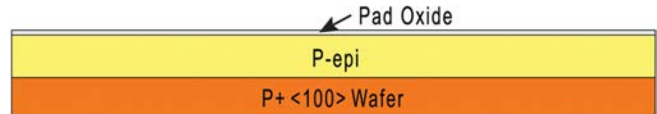


Figure 54. Schematic of the physical cross-section of the twin-tub two inverter device structure [152].

The process described in this section is simplified as necessary and does not address the many advanced materials and methods that have been developed for sub-100 nm device designs.

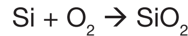
C. CMOS Process Steps

1. Pad Oxide (Thermal Oxidation)



The initial step in the CMOS process is the formation of a “pad” thermal silicon dioxide layer on the wafer surface. The pad oxide relieves stress between the substrate and the subsequent silicon nitride layer (see below), diminishing stress-induced dislocations in the substrate (thick nitride layers can induce such dislocations [168] [169]). Silicon dioxide’s excellent electrical properties and the fact that thin films of the material can be formed by direct oxidation of the substrate have made thermal silicon dioxide the primary insulating film material employed in semiconductor device manufacture. Thermal oxides have high electrical resistivity (>10^{20} Ω-cm), a high energy band gap (~9 eV) and a high breakdown field strength (>10 MV/cm). The films exhibit conformal growth on exposed silicon surfaces and, when grown on H-passivated silicon surfaces, have extremely regular, stable and reproducible Si/SiO_2 interfaces. Their high etch selectivity with HF is also a very desirable characteristic in device manufacturing processes.

Pad oxides are 10 - 50 nm thick, typically grown using a “dry” oxidation process:



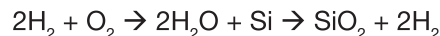
$$Si + O_2 \rightarrow SiO_2$$

Dry oxidations are performed at temperatures ranging from 900°C to 1200°C at high oxygen pressures. Dry oxidation processes exhibit relatively low oxide growth rates when compared against other thermal oxidation processes, typically around 14 – 25 nm/hr. Because of this, dry oxidation processes are normally used only when silicon dioxide film thicknesses of less than 100 nm are needed. In general, dry oxidation processes produce silicon dioxide films with the highest quality electrical and material characteristics.

Thermal silicon dioxide films can also be produced using the reaction of the substrate with steam in what is known as a “wet” oxidation. Wet oxidations can be performed using entrained water:

$$Si + 2H_2O \rightarrow SiO_2 + 2H_2$$

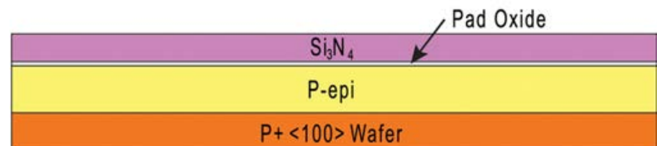
or using the in situ generation of steam through the reaction of hydrogen and water:



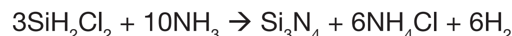
$$2H_2 + O_2 \rightarrow 2H_2O + Si \rightarrow SiO_2 + 2H_2$$

A more detailed discussion of thermal oxidation processes is provided in Section B, Chapter IV.A.

2. Silicon Nitride



The next step is the deposition of a layer of silicon nitride over the pad oxide. This layer acts as a stop for the chemical mechanical polishing (CMP) step later in the process. Silicon nitride thin films can be deposited by Low Pressure Chemical Vapor Deposition (LPCVD) using the chemical reaction:



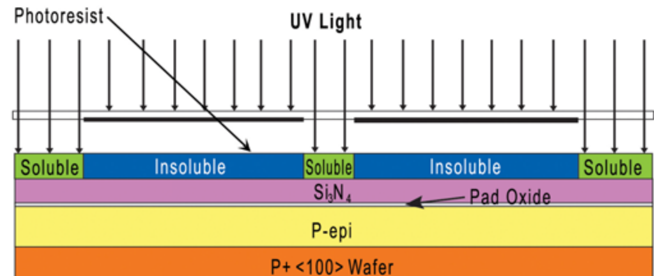
$$3SiH_2Cl_2 + 10NH_3 \rightarrow Si_3N_4 + 6NH_4Cl + 6H_2$$

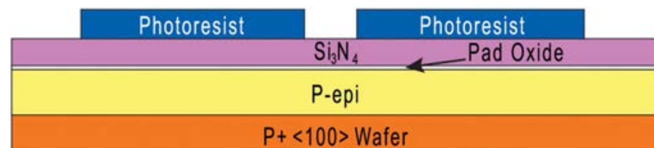
The equipment and processes employed for silicon nitride and other LPCVD thin films are described in greater detail in Section B, Chapter V.

3. Shallow Trench Isolation (STI) Process

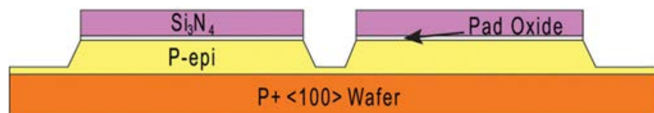
In a CMOS device, shallow trenches filled with silicon dioxide are used to electrically isolate the n- and p-type active areas on the substrate surface. The following procedure is used to create these trenches in technologies below 300 nm.

The silicon nitride layer is patterned using a photolithography process. The first step in photolithography is the deposition of a layer of photoresist. Photoresists are light sensitive organic materials that become either more (positive resist) or less (negative resist) soluble in selected solvents on exposure to appropriate wavelengths of light. In the first step of the STI process, a photoresist layer is deposited on the silicon nitride using a method known as spin coating. In this process the substrate is rotated at a high angular velocity (up to 5000 rpm) while a viscous liquid solution of the photoresist is dispensed onto the center of the substrate. Centrifugal force drives the photoresist solution to the edges of the substrate and a layer of photoresist with a very uniform thickness across the substrate is deposited on the surface. The target thickness of the layer varies, depending on the particular CMOS process. Following formation of the photoresist layer, the substrate is subjected to a "soft bake" at an elevated temperature that removes the solvent in which the photoresist was dissolved. Once the photoresist layer has dried, it is patterned using light exposure. Typically, CMOS processes use ultraviolet (UV) light and step-and-repeat patterning using a tool called a "stepper." After exposure, the substrate, with resist, is subjected to another bake step that further hardens the photoresist layer that remains behind in the unexposed regions. Next, the substrate undergoes a development step that dissolves away the photoresist in the exposed areas, forming the desired pattern in the photoresist mask on the silicon nitride layer. Advanced patterning tools and the photolithography process in general are described in detail in Section B, Chapter VI.



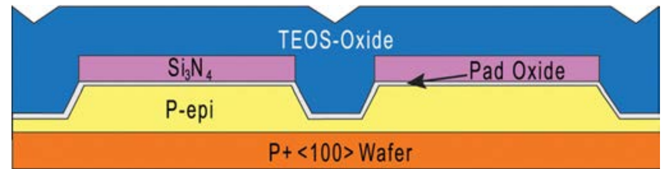


Once the mask has been formed, the isolation trenches can be created. This is done by etching the trench into the substrate, then backfilling the trench with deposited silicon dioxide. The substrate—with its masking photoresist and underlying silicon nitride layers—first undergoes a plasma etch process to remove material in those areas not covered by photoresist (silicon nitride, pad oxide and epitaxial silicon), creating the isolation trenches. A sequential series of plasma etch chemistries is required to remove these different materials. The silicon nitride layer is removed using a fluorine atom etch that employs sulfur hexafluoride as the fluorine source. The pad oxide layer is removed using a fluorine atom etch that employs carbon tetrafluoride as the source gas. Finally, the silicon in the epitaxial layer is removed using a mixture of difluoroethylene and sulfur hexafluoride as the fluorine atom source. The reasons for these different chemistries rests in the need to optimize the etch rates and etch directionality for each individual layer being removed in the trench formation process. Etch processes are described in Section B, Chapter VII.



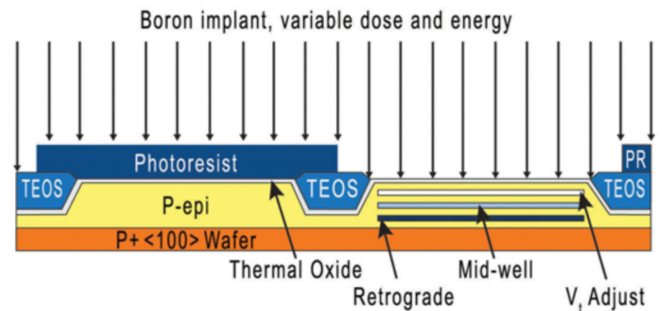
Following trench formation, the photoresist layer is removed in a plasma ashing step (see Section B, Chapter II.B) and a wet clean is performed that leaves a passivated silicon surface in the trenches (typically an RCA clean incorporating piranha, SC-1, SC-2 and DHF). A trench liner of thermal silicon dioxide is grown on the exposed epitaxial silicon within the trench, followed by a trench filling oxide deposition step

in which an organometallic silicon source (tetraethoxysilane, TEOS, $Si(OC_2H_5)_4$) is reacted with ozone, O_3, in a high density plasma deposition process known as HDP-CVD to deposit undoped silicon dioxide in the trenches. The process characteristics and equipment requirements for TEOS-Oxides and HDP-CVD are discussed in detail in Section B, Chapter V.



The final step in the STI process is the use of chemical mechanical polishing (CMP – see Section B, Chapter VIII) to planarize and polish the surface. In this step, the purpose of the silicon nitride layer becomes apparent as the much greater hardness of the nitride allows it to act as a stop that prevents the CMP process from removing the TEOS-oxide in the shallow trenches. Following the CMP process, the remaining nitride is stripped off selectively using phosphoric acid (H_3PO_4) at 140°C and the pad oxide removed using HF. A new layer of thermal oxide is then grown on the substrate surface using dry oxidation.

4. N- and P-Well Process

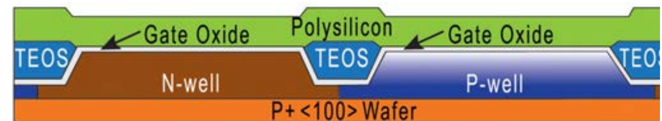


The next step in the CMOS process flow is the formation of the active area n- and p-wells. These wells are formed in the open substrate surface areas defined by the shallow trenches. In the first step of well formation, a photoresist is deposited and patterned (as described above) so that the n-well area is masked and the p-well area is exposed. The substrate then undergoes an ion implantation process to deposit boron into the substrate in the exposed p-well area. In the ion implantation process, a boron (or phosphorus) source is first ionized under high vacuum conditions to create $^{11}B^+$ ions in the gas phase. The source can be a solid target or a gaseous boron compound, depending on the equipment used. The $^{11}B^+$ ions are passed through a magnetic field analyzer that eliminates any contaminating ionic species before being accelerated using electromagnetic fields to a selected energy. The substrate to be implanted is located in a process chamber at the end of the ion acceleration column and it undergoes ion bombardment by the $^{11}B^+$ ions coming out of the acceleration column. The $^{11}B^+$ ions penetrate into the silicon substrate to a depth that is dependent on the energy imparted to them in the accelerating column. The final position and concentration of dopant atoms in the substrate is dependent on the ion energy and exposure time, respectively. In modern CMOS devices, different ion energies are employed to create zones in the substrate that have different dopant concentrations at different depths. Following p-well ion implantation, the n-well for the device is created in a similar manner. The final step in well formation is a rapid thermal anneal of the substrate that removes any lattice damage that may have occurred in the implantation process and establishes a continuous vertical dopant concentration profile in the well. Ion implantation equipment and processes are described in detail in Section B, Chapter III. It should also be noted that certain modern device technologies employ in situ doped layers to avoid the need for ion implantion.

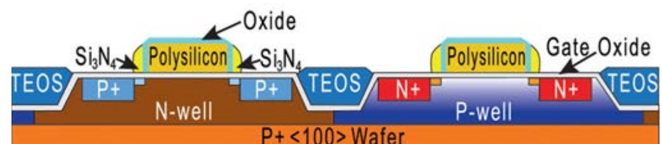
mks

5. Gate Process

The first step in the gate formation process is an RCA strip and clean (incorporating piranha, SC-1, SC-2 and DHF) that removes any residual contamination and the thin thermal oxide layer, leaving a hydrogen-passivated silicon surface. Following this, a thin dielectric layer that will become the gate oxide is grown on the exposed silicon in the well areas. Older device structures employed a thermal oxide layer for this purpose; however, as devices shrank in size and the gate oxide layer became ever thinner to accommodate lower gate voltage, leakage due to quantum tunneling of electrons through the gate oxide became a serious problem. For this reason, more advanced CMOS devices use multilayer gate oxide structures that employ barrier layers and high dielectric constant (high-k) materials. High-k materials achieve greater electrical fields in the substrate at higher physical oxide thickness than can be achieved using silicon dioxide. The greater physical thickness of the gate prevents quantum tunneling and excessive leakage. For the purpose of describing the CMOS process flow, we will assume the simpler case in which silicon dioxide is suitable for gate formation. The silicon dioxide gate oxide layer is grown, typically to a layer thickness of 5-10 nm, using a thermal oxidation process that is made more controllable through the use of slower oxidizers than oxygen, e.g., nitrous oxide, N_2O, as the oxidizing agent. Modified processes and applications for thermal oxidation and high-k dielectrics are discussed in detail in Section B.



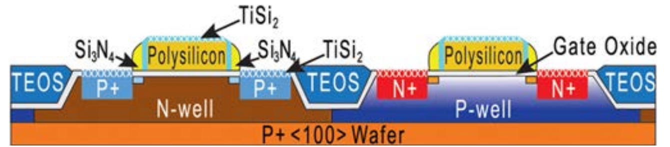
The next step in the gate process is the formation of the gate electrode using highly doped polycrystalline silicon. First, an undoped polysilicon layer (400 – 500 nm) is deposited over the gate oxide on the substrate. Polysilicon is usually deposited using a LPCVD process carried out at 200 – 300 mTorr and 600°C – 620°C. In advanced devices, a barrier layer such as TiN may be deposited on the oxide prior to the polysilicon deposition. Barrier layers are employed in advanced device structures to reduce or eliminate impurity diffusion from the polysilicon into the gate oxide. Polysilicon properties and processes, and LPCVD processes and equipment are described in detail in Section B, Chapter V.A. The substrate is next coated with photoresist, and the resist is patterned and developed to define the gate structure in the active areas. Using a plasma etch process that employs fluoride-based chemistry (i.e., CF_4, CH_2F_2, SF_6, etc.), the polysilicon layer is etched away except in those areas protected by the photoresist. Following the polysilicon etch process, the substrate undergoes a plasma ashing step to remove the photoresist. The polysilicon is then oxidized in a thermal oxidation process, forming a layer of oxide surrounding the polysilicon core of the gate structure. Next, a layer of LPCVD silicon nitride is deposited over the gate structure, followed by a plasma etch of the nitride layer that leaves nitride side-wall spacers on either side of the polysilicon gate. This structure is subjected to an ion implant process in which the polysilicon gate is heavily implanted with phosphorus to increase its electrical conductivity. Variations on the ion implant process flow are used to form lightly doped drain and source regions under the sidewalls, but these are not included in this discussion in order to simplify the presentation of fabrication principles.
The ion implant process creates the n+ and p+ source/drain regions in the P-well and N-well, respectively. A piranha clean is then performed to remove any remaining photoresist residues.



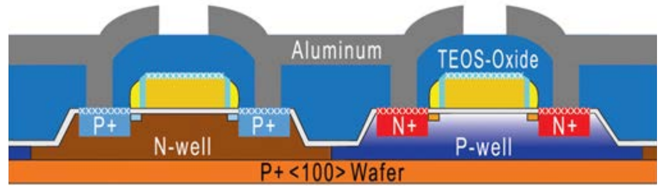
6. Contact Process

The substrate undergoes a short (~ 1 min) oxide etch using buffered oxide etch (BOE) that removes the thermal oxide from the top of the polysilicon gate and from the surfaces of the source and drain regions in the substrate. A sputtering process (PVD – see Section B, Chapter V.B) is used to deposit a layer of titanium over the gate, source and drain areas. This is followed by a rapid thermal process that converts the areas

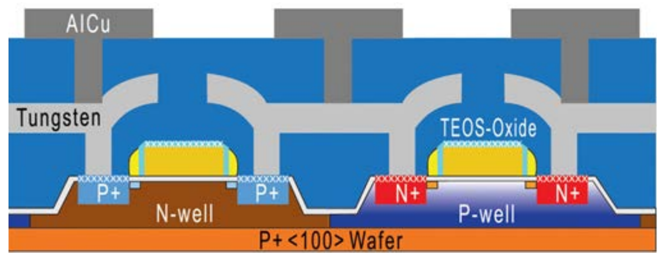
where titanium is in contact with silicon to titanium silicide, $TiSi_2$, i.e., the top layers of the gate, source and drain. The titanium layer is then etched in sulfuric acid, removing any remaining elemental titanium but leaving behind the $TiSi_2$. Titanium silicide improves the ohmic contact between metallization and the gate, source and drain. It should be noted that some processes use metal gates or add gates at the end of the device fabrication process (gate last processes).



The next step in this process is the deposition of a SiO_2 dielectric layer using a TEOS-CVD process. The TEOS-oxide process is chosen for this layer due to excellent step coverage characteristics that allow it to conformally cover the gate structure. The TEOS-oxide dielectric layer is then patterned and etched to open contact holes to the source/drain regions followed by an RCA clean to remove photoresist residues. Next, Tungsten is sputter deposited and is the first metal layer. The tungsten layer is patterned using photolithography so that the source/drain contacts are protected by PR while the tungsten/TEOS-oxide layer on top of the gate is exposed. The tungsten is then etched away using a reactive ion etch. While it is not evident in the accompanying cross-sections, the contact to the gate is also created in this segment of the process.



A TEOS-based CVD process is used to deposit a second layer of dielectric, commonly referred to as an intermetal dielectric. This dielectric layer is patterned using a conventional photolithography process to create contact holes to the first level metal which contacts the source/drain regions. A sputter deposition process and photolithography are employed to deposit and pattern an aluminum-copper alloy in this second-level metal process.



In summary, the CMOS structure is complete and forms the building blocks for complex IC structures.

This page intentionally left blank.

Section B

Process Equipment, Technology, and MKS Instrument's Product Applications

mks

This page intentionally left blank.

I. Vacuum Technology

This section provides a brief overview of vacuum science and technology. There are excellent treatments on the underlying science and technology of vacuum characteristics and technology available. See references [170] [171] [172] for in-depth treatments of both the science and practical aspects of vacuum generation, maintenance and measurement.

A. Basic Concepts

In simple terms, vacuum may be defined as:

A space or container from which the air has been completely or partially removed.

The first basic concept that the reader needs to understand is that of gas pressure. If you fill a balloon with gas, it is the pressure of the gas within the balloon that keeps it inflated. This pressure is a measure of the cumulative force of individual gas molecules colliding with each other and with the walls of their container (in this case the balloon). The force (pressure) that the gas exerts is a combination of the number of molecules present within the container and the velocity of their movement. The velocity of the molecules in the gas is given by a distribution that is dependent on the molecule's mass and temperatures, as shown in Figure 55.

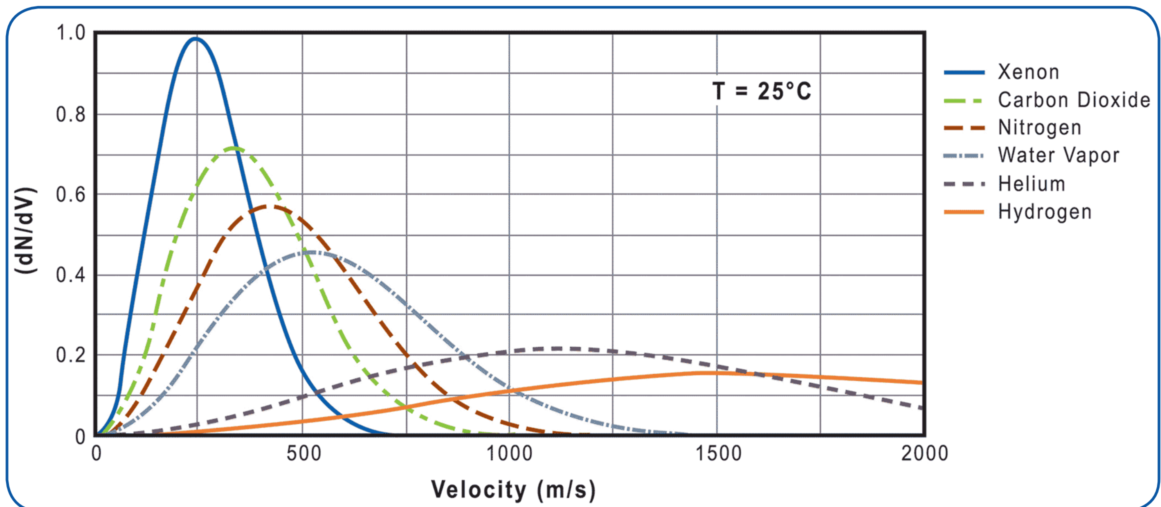


Figure 55. The relative velocity distribution of various gases at 25°C [172].

The pressure on a container's surface (in this case, the inside of the balloon) is defined as the rate at which momentum, mass × volume, *mv*, is transferred to the surface. Since molecules hitting the surface do so at many different angles, the amount of momentum transferred in any given collision is also dependent on the angle of incidence of the collision. Without going into the details of the derivation (these can be found in references [170] and [172]), pressure can be expressed in molecular terms by the equation

$$P = nkT,$$

where P is the pressure (in Pascals), n is the number density of gas molecules (in m^{-3}), k is a constant (known as the Boltzmann constant and having a value of 1.3806×10^{-23} joules/K) and T is temperature (K).

The key characteristic of vacuum for use in semiconductor processes is that a vacuum has so few atoms or molecules that they do not affect any process being carried out within the vacuum. Thus, one reason for vacuum processing in semiconductor manufacturing is the assurance of high purity in the finished product. The statement also implies other, more complex, advantages to the use of vacuum in semiconductor processing: e.g., ballistic transport of ions between a source and a substrate in ion implantation, or directional etching made possible by the lack of atom/ion scattering under vacuum conditions.

So, how many atoms or molecules is few? To understand this question, we first must understand how many atoms or molecules there are in a given volume of space at atmospheric pressure. Atmospheric pressure is the pressure around you on a normal day. We can get at this number (roughly) using the Ideal Gas law:

$$PV = nRT,$$

where P is the pressure, V is the volume, n is the number of moles of substance in that volume, R is a constant known as the "Universal Gas Constant" and T is temperature. On a normal day, the pressure of the atmosphere around you is commonly designated as 1 atmosphere (1 atm), so P = 1 atm. (Note that pressure can be expressed in other units. These include pounds/sq. inch (14.7 psi = 1 atm) and Pascals (101.3 kPa = 1 atm)). The volume that we will consider is a cubic meter, so V= 1 m^3, and the temperature is typically 25°C or 298K (K denotes degrees on the Kelvin temperature scale which starts at absolute zero or -273.15°C), so T = 298K. R, the Universal Gas Constant, has a value in these same units, of 8.205746 × 10^{-5} m^3-atm/°mole. This leaves only n, the number of moles of material present in the volume, to be determined. Re-arranging the Ideal Gas equation:

$$n = PV/RT = (1 \text{ atm} \times 1 \text{ litre})/(8.205746 \times 10^{-5} \text{ m}^3\text{-atm/K-mole} \times 298\text{K}).$$

Doing the calculation and canceling the units yields the following answer for the number of moles of gas that is present in a cubic meter of volume at 1 atmosphere pressure and room temperature:

$$n = 40.8945 \text{ moles.}$$

This number, in turn allows us to determine how many molecules are present in the cubic meter of gas since a mole of any material contains Avogadro's number, N_A, (6.022 × 10^{23}) of molecules:

$$\text{number of gas molecules in a m}^3 = n \times N_A = 40.8945 \times 6.022 \times 10^{23}$$

$$= 2.4626849 \times 10^{25} \text{ molecules in a cubic meter of gas at room temperature}$$

$$= 6.9735479 \times 10^{23} \text{ molecules in a cubic foot of gas at room temperature}$$

Description	Pressure (Torr)	Pressure (Pa)	Number of Molecules per m³ of Gas
Atmospheric Pressure	760	101.3 kPa	2.5×10^{25}
Low (Rough) Vacuum	25 to 760	3 kPa – 100 kPa	8.1×10^{23} – 2.5×10^{25}
Medium Vacuum	$1\text{x}10^{-3}$ to 25	100 mPa – 3 kPa	3.2×10^{19} – 8.1×10^{23}
High Vacuum	$1\text{x}10^{-9}$ to $1\text{x}10^{-3}$	100 nPa – 100 mPa	3.2×10^{13} – 3.2×10^{19}
Ultra-High Vacuum (UHV)	$1\text{x}10^{-12}$ to $1\text{x}10^{-9}$	100 pPa – 100 nPa	3.2×10^{10} – 3.2×10^{13}
Extremely High Vacuum	$< 1\text{x}10^{-12}$	< 100 pPa	$< 3.2 \times 10^{10}$
Outer Space	$<3\text{x}10^{-17}$ to $1\text{x}10^{-6}$	< 3 fPa - 100 μPa	970,000 - 3.2×10^{16}

Table 14. Vacuum classifications and molecules/litre.

The term vacuum, in the formal sense, describes any pressure less than normal atmospheric pressure. In pratical application, it is classified as one of three kinds of vacuum: either low (rough) vacuum, medium vacuum or high vacuum (Table 14). You can see from the values in the right-most column of Table 14 that even in the furthest reaches of outer space there are still nearly a million molecules of gas in a cubic meter. Most semiconductor processes are conducted under either medium or high vacuum conditions.

B. Creating a Vacuum

The creation of a vacuum can be simply stated as the removal of gas molecules from a defined volume. This simple statement masks some very complex phenomena associated with molecular flow in the gas phase. An in-depth understanding of the creation of vacuum requires consideration for physical characteristics such as gas viscosity, flow type (degree of turbulence), the thermal conductivity of the gas, and gas diffusion properties. An in-depth discussion of these factors is well beyond the intended scope of this work and the reader is again referred to references [170] [171] and [172]. In addition, a number of highly useful articles on practical aspects of vacuum technology are available on-line from the Normandale Community College in Bloomington MN [173].

In practice, industrial vacuums are produced using different kinds of vacuum pumps. Vacuum pumps can be based on a number of different principles [170]:

- Compression-expansion of the gas (piston pumps, rotary pumps, Root's (dry) pumps)
- Drag by viscosity effects (vapor ejector pumps)
- Drag by diffusion effects (vapor diffusion pumps)
- Molecular drag (molecular drag or turbomolecular pumps)
- Ionization effects (ion pumps)
- Physical or chemical sorption (sorption pumps, cryopumps and gettering processes)

In practice, within the semiconductor industry, the reader will primarily encounter mechanical pumps that operate on the compression-expansion principle, molecular drag pumps such as turbopumps, and physical sorption pumps such as cryopumps. We will therefore limit our discussion of pumping options for the creation of vacuum to these methods.

Mechanical Compression-Expansion Vacuum Pumps

Mechanical pumps that operate on the compression-expansion principle are by far the workhorse for vacuum creation within the semiconductor industry. Those wishing a detailed discussion of the construction of such mechanical pumps are referred to the seminal article by Harris and Budgen [174]. Figure 56 shows the internal construction of common mechanical vacuum pumps. Rotary vane and rotary piston pumps are known as two-sealed pumps with piston pumps commonly used for larger pumping capacity applications. Rotary vane pumps are oil-sealed units that operate on the principle of rotary displacement of the gas entering the pump. Figure 56(a) shows the basic components of this pump which include an inlet, a stator (housing), an eccentric rotor with vanes that move radially outward under a combination of centrifugal and spring forces, and an outlet. As the rotor rotates, a vane passes the inlet and gas enters the working chamber from the inlet. Continued rotation of the vane increases the volume of the working chamber, drawing in more gas until the eventual passage of the second vane past the inlet seals the gas between the two vanes within the working chamber. Further rotation compresses this captured gas increasing its pressure until it opens the discharge valve against atmospheric pressure. Oil in the outlet chamber prevents backflow of the discharged gas; some oil leaks back into the working chamber lubricating the system and sealing the vanes to the interior surface of the stator. An instructive animation of rotary vane pump operation has been produced by Edwards Vacuum and is available on YouTube [175]. Rotary vane pumps are typically used in applications requiring low or medium vacuum operation and come in single and two-stage pumping configurations; two-stage pumps are capable of achieving lower base pressures.

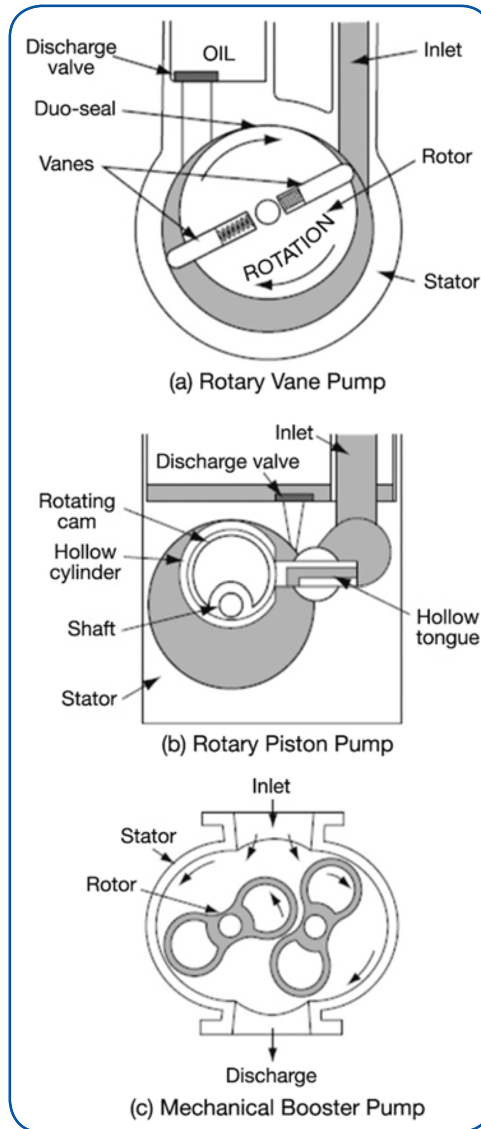


Figure 56. Compression-expansion pumps: (a) rotary vane pump; (b) rotary piston pump; (c) mechanical booster (Roots) pump [174].

While they can still be found in many semiconductor applications, oil-sealed rotary vane pumps have largely been supplanted in the industry by "dry" pumps owing to the fact that dry pumps offer less risk of oil contamination in the process (Figure 57). Dry pumps are also compression-expansion pumps; however, they employ different mechanisms to create this action. The most common mechanical arrangements for dry vacuum pumps in use in semiconductor processing include screw-type pumps (e.g., Edwards CSX Series [176] and Oerlikon Leybold's DRYVAC [177] pumps) and multi-stage lobe (Roots) pumps (e.g., Pfeiffer A4 Series [178]). In addition, claw pumps (e.g., Busch Mink Series [179]) are used for some vacuum applications.

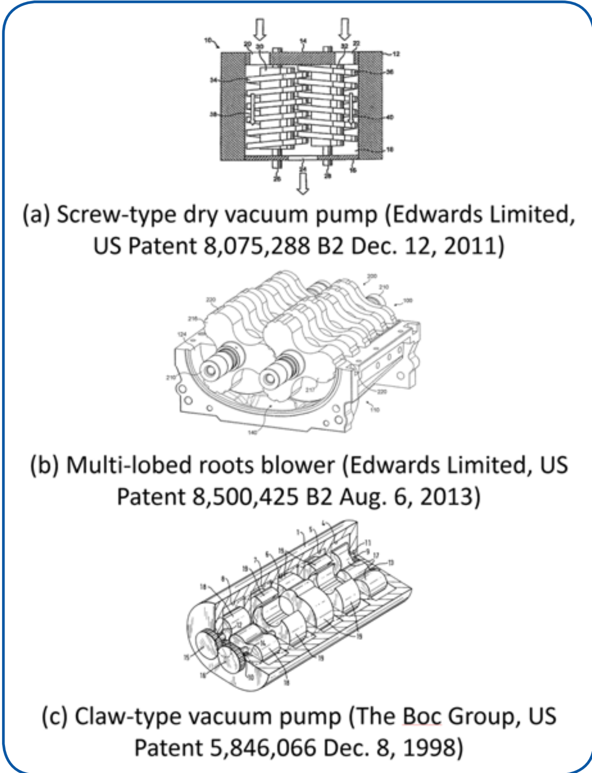


Figure 57. Different kinds of dry vacuum pumps.

All mechanical compression-expansion pumps are limited in the base vacuum pressure that they can achieve, with values of about 0.1-1 millitorr (~0.1-1 x 10^{-3} mbar) being typical. Note that different pump size/configurations are capable of handling different amounts of gas flow through the pump while maintaining vacuum, but this is a separate property from the base pressure achievable. Figure 58 shows typical pumping characteristics for a rotary vane pump. Pumping curves for screw and claw type pumps, while they may differ in relative pumping speeds and high pressure pumping characteristics for comparable models, show similar pumping curve shapes at pressures below 1 Torr (mbar), the pressure regime of interest for most semiconductor manufacturing processes. The fact that compression-expansion pump pumping speeds begin to fall off at pressures below 0.1 Torr is significant for semiconductor manufacturing. This characteristic means that prolonged times are needed to achieve base pressure when using only a mechanical rotary vane, screw or claw pump. This problem can be solved by the addition of a Roots blower to the vacuum pump. Figure 59 shows a pumpdown curve for a mechanical compression-expansion pump versus systems using two different size

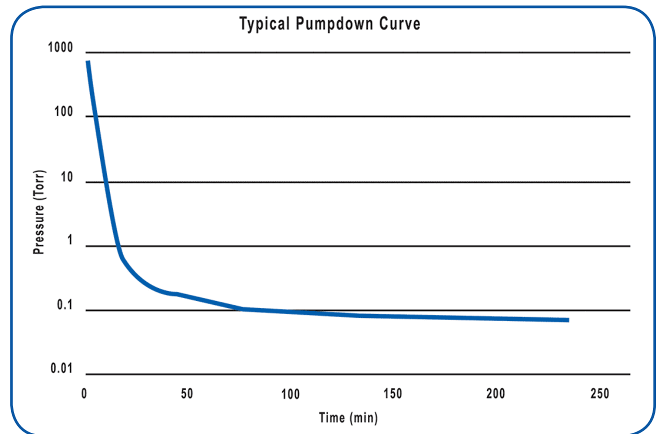


Figure 58. Typical pumpdown curve for a mechanical compression-expansion vacuum pump.

pump/Roots blower combination. It can be seen that much higher pumping speeds are achieved at low pressures when the pump/blower combination is employed. This is the reason that you will almost always see a Roots blower in combination with a compression-expansion pump, as shown in Figure 60, with the blower exhaust positioned at the entry to the vacuum pump. The purpose of the blower is two-fold in this application. First, it increases the overall gas handling capacity of the system over and above that of the vacuum pump alone. Secondly, it reduces pump down time in the system by maintaining the gas pressure at the entry to the vacuum pump at values significantly higher than those in the process chamber. Thus with the blower/pump combination—when the pressure in the chamber being pumped down is below 10 millitorr for example, and the pumping speed of the vacuum pump at that pressure is 1 or 2% of its value at pressures above one Torr—the addition of a blower at the mouth of the vacuum pump raises the effective pressure to values at which the pump retains relatively high pumping speeds.

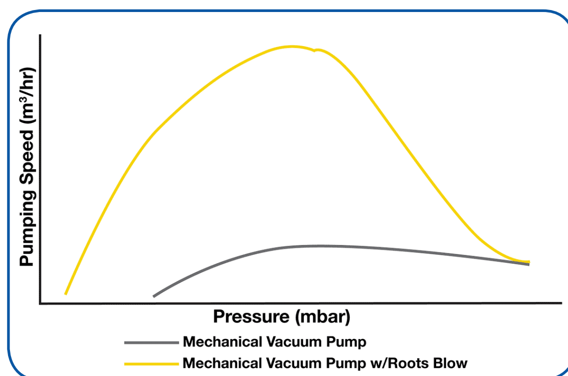


Figure 59. Typical pumping speed versus pressure curve for a rotary vane vacuum pump with and without a Roots blower (Edwards Vacuum).

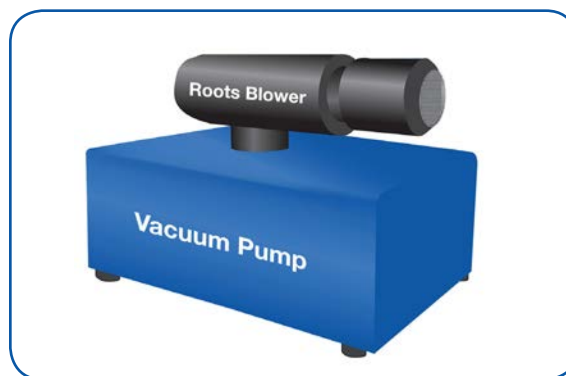


Figure 60. Typical pump/blower combination employed in semiconductor processing [182].

Turbomolecular Vacuum Pumps

Mechanical compression-expansion pumps are used in almost all systems to achieve base pressures of 10^{-4} – 10^{-3} Torr. For processes requiring lower pressures, other approaches to pressure reduction must be employed. Turbomolecular pumps are pumping systems that are effective at pressures below 10^{-3} Torr which can be employed in certain semiconductor processing/analytical equipment.

Within a turbomolecular pump (Figure 61), a number of circular discs with angled blades (referred to as rotor discs) rotate at a very high speed (up to 90,000 rpm). Between each moving disc/blade assembly is a mirror-image static disc/blade assembly known as a stator disc. The moving blades impact gas molecules within the pump and physically impart additional momentum (proportional to the blade speed) to the molecules, moving them in the direction of the stator disc. The pump is designed to ensure that the mean free path of molecules in the gas within the pump is greater than the distance between the rotor disc blades and the stator disc blades. This assures that the momentum imparted by the rotor is not lost in molecular collisions, but rather moves the gas molecules through the stator blades. Here they encounter another rotating blade that moves them towards another stator further down the pump and closer to the exhaust. The gas is thus compressed through a number of stages in the turbomolecular pump until it finally exits the turbomolecular pump and is drawn away by the backing mechanical pump. Depending on the model, turbopumps can handle gas loads of up to a few thousand liters/second and are capable of reducing chamber pressures to around 10^{-10} Torr. For a greater, in-depth understanding of turbomolecular pumps, refer to the literature [180] [181] and numerous commercial sites on the Internet.

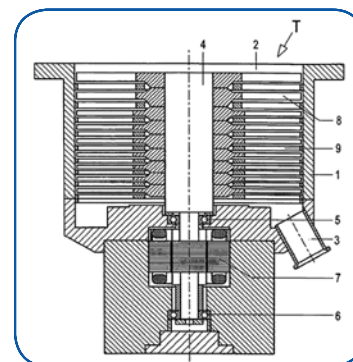


Figure 61. Cutaway view of a turbomolecular pump [183].

Cryopumps

Cryopumps are vacuum pumps that capture (as opposed to transport) the gas in a process chamber, depositing the condensed gas on internal surfaces in the pump that have been cooled to cryogenic temperatures (below 120K). Cryopumps can be cooled by liquid helium, liquid nitrogen or by a stand-alone cryocooler, depending on the temperature required by a particular pumping application. Like the turbomolecular pump, cryopumps can only operate under low pressure conditions, typically less than 10^{-3} Torr. The pumping mechanism is the steady depletion of gas molecules within the volume of the pump as they are deposited on the cold surfaces of the cryopump. This promotes continuous diffusion of gas molecules from the chamber being pumped into the cryopump, thus evacuating the chamber. Cryopumps can pump all gases from a process chamber, including noble gases, so long as the temperature is low enough to condense the gas. It provides a perfectly clean approach (i.e., no oil or other contaminants that might backstream into the process chamber) for the generation of high vacuum conditions between 10^{-3} and 10^{-9} Torr. Figure 62 shows a cutaway view of a typical cryopump structure. Unlike mechanical vacuum pumps and turbomolecular pumps, cryopumps are not capable of continuous pumpdown of a chamber. Cryopumping is, of necessity, a batch process since the pump must be periodically "regenerated" to remove the build-up of condensed gases. During the regeneration cycle, the cryopump is isolated from the process chamber and its temperature is raised to room temperature or higher to convert all of the condensed material to the gas phase so that it may be pumped away. A mechanical backing pump is required to reduce the system to the cryopump operating pressure and to remove built-up gases during the regeneration cycle.

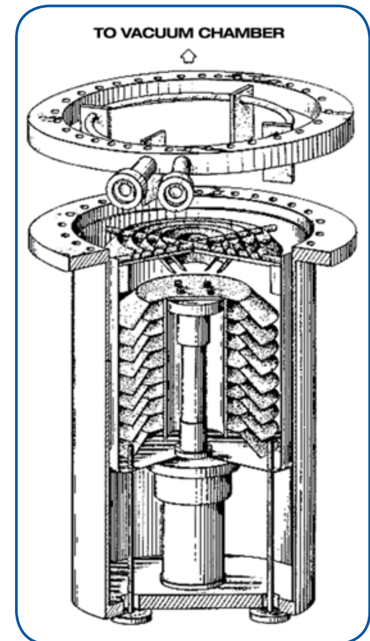


Figure 62. Cutaway view of an industrial cryopump [184].

MKS Product Applications in Vacuum Production

While MKS is not directly involved in the production of vacuum pumps, it does have robust products for roughing, foreline and ultra-high vacuum (UHV) applications. MKS's products in this area include medium and UHV vacuum components and fittings; heater jackets for vacuum components; and custom fabrication of vacuum systems.

MKS Medium and UHV Vacuum Components and Fittings

MKS offers a line of stainless steel, aluminum and brass vacuum components. Series 23 ButtWeld components and fittings (Figure 63) include 45° and 90° elbows, crosses, tees, reducers, reducing tees, end caps, metal hose and tubing. All Series 23 components are made of type 304 stainless steel and are specifically prepared for use in high vacuum process or UHV applications. They possess the following properties critical to vacuum and UHV applications:

- Low outgassing under vacuum
- Chemically resistant
- Annealed for consistent leak-free welding
- Minimized corrosion concentration points

Figure 63. Series 23 ButtWeld Components.

MKS also produces a line of ISO-KF and ISO-Universal vacuum components and fittings that are suitable for medium vacuum applications. The ISO-KF system (Figure 64) refers to small, quick release vacuum flange components (up to 2") that are manufactured to DIN 28403 and ISO 2861 standards. These fittings are suitable for use in vacuum applications down to pressures of about 10^{-6} Torr and up to pressures of about 1.5 atm. MKS Series 31 ISO-KF components and fittings provide the basic building blocks to produce all necessary vacuum piping for the following applications:

*Figure 64.
Series 31 ISO-KF Vacuum Components.*

- Roughing and foreline pumping
- Small systems
- Systems requiring frequent cleaning
- Research systems in which configurations change frequently

MKS Series 76 ISO-Universal products are similar to ISO-KF but specifically designed for larger applications such as connection to piping for pumping systems and cleanrooms, roughing, and foreline pumping for large systems and high vacuum conductance interconnects.

*Figure 65.
Series 88 CF UHV Vacuum Components.*

In addition to products aimed at medium vacuum applications, MKS produces the Series 88 CF UHV high purity stainless steel components and fittings (Figure 65) which are suitable for use in ultra-high and extremely high vacuum (XHV) applications. These are a full line of components, similar to those described above, but employing metal sealed CF vacuum flanges and fittings. CF flanges consist of a soft metal gasket and knife edges on the mating faces of the flange that cut into the metal to provide an extremely tight vacuum seal. The use of metal gaskets as the vacuum seal also makes possible a high temperature bake-out of the vacuum system, up to 500°C. Such bake-outs are the only way to remove adsorbed moisture and other volatile materials from the internal surfaces of a vacuum system and achieve UHV/XHV conditions in the system. With appropriate system preparation and bake-out, Series 88 products can accommodate vacuum systems that operate at base pressures below 10^{-13} Torr. Series 88 CF UHV components are suitable for use in the following applications:

- Difficult applications with corrosive, radioactive, toxic or air sensitive materials
- Materials creation and processing such as molecular beam epitaxy
- Analytical instrumentation including Auger electron spectrometers, mass spectrometers, and electron microscopes
- High energy physics installation such as accelerators and fusion devices

MKS Heater Jackets for Vacuum Components

Good and repeatable vacuum conditions can only be achieved in a system when volatile materials such as adsorbed gases and moisture and/or more complex volatile chemical compounds are removed from the interior ("wetted") surfaces in the vacuum system. If not, these volatile materials will desorb from the surfaces at a slow rate, constantly contaminating the vacuum environment and preventing the system from achieving its ultimate base pressure.

Adsorbed volatile materials can be removed from vacuum-wetted surfaces by the application of heat. MKS's advanced heaters and heater control systems (Figure 66) are well-suited for use in a variety of medium vacuum applications including ALD, CVD and etch in semiconductor device fabrication as well as many other deposition applications. The heaters are thin and conformal, use less energy than conventional heat tracing/insulation installations, and have superior control of temperature uniformity.

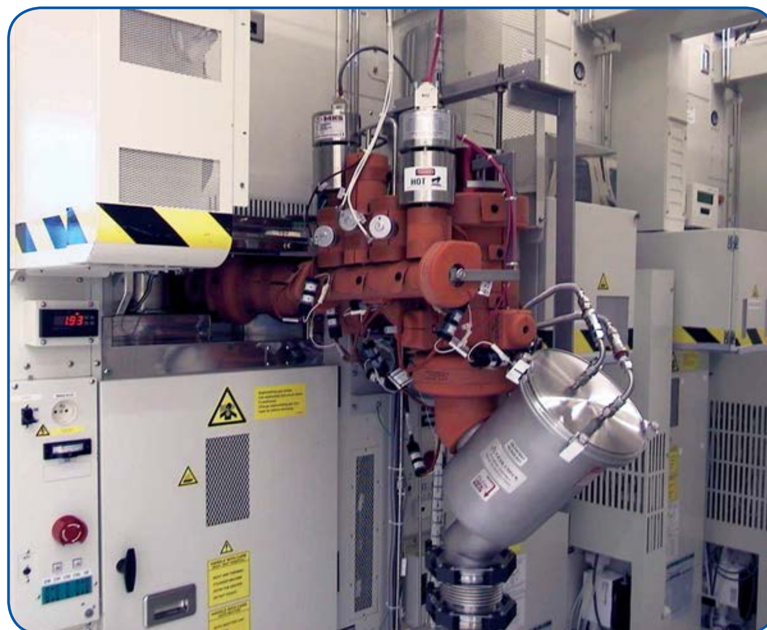


Figure 66. MKS Heater Jackets (orange) installed on a CVD system.

C. Measuring Vacuum

The reader is referred to two excellent treatises on vacuum pressure measurement that have been made available from NIST [185] and the National Physical Laboratory in the UK [186].

Vacuum Measurement Principles

Direct Measurement

The first true pressure measurement was made by Evangelista Torricelli in the 17th Century, when he invented the mercury barometer and measured atmospheric pressure. The mercury barometer is simply a glass tube that is filled with mercury and placed upside down in a dish of mercury without allowing any air to enter (Figure 67). What Torricelli found was that the force that the atmosphere exerts on the mercury in the dish (at sea level) is sufficient to support a column of mercury in the tube that is 760 mm high. The weight (force) of the column of mercury exactly counter-balances the pressure (force) that the atmosphere exerts on the mercury in the dish. If the pressure on the surface of the mercury in the dish is less than normal atmospheric pressure, the mercury column will have a height that is less than 760 mm, since the force on the mercury in the dish will be less. If there was no gas pressure exerted on the mercury in the dish (i.e., if it were in a vacuum) the level of the column would be equal to the level of mercury in the dish. So, vacuum – i.e., pressures lower than atmospheric pressure – has been historically divided into 760 units, based on the mercury height in a manometer. This experiment is the reason that today we use a unit of measurement for vacuum pressure that is named after Torricelli – the Torr – and why a Torr is exactly 1/760th of normal atmospheric pressure at sea level. A derivative of the Torr often used in semiconductor process vacuum measurements is the millitorr or 1/1000th Torr. Pressures lower than 1.0 millitorr are usually expressed as scientific notification, (e.g., 1.0×10^{-6} Torr). Vacuum and meteorological measurements in the European and Asian systems usually refer to pressures in “atmospheres” where 1 atmosphere (referred to as 1 bar) is the normal atmospheric pressure at sea level. Vacuum measurements in this system are usually reported in terms of 1/1000th’s of an atmosphere (the millibar). The Pascal, another unit often used for pressure measurement is the SI unit of pressure measurement, equal to 1 Newton/m^2.

Since Torricelli's time, liquid filled manometers have remained in use as a fundamental standard for absolute vacuum pressure measurement, with many customized designs having been developed [187]. Since liquid manometers make a direct and absolute measurement of vacuum and pressure, they are often considered as the fundamental measurement standard to which measurements made by other kinds of pressure measurement devices can be referenced. Many kinds of vacuum measurement devices, commonly called vacuum gauges, have been developed since Torricelli's time.

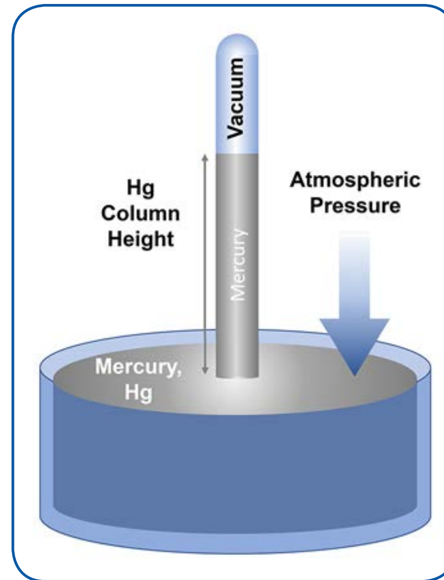


Figure 67. The mercury manometer.

Mechanical Deformation

Probably the most common type of vacuum gauges encountered in technological environments are those that employ mechanical deformation as the underlying principle for measuring pressure. These include diaphragm, bellows, and Bourdon gauges. Since pressure is a measure of force per unit area, pressure has the ability to deform different kinds of material elements in a reproducible way. The degree of deformation that an element undergoes is proportional to both the material properties of the element and to the pressure exerted on it. In this way thin, flexible elements can be used to measure low pressures (thicker, stiffer ones can be similarly used for measuring high pressures). The degree of deflection of these elements can be measured in a variety of ways, including direct mechanical measurement, variation in electrical properties of a device containing the element and deflection of optical probes. Figure 68 shows some of the structures that can be employed as deformation elements in a pressure measurement device. Mechanical deformation gauges include capacitance manometers, Bourdon gauges, resonant diaphragm gauges, bellows gauges, piezoelectric gauges, and others.

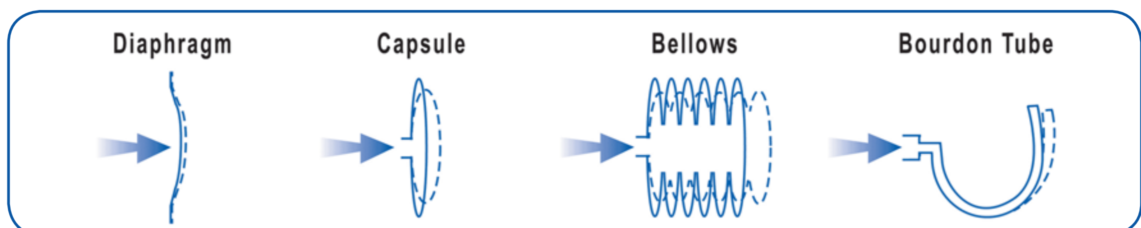


Figure 68. Common mechanical deformation elements used in pressure measurement devices and their mode of deformation with changing pressure [186].

Capacitance manometers detect the deflection of a diaphragm using changes in capacitance between the diaphragm and a powered electrode. They are commonly used for pressure/vacuum measurement. Strain-based gauges are a variant on this approach that are commonly encountered in positive-pressure (i.e. greater than atmospheric pressure) applications. Strain gauges employ a thin diaphragm with a strain sensing electronic circuit mounted on its backside. A change in pressure causes the diaphragm to deflect producing strain that is detected by the sensor.

Gas Density Measurements

Some vacuum gauge designs are based on a response to gas density and some species dependent molecular property such as specific heat. For instance, in thermal conductivity gauges, gas pressure is determined by measuring the energy transfer from a hot wire to the surrounding gas. The heat is transferred into the gas through molecular collisions with the wire and the frequency (and therefore the degree of heat transferred) of these collisions is dependent on the gas pressure and the molecular weight of the gas molecules. These gauges only exhibit simple proportionality between pressure and heat transfer at relatively low pressures, with typical measurement ranges lying between 10^{-4} and 10 Torr. Thermal conductivity gauges include thermocouple, thermistor and Pirani gauges. They are generally relatively inexpensive and reliable. Figure 69 provides a representation of the physical configuration of a Pirani gauge head and the electrical circuit used for pressure measurement. The measurement uses the thermal element as one arm of a Wheatstone bridge. Using the bridge, changes in the electrical resistance of the element can be measured and these changes are proportional to its temperature. Using this information and the known electrical characteristics of the element, the heat transfer can be calculated and related to the gas pressure. Thermocouple and thermistor pressure gauges are very similar to Pirani gauges with the primary difference being that the hot wire temperature is measured directly using either a thermocouple or thermistor attached to the element.

Another type of density measurement involves ionization of the surrounding gas. The principle of operation for one variety of this type of gauge, the hot cathode gauge, is shown in Figure 70. A filament (the cathode) emits electrons by thermionic emission and a positive electrical potential on the grid accelerates these electrons away from the filament. The electrons oscillate through the grid until they eventually strike either the grid or a molecule of gas. When an electron impacts a gas molecule, a positively charged cation is created that is accelerated toward and collected by a negative electrode known as the collector. The electrical current created in this manner is directly proportional to the number of ions that are created in the gas phase which, in turn, is directly proportional to the gas pressure.

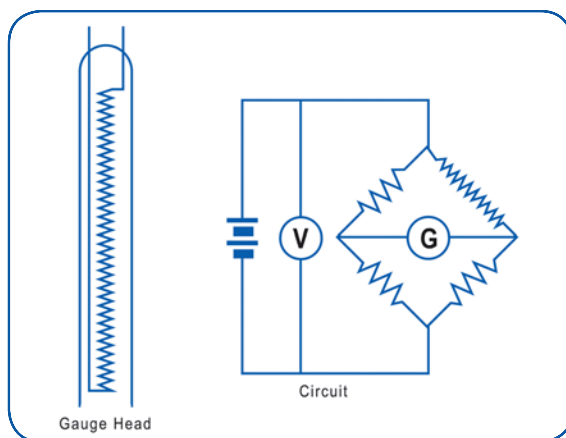


Figure 69. Physical and electrical configuration of a Pirani thermal conductivity gauge [186].

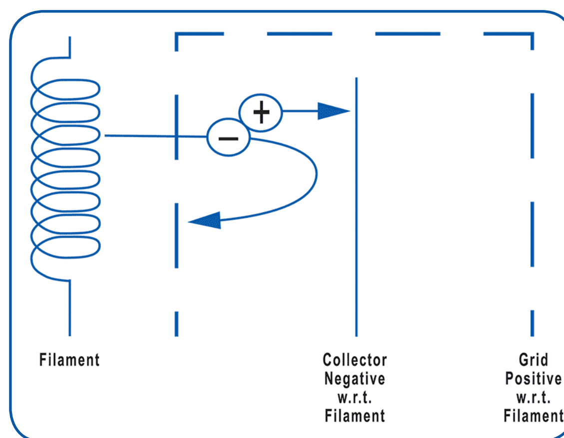


Figure 70. Operating principle for a hot cathode ionization gauge [188].

Partial Pressure Measurements

Residual gas analyzers (RGAs) are often used to measure pressures, especially in the semiconductor industry where they are commonly available for other process purposes. RGAs are compact mass spectrometers that can detect and analyze residual gases within a vacuum system. Within the RGA, incoming gases are ionized in a manner similar to that used in ionization gauges and the ions are mass filtered using an electromagnetic quadrupole. The quadrupole can be controlled to create a variable electromagnetic field that can be swept to detect ions of different mass to charge ratio. After exiting the quadrupole, the ions are detected using a Faraday cup or other more sensitive detectors. A detailed discussion of the operating principles and characteristics of RGAs is discussed in Section E. For additional information, see Bibliography reference [172].

Other Measurement Principles

Other specialty methods have been used to measure vacuum pressures in the UHV to XHV pressure regime. These are variations of ionization gauges.

Major Vacuum Gauge Types

The selection of a measurement method for a given vacuum application depends largely on the expected vacuum environment and the degree of accuracy required for the measurement. For example, a gauge that is expected to measure process pressures during a deposition or etch must, by necessity, be exposed to the process gases. This may or may not impact the measurement accuracy and it may cause physicochemical problems for the components of the gauge, depending on the type of gas that is employed in the process. If a vacuum gauge is needed for the measurement of base pressure in a system in which the chemical nature of residual gases is unknown, gauges in which the measurement has a dependence on the molecular properties of the gas being measured may not be suitable. In a manufacturing environment, it is also important that gauge selection be cost effective. In the case of vacuum gauges, depending on the design and associated ancillary equipment, gauge costs can vary by orders of magnitude. Once the specifications for the vacuum gauge range, accuracy, process compatibility, and cost have been defined, it is important to select the gauge that best matches these specifications. Table 15 provides a comparison of the different gauge types for relative performance in several areas along with the relative cost for such gauges. This section will provide more detail on the vacuum gauge types commonly found in semiconductor device fabrication applications.

	Mechanical Gauges	Thermal Gauges	Strain Gauges	Capacitance Gauges	Ionization Gauges	Spinning Rotor Gauges
Pressure/Vacuum Range (Torr)	0 to 500 atm	10^{-4} to 760	0 to 500 atm	10^{-4} to 500 atm	10^{-9} to 0.01	10^{-7} to 1
Gas Sensitivity	None	High	None	None	High	High
Accuracy	Poor to Good	Fair	Fair to Good	Excellent	Fair	Excellent
Main Power	No	Yes	Yes	Yes	Yes	Yes
Electrical Output	No	Yes	Yes	Yes	Yes	Yes
Corrosion Resistance	Poor	Fair	Good	Excellent	Poor	Poor
Physical Size	Good to Poor	Good	Excellent	Excellent	Fair	Fair
Overall Safety for Personnel	Fair	Poor	Fair	Good	Fair	Good
Relative Cost	Fair to Excellent	Good	Fair to Excellent	Fair	Fair	High

Table 15. Performance and relative cost of major pressure gauge types.

Because of the almost universal requirement for pressure sensors that can provide signals for process control and automation in semiconductor manufacturing, simple mechanical gauges such as the Bourdon gauge and bellows gauge find only limited use in semiconductor device fabs and will not be discussed further.

Thermal Conductivity Gauges

Only the Pirani gauge will be described in this section since thermocouple and thermistor gauges find limited application in the semiconductor fab. Pirani gauges were first developed in the early 1900's. As noted earlier, the sensing element is a wire of known resistance and known temperature coefficient of resistance (i.e., how its resistance varies with temperature). The element forms one leg of a balanced Wheatstone bridge. When gas molecules collide with the heated element, they extract heat from it, changing its resistance which unbalances the bridge relative to its reference state. Since the number of collisions and hence the amount of heat transferred to the gas is proportional to the gas pressure, the voltage required to maintain the bridge in balance is proportional to the pressure.

Pirani gauges have limitations that must be recognized when they are used. Since different gases have differing abilities to transfer heat away from the filament (i.e., different specific heat capacities), the Pirani gauge response depends on the gas that is present in the system. Thus, users must calibrate the Pirani gauge for the expected residual gas in the system. At the same pressure, a system that has been purged with helium will yield a different Pirani gauge reading than one purged with nitrogen (see Figure 71, Figure 72). Since the Pirani element operates at temperatures of between 100 and 150°C, care must be taken that reactive gases that can break down and deposit solid material on the element are excluded from the gauge. Should such deposits occur (or should corrosive gases react with the element), the accuracy of the Pirani will be diminished. Since the Pirani depends on heat dissipation from the element by the surrounding gas, the Pirani loses accuracy when the pressure drops below about 10^{-4} Torr where heat transfer is significantly reduced. At high pressures (10 Torr and above), the mean free path of gas molecules becomes reduced to a point where nonlinearities enter into the pressure-voltage relationship and reduce the gauge's sensitivity. Advanced Pirani gauges are constructed so as to permit convective forces within the gauge to assist molecular flow. These latter designs enable Pirani gauges to be used with good accuracy up to 760 Torr pressure. Pirani gauges are normally received calibrated for nitrogen and calibration curves are required for use with other gases. Pirani gauges are relatively fast with response to pressure changes occurring in a tenth of a second or less [189].

The relatively low cost, small size and good sensitivity of Pirani gauges make them the gauge of choice for routine vacuum measurements in the low to medium vacuum range. Their low accuracy (5% of reading) relative to other types of vacuum gauge limits their use to non-critical applications. Their sensitivity to reactive gases means that they must be isolated from process gas streams.

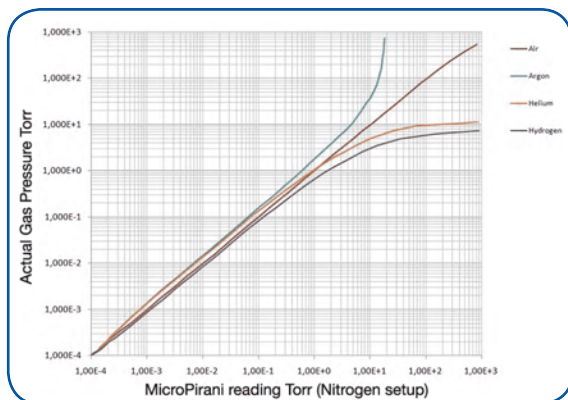


Figure 71. Pirani gauge actual pressure versus reading on a nitrogen calibrated gauge for various gases.

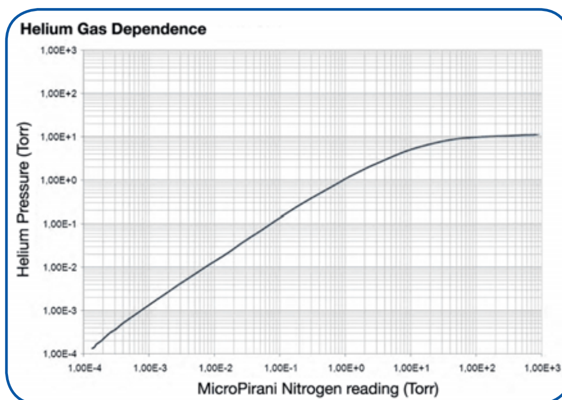


Figure 72 The response of a Pirani gauge to helium versus its response to nitrogen.

Capacitance Manometers

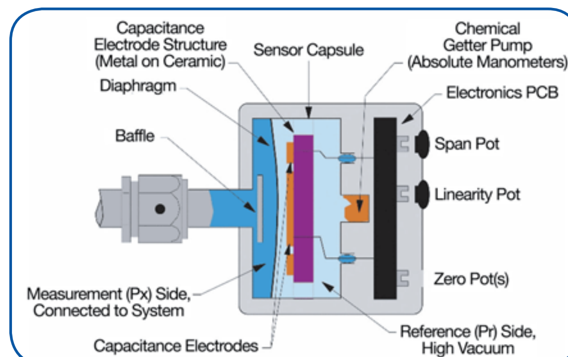


Figure 73. Capacitance manometer schematic showing the internal components and functional zones.

Capacitance manometers are electro-mechanical gauges that measure both pressure and vacuum. The capacitance gauge translates a pressure-modulated movement in a thin diaphragm into an electrical signal proportional to the pressure. Figure 73 shows the major components of a commercial capacitance manometer. The pressure sensor is the thin diaphragm that is exposed to the pressure or vacuum being measured via the inlet tube. An electrode, typically a ceramic disk with conductive pathways, is mounted in the reference cavity behind this diaphragm and is powered with an electrical signal (either AC or DC). Pressure differences between the process and the reference cavity deflect the diaphragm slightly, changing the distance between it and the electrode. Changes in this distance produce variations in the capacitance between the diaphragm and the electrode and this variation in electrical signal is proportional to the pressure change. This signal is amplified by the on-board electronics and output to the device's electrical connector for transmission to pressure indicators and process controllers.

Since the diaphragm deflection is proportional to force (pressure), pressure measurements by the capacitance manometer are independent of the composition of the gas being measured. Thus, a capacitance manometer's output doesn't change if the gas composition in the process changes. In many chemical and thin film processes the chemistry can change on a second-by-second basis; capacitance manometers provide a critical advantage in such processes. Capacitance manometers typically read pressure or vacuum over a five-decade range; i.e., from 1000 Torr to 0.01 Torr.

There are two basic types of capacitance manometer sensors:

- *Absolute* capacitance manometers in which the reference cavity is evacuated to high vacuum so that the pressure measurements will always be referenced to vacuum
- *Differential* capacitance manometers that do not have a reference cavity – just a tube or passage that can be connected to any pressure or vacuum source. These manometers read the *difference* in pressure between the inlet tube and the reference cavity at the backside of the electrode. They are routinely used as safety switches and for airflow pressure drop measurements

Capacitance manometers are frequently used as reference devices for other types of gauges (i.e., they are the calibration standard against which other gauge types are measured). An unheated 1000 Torr capacitance manometer has an accuracy specification (which includes repeatability) of about 0.25% of reading. This can be compared with an accuracy of 5-25% of reading for a Pirani or thermocouple gauge of the same range – a ratio of 100X.

NOTE: Thermal or mechanical gauges should NEVER be used to calibrate a manometer – the manometer is far more accurate.

Capacitance manometers are vacuum measurement workhorses in the semiconductor industry. Because of their insensitivity to gas composition absolute capacitance manometers are found in almost every semiconductor process tool where they are used to monitor in-process pressures. Differential capacitance manometers find broad application in areas requiring pressure-based switching and control such as chamber load locks.

Ionization Gauges

UHV and XHV pressure measurements are routinely performed using ionization gauges configured as either hot cathode gauges (HCGs) or cold cathode gauges (CCGs). Both sensor types determine pressure from measurements of an ion flux created by collisions between energetic electrons and residual neutral gas molecules within the gauge (see our earlier discussion). HCGs employ thermionic emission from a filament as a source of electrons while CCGs use a circulating space charge to create a free electron plasma.

When the pressure of a vacuum system is below about 10^{-4} Torr, direct pressure measurement methods such as capacitance manometers and Pirani gauges no longer work well and it becomes necessary to use methods that basically count the number of molecules in the gas space. This latter quantity is called the number density and is commonly given the symbol *n*. In terms of the number density, the pressure can then be calculated as:

$$p=cnT,$$

where *p* is the pressure, *c* is a constant and *T* is the temperature. Ionizing the gas and collecting the ions is a convenient way of establishing the number density in the gas phase since there is a well-defined statistical relationship between the collisional dynamics that create ions and the number density of molecules in a gas and hence the pressure of that gas. Ionization gauges perform this measurement using free electrons that have been produced by either thermionic emission or plasma generation. The ion current that is collected by the negatively biased collector can be related to pressure once the gauge has been calibrated. The basic gauge equation for ionization gauges is:

$$I_C = KnI_e,$$

where I_c is the ion current, *K* is a constant relating to the ionization probability, *n* is the number density of gas molecules and I_e is the ionizing electron current.

Figure 74 shows a typical Bayard-Alpert (B-A) hot cathode ionization gauge. It has three electrodes: the cathode or filament, the collector and the anode grid. Energetic electrons emitted from the cathode (biased at about 30 Vdc) are accelerated towards the anode grid (typically 180 Vdc) colliding with and ionizing molecules present in the gas phase. The positive ions that are created in the collisions are accelerated towards the collector (0 Vdc) located along the axis of the anode grid producing an ion current that is measured by the gauge electrometer. This gauge configuration produces a strict, linear relationship between collector current and pressure. HCG measurements are gas-dependent since different atoms and molecules exhibit different ionization efficiencies (due to factors such as electron-molecule collision cross-section and ionization potential). They are usually calibrated for either nitrogen or argon and attention must be paid to the relationship between the gauge calibration gas and process gas protocols when selecting a HCG. Gas correction factors are available that permit field adjustment of HCG readings.

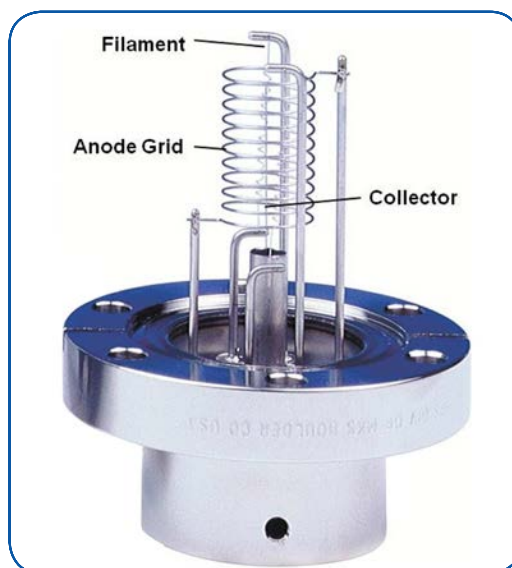


Figure 74.
Typical Bayard-Alpert hot cathode ionization gauge.

The useful operating range of measurement of the HCG is defined, in the upper limit, by restrictions on the mean free path of the ions when the number density in the gas phase becomes too high. At high number densities, an ion has a high probability of interacting with either a free electron or a neutral molecule before reaching the collector. When this happens, the initial electron-molecule ion creation event no longer contributes to the ion current, introducing inaccuracies in the relationship between that current and the system pressure. In practical terms, this upper limit for the use of ionization gauges is about 10^{-3} Torr. Historically, the lower limit of HCGs was considered to be reached when either electrical leakage currents in the gauge head or the controlling electronics become comparable with the ion current being measured. However, advances in modern electronics have extended the lower limits of these gauges into the XHV regime. In HCGs, the lower limit is the so-called X-ray limit. In an HCG, X-rays are emitted when electrons impact the grid. A small fraction of these X-rays are intercepted by the ion collector in the HCG where they cause electrons to be photoelectrically ejected. This photoelectron current leaving the collector is electrically detected as identical to the positive ion current, adding to the current and producing a false pressure reading. This phenomenon results in the standard lower limit for pressure measurement using the best conventional HCGs having a value of around 10^{-11} Torr.

With good operating protocols and proper calibration, the sensor-to-sensor reproducibility of Bayard-Alpert HCG readings is typically about 2%. Repeatability of readings is 1-2%, limited mainly by uncontrollable random sensitivity variations.

Care must be taken to avoid exposure of the working components of an HCG to process gases while the HCG is operating (other than purge gases such as nitrogen and argon). Reactive gases can damage gauge components when they are at operating temperature, either through unwanted deposition or through corrosion. Even when the gauge is not "on," process and ambient gases will adsorb on the components of the HCG and these materials will slowly desorb from the gauge surfaces during normal operation, increasing the local pressure within the gauge. "Degassing" can be used to drive the molecules adsorbed on the gauge surfaces back into the process chamber, where they can be pumped out of the system. Degassing can be done as required or as part of a regular pumpdown sequence. Regular degassing helps prevent process deposits from collecting and ensures that the HCG provides the most sensitive and repeatable pressure indications.

Cold cathode gauges, also known as Penning gauges (after their inventor), have been in use since the late 1930's. CCGs differ from HCGs in that they employ high voltages between the electrodes (in the kV range), in place of a hot filament, to generate free electrons and a crossed magnetic field to create a controlled electrical current path in the gauge. CCGs rely on randomly emitted electrons caused by cosmic rays, a radioactive source, and field emission to initiate stable electron plasma within the current path of the gauge. The discharge within a CCG normally ignites within a very short time at 10^{-6} Torr or above, a few minutes at 10^{-8} Torr, and longer times at lower pressures. In operation, the crossed fields trap a near-constant circulating electron current in this path. Electrons moving within the conduction path collide with residual gas molecules, producing the ion flux that is measured. Figure 75 shows a schematic that can help in understanding the ion generation mechanism within CCGs. In the Figure, Electron 1 (whose path begins on the left side of the circle) is accelerated by the electric field in approximately cycloidal jumps around the anode at a constant radius until it encounters a neutral molecule. If it has sufficient energy, the electron causes an ionization event which produces a positive ion and an additional electron (2) which undergoes acceleration in the conduction path in a manner similar to electron 1. The ion generated in the collision moves to the cathode where its impact generates a secondary electron which similarly enters into the conduction path. Through

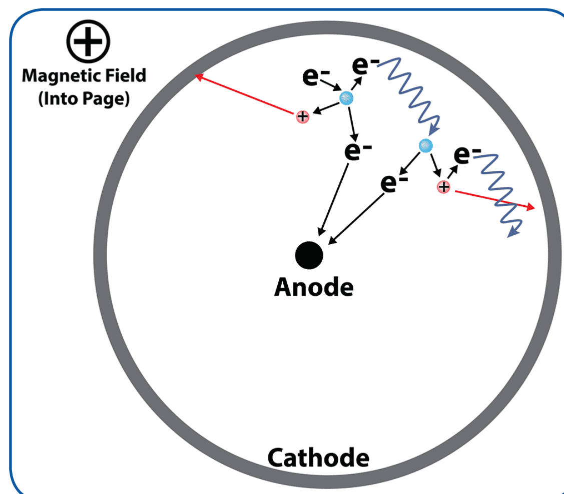


Figure 75. Electron path and ion generation in an inverted magnetron sensor.

these mechanisms, the circulating electron current quickly builds to a stable value, determined by the fields and geometry within the sensor. Once a critical electron density is reached, space charge depression of the electric field within the device prevents any further production of electrons from the cathode. The circulating electron current becomes very stable and nearly independent of pressure and surface conditions.

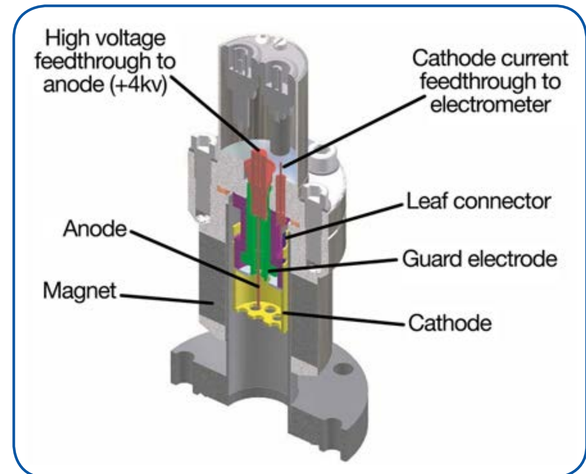


Figure 76. Cutaway view of an MKS inverted magnetron cold cathode sensor.

Since its invention, the cold cathode gauge has undergone numerous evolutionary developments that have extended its range and accuracy. The inverted magnetron sensor geometry is now the most commonly used for CCGs, owing to its wide range and freedom from spurious effects. The advent of inverted magnetron designs enabled the development of compact CCGs capable of measurement below 10^{-10} Torr. Other improvements, such as isolation of the collector and the use of guard electrodes have further broadened the spectrum of application for CCGs. Today, CCGs mated with simple, reliable controllers can be used in a wide variety of vacuum measurement applications, especially those that require a combination of ruggedness, durability and high sensitivity/accuracy. Figure 76 shows a cutaway view of MKS's inverted magnetron cold cathode sensor design.

Cold cathode gauges exhibit upper limits on pressure measurements similar to those observed for hot cathode gauges. Typically, CCGs are effective for measuring operating pressures that range between 10^{-10} and 10^{-2} Torr. As with HCGs and for similar reasons, CCG readings are gas dependent. Because of the different operating principle, however, CCGs require different gas correction factors than those for HCGs. Repeatability for CCGs is not normally as consistent as that of HCGs, with typical values of about ±5% and reproducibility sensor-to-sensor of 20-25% is not uncommon. For high accuracy operations, CCGs are normally calibrated against a transfer standard such as a spinning rotor gauge or a high accuracy HCG.

Spinning Rotor Gauges

Spinning rotor gauges (SRGs) also known as molecular drag or viscosity gauges, measure the number density of the surrounding gas. Like Pirani gauges, SRGs measure pressure by measuring the transfer of energy from a sensor to the surrounding gas. In the case of the spinning rotor gauge, a small steel ball is magnetically levitated within a non-magnetic tube that is mounted horizontally and connected to the vacuum system. Figure 77 shows a schematic of the SRG. In the measurement procedure, the ball is spun to a few hundred hertz using a rotating magnetic field then the drive field is turned off and the rate of deceleration of the ball is measured using magnetic sensors. Since the deceleration is caused by the transfer of energy from the ball to gas molecules during collisions, it can be related, using gas kinetic theory, to the number density of the gas and thus to the pressure. Like Pirani and ionization gauges, SRGs are sensitive to the gas species. SRGs are often used as a reference gauge for calibrating other gauge types.

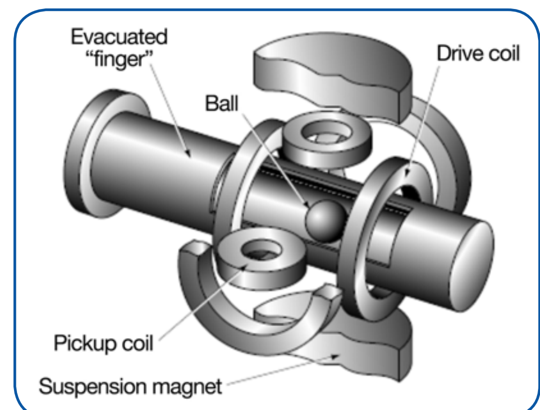


Figure 77. The spinning rotor gauge (reproduced with permission from the Institute of Measurement and Control) [186].

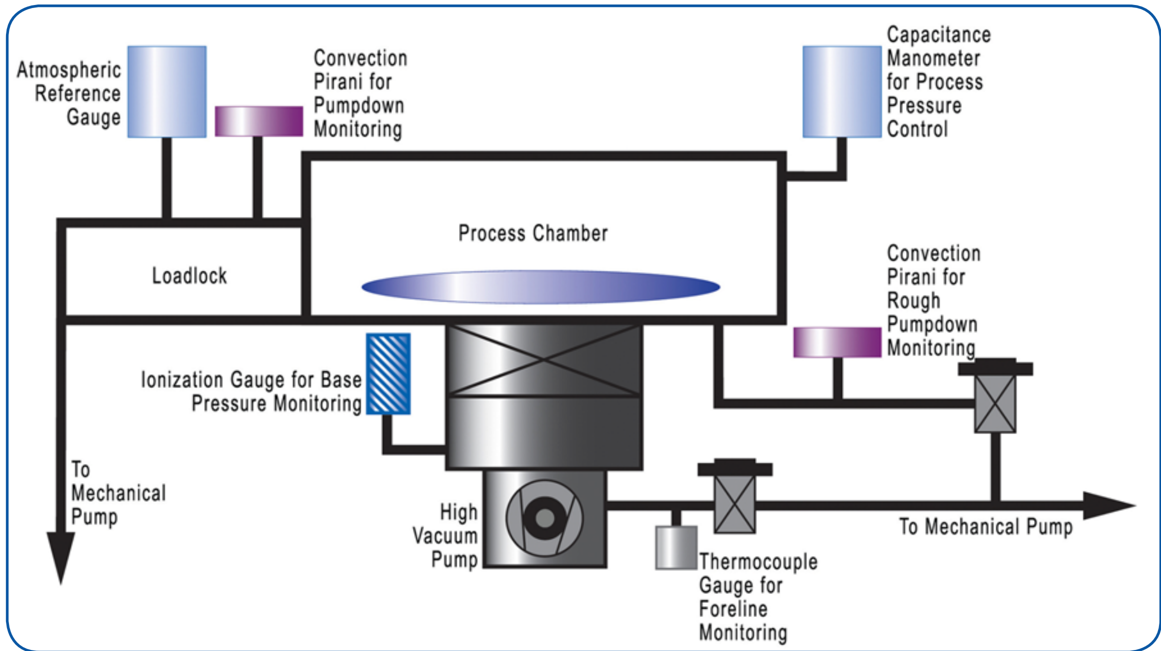


Figure 78. Applications for different vacuum gauge types in a semiconductor process tool.

Vacuum Gauge Applications in Semiconductor Processing

Vacuum gauges are critical components of many semiconductor processing tools, especially deposition and etch tools. Figure 78 provides a simple schematic to help understand vacuum gauges in a semiconductor process and the applications that are available.

MKS Product Applications for Vacuum Pressure Measurement

MKS offers a wide array of vacuum gauges for application in semiconductor and other industries. These products are detailed in the product overview available on the MKS website, entitled: "Advanced Vacuum Measurement Solutions". MKS products cover essentially all of the semiconductor industry's vacuum gauge requirements with products ranging from capacitance manometers to spinning rotor gauges.

MKS Baratron® Capacitance Manometers

MKS Baratron® capacitance manometers are an ideal choice for measuring pressure and vacuum in many industrial, medical, semiconductor, LED, and critical thin film applications, including semiconductor device manufacturing, optical coating, flat panel display, and solar cell manufacturing processes. They are extremely accurate and stable over long periods of time. MKS Baratron manometers are constructed from corrosion resistant materials, and are insensitive to the types of aggressive process gases, such as halogens, typically used in semiconductor etch processes. Standard Baratron manometers such as the one shown in Figure 79 measure absolute pressure and are widely used for stand-alone pressure sensing as well as being

Figure 79. Baratron® Capacitance Manometer.

Figure 80. 226A Differential Capacitance Manometer.

integrated into process control equipment. Baratron manometers for typical semiconductor applications have full scale pressure ranges from 20mT up to 1000Torr. They come in a variety of application-specific configurations including both unheated and heated versions. MKS also supplies capacitance manometers that measure differential pressure (Figure 80), as well as ancillary equipment for the use of these manometers in process settings, including power supplies, readouts and other accessories. MKS also provides calibration subsystems and services.

MKS, Granville-Phillips® Pirani and MicroPirani™ Gauges

MKS, Granville-Phillips® produces the 275 Convectron® Pirani vacuum gauge (Figure 81) and Series 475 gauge controller offering exceptionally stable performance for low vacuum measurement. Unlike traditional thermocouple and Pirani gauges that use only conductive heat loss, Convectron® gauges take advantage of heat loss due to convection at higher pressures to extend the range of accurate, repeatable measurement to atmospheric pressure. MKS Granville-Phillips also offers the Convectron system as Mini-Convectron® modules with built-in controls and a DeviceNet™ interface. Key features of the Convectron gauge include:

Figure 81. Convectron® Vacuum Gauge.

- Vacuum pressure measurement from atmosphere to 1×10^{-4} Torr
- Individually calibrated gauges ensure the highest measurement performance
- Easy installation in space-restricted locations
- Wide selection of vacuum fittings simplifies installation
- Compatible with several of MKS's vacuum gauge controllers
- Temperature compensation offsets the effects of ambient temperature changes
- Rugged construction

MKS, Granville-Phillips also offers a line of modular vacuum gauges and transducers that are based on Micro-Electro-Mechanical System (MEMS) design and which have built-in controllers and extremely small footprints. These micro- and mini-transducers and modules are ideal for size-critical applications and include a variety of sensors and transducers. We will describe two representative examples of these transducers, beginning with the 925 MicroPirani™ vacuum transducer (Figure 82). This transducer has an extended range of 10^{-5} Torr to 1 atmosphere and three set-point relays for process control. It can be used in a variety of applications including:

Figure 82. 925 MIcroPirani™ Transducer.

- General vacuum pressure measurement
- Foreline and roughing pressure measurement
- Gas backfilling measurement and control
- Mass spectrometer control
- Activation of UHV gauges
- System process control (interlocks, valves, pumps)
- Control system pressure

The second example is the 910 DualTrans™ MicroPirani Piezo vacuum transducer. It comes in the same form-factor as the 925, however, it contains two gauges in one package: a MEMS MicroPirani sensor and a Piezo sensor. This combination extends the measurement range of this transducer to measure

mks

pressures from 1.0 x 10^{-5} to 1500 Torr. This wide range allows the 910 to be used in vacuum chamber applications requiring absolute vacuum/pressure measurement with switching capabilities.

The full line of sensors and transducers based on MEMS MicroPirani™ technology is detailed on the MKS website for all Granville-Phillips indirect vacuum gauges.

MKS, Granville-Phillips® UHV Gauges

MKS, Granville-Phillips offers a line of Bayard-Alpert hot cathode ionization gauges including conventional glass-sealed and nude gauges, all-metal Stabil-Ion® gauges, and the world's smallest Micro-Ion® gauges.

MKS, Granville-Phillips Series 274 Bayard-Alpert ionization gauges (Figure 83) are available in three configurations: glass tubulated; nude with electron bombardment (EB) degas; and nude with I^2R (resistive) degas. The 307 Bayard-Alpert vacuum gauge controller and the 350 Bayard-Alpert UHV vacuum gauge controller offer a number of options for control of one or more of these gauges. Figure 83 compares the tubulated and nude Bayard-Alpert ionization gauge configurations. All series 274 gauges provide excellent vacuum measurement with comparable accuracy. The nude gauge with EB degas offers somewhat lower ultimate pressure measurement capability with a lower X-ray limit of about 2 x 10^{-11} Torr. The glass tubulated gauge and nude gauge with I^2R degas have a slightly higher lower limit of 3-4 x 10^{-10}. Both types of gauge have an upper pressure measurement limit of about 2 x 10^{-2} Torr. The Series 274 gauges are configured for high reliability, with a bifilar grid and either a burn-out resistant filament or dual tungsten filaments. These gauges find ready acceptance in both research and industrial applications, with proven performance and reliability characteristics.

*Figure 83.
274 Series B-A
Ionization Gauges.*

MKS, Granville-Phillips 370 Stabil-Ion sensor (Figure 84) and 370 Stabil-Ion vacuum gauge controller products are the most accurate hot cathode ionization gauges available. These gauges are housed in an all metal tube that eliminates glass breakage, decomposition and helium permeation. The metal housing acts as an electrostatic shield, protecting the gauge from external electric fields. The combination of metal housing, an internal design in which the physical relationship between the grid and the filament has been significantly stabilized relative to other B-A designs and significantly improved control electronics result in exceptional reliability and accuracy for this gauge. The 370 Stabil-Ion has an accuracy when calibrated for nitrogen, of 4% of the reading and a repeatability of 3% for measurements between 1 x 10^{-8} Torr to 1 x 10^{-4} Torr. By comparison, many other B-A gauges have reading inaccuracies in the 30-40% range. The range of the 370 Stabil-Ion gauge (when equipped with a Convectron® option) is 2 x 10^{-11} to 999 Torr.

*Figure 84.
370 Stabil-Ion® B-A
Vacuum Gauge.*

MKS, Granville-Phillips also produces a line of very small B-A ion gauges and gauge controllers for size critical applications, the Series 355 Micro-Ion Bayard-Alpert vacuum gauge (Figure 85) and the 358 Micro-Ion vacuum gauge controller.

Having only 10% of the volume of a conventional gauge, these are the world's smallest hot cathode ionization gauges. However, they offer comparable range, accuracy and repeatability to most full size B-A gauges. Their range when calibrated with either air or nitrogen is 5 x 10^{-10} to 5 x 10^{-2} Torr (to 1000 Torr with Convectron option). The 355 Micro-Ion achieves its high performance through the combination of a unique, patented dual ion collector / dual filament design and ultra-clean construction. Also available is the Series

*Figure 85.
355 Micro-Ion® B-A Gauge
showing external and
cutaway views.*

354, 392, 390 Micro-Ion modules product line that combines Micro-Ion ionization gauge technology with a miniature Pirani Conductron® heat-loss sensor to provide accurate, continuous pressure measurement from high vacuum to atmosphere.

In addition to Bayard-Alpert hot cathode ionization gauges, MKS, Granville-Phillips supplies a cold cathode vacuum gauge system, the Series 943 which includes either the 431 or 423 I-Mag® cold cathode sensors (Figure 86). The sensors are based on an inverted magnetron design with proprietary features that improve the accuracy relative to other CCGs. The 431 sensor has a higher operating temperature with customized models available for research applications. The 423 I-Mag sensor is designed as a lower cost, more compact and rugged sensor suitable for OEM and industrial applications. These CCGs can be used in the following applications:

*Figure 86.
431 and 423 I-Mag®
Cold Cathode Sensors.*

- High energy physics
- Analytical instruments
- Laser production
- Ion implantation
- Mass spectrometry
- PVD

MKS also produces the SRG-3 Spinning Rotor Gauge (Figure 87) for applications ranging from metrology to use as an in situ standard on high vacuum process chambers to ensure accuracy of vacuum readings from other methods. The high accuracy and repeatability of the SRG method makes it a reference standard that is used in a variety of circumstances. The SRG-3 system offers the user consistent traceability as required by ISO9000.

*Figure 87.
SRG-3 Spinning Rotor
Gauge System.*

MKS, Granville-Phillips® Multipurpose Vacuum Gauge Controllers

MKS, Granville-Phillips offers a number of options for the combined display of different vacuum gauges and types. These control options can be found on the Controllers for Vacuum Gauges and Transducers section of the website.

MKS, Granville-Phillips® RGA Partial Pressure Analyzers

MKS, Granville-Phillips offers the 835 Vacuum Quality Monitor (VQM™) mass spectrometer (Figure 88) for the measurement of partial pressures in vacuum systems. This system uses an autoresonant ion trap mass spectrometer gauge rather than a quadrupole mass analyzer. When coupled with a total pressure gauge such as a Stabil-Ion Bayard-Alpert gauge, the VQM provides a measurement of the partial pressures of each gas component in the vacuum system. The VQM can detect gases such as H_2, H_2O, CO_2 and CO at pressures as low as 10^{-14} Torr. The MKS website provides a useful instructional video for this system.

*Figure 88.
835 VQM® System.*

D. Vacuum/Pressure Control

Stable and precise control of a processing system's vacuum/pressure is critical for high-yield semiconductor device fabrication. Processes such as SAPCVD, LPCVD and etch exhibit optimal behavior at well-defined process pressures and it is critical to maintain and transition process pressures in a well-controlled, stable manner. Similarly, advanced processes such as ALD must have tight control over system pressures during gas switching steps. The required vacuum/pressure control in these and other semiconductor unit processes is accomplished using closed-loop control for a number of variables that affect the vacuum process.

To understand the issues involved in vacuum/pressure control in semiconductor process equipment, consider the simplified schematic for a hypothetical process control system in a deposition or etch tool as depicted in Figure 89 (for the sake of simplicity, plasma processing is not considered). The Figure shows the equipment components with connections denoting the primary inputs and outputs used by an automatic process controller for the different control functions in the system. The process controller in this particular instance requires vacuum/pressure information from the different vacuum gauges depicted in Figure 89 as input for automated gas switching, purge/reactant gas flow control, wafer transfer, and vacuum/pressure control.

Vacuum/pressure gauges provide key information that is used to trigger different actions in a process sequence. Consider the following (much simplified) process recipe steps for a hypothetical deposition process performed in the process tool shown in Figure 89.

1. Load substrate in loadlock and seal loadlock
2. Pump down loadlock to rough pump base pressure
3. Equalize loadlock and process chamber pressures
4. Transfer substrate to process chamber and seal chamber
5. Purge process chamber and pump to mechanical pump base pressure
6. Open valve to high vacuum pump and bring process chamber to high vacuum base pressure
7. Isolate high vacuum pump
8. Open isolation valve to mechanical pump
9. Establish purge gas flow at expected process pressure
10. Bring chamber and process to process temperature setpoint under purge gas flow
11. Switch from purge gas to process gas
12. Run deposition process for pre-set time
13. Shut off process gas and pump to rough pump base pressure
14. Cycle purge process chamber to remove residual process gas
15. Pump chamber to rough pump base and equalize loadlock and chamber pressure
16. Transfer substrate to loadlock and seal process chamber
17. Isolate loadlock and backfill to 1 atm with purge gas
18. Remove finished substrate

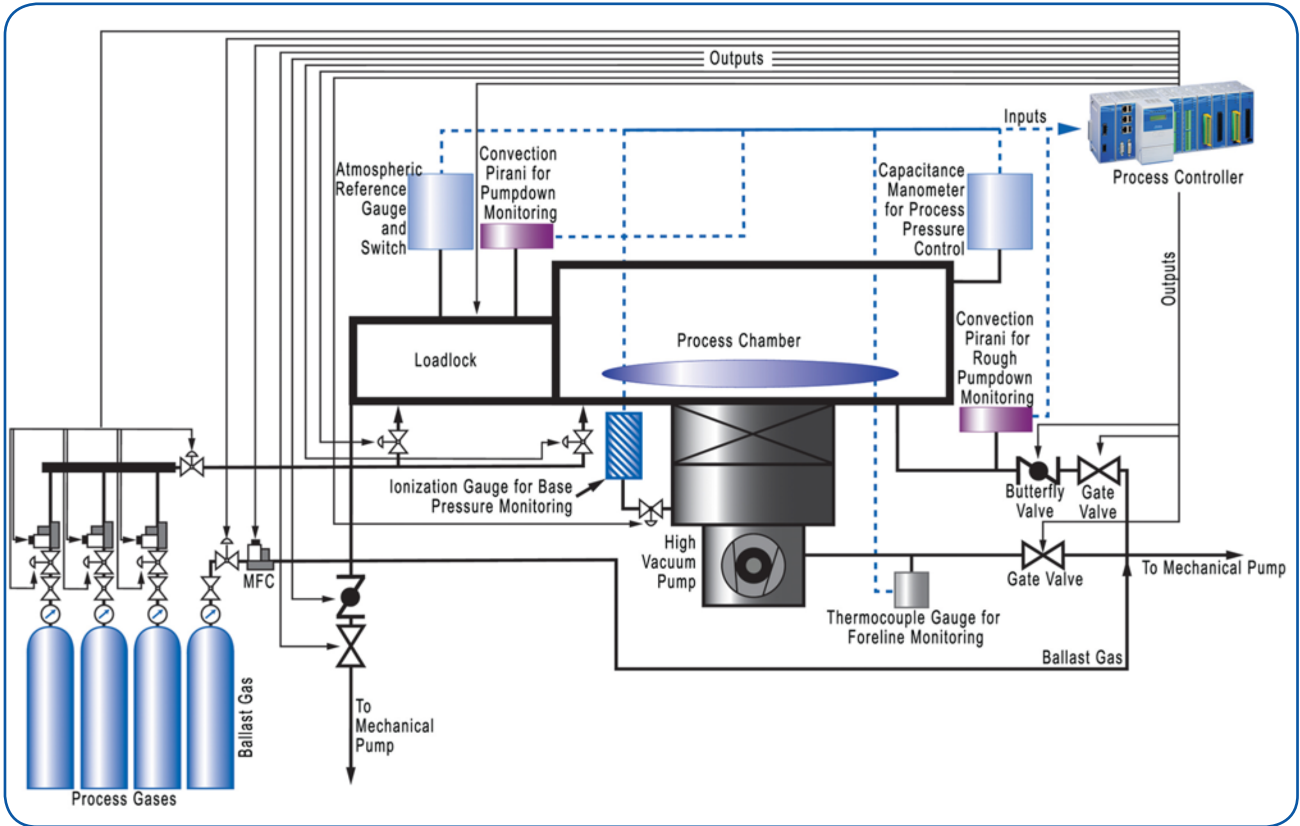


Figure 89. A hypothetical deposition or etch tool showing vacuum sensing and control elements.

It is easy to see that there are many times in this process sequence when the process controller needs information on the pressure in either the loadlock or the process chamber in order to perform a process step. The simple task of loading the substrate into the loadlock requires that the pressure in the loadlock be equal to or greater than that in the surrounding fab environment, otherwise atmospheric pressure would prevent the outer loadlock door from opening. A differential pressure gauge comparing the loadlock and atmospheric pressures can tell the controller if the pressures are equal. If not, the controller can add additional purge gas to bring the pressures into equilibrium. Once equilibrium is established, a simple switch can release the loadlock door and allow substrate loading.

Since overall control of the process is not our focus in this section, we will only give further consideration to steps in which vacuum/pressure control is needed. Step 2 in the process above requires that the loadlock be pumped from atmospheric pressure down to the base pressure of the rough pumping system which is typically around 10^{-3} Torr. If the process simply opened a valve between the loadlock and the mechanical pump, the loadlock would quickly pump down to base pressure. However, the pneumatic shock and accompanying turbulence would disturb any particles in the loadlock, contaminating the substrate surface. This is avoided by adjusting the pumping speed on the loadlock from relatively low at the beginning of the pumpdown to full speed once the loadlock pressure is close to base pressure. Once the system is under vacuum, pneumatic shocks during purge and process gas switching are normally avoided through the use of precise and controlled adjustments of the gas flow rates. Steps 9 and 10 in the process sequence also require vacuum/pressure control. In these steps, the gas flow and pumping speed must be adjusted to maintain an optimal pressure during the deposition process. While there is some relationship between gas flow and process pressure, the optimal process pressure does not usually occur at the optimal gas flow values and these two variables must be controlled separately.

Semiconductor processing systems commonly use one of three main approaches to vacuum/pressure control in process steps such as those discussed above:

1. Control the pumping speed by changing rotational speed of a mechanical blower
2. Control the pumping speed by adjusting the flow rate of ballast gas feed between the process chamber and pump
3. Control the pumping speed by adjusting the vacuum conductance of the vacuum downstream foreline using a throttling valve

	Blower Speed Control	Ballast Gas Control	Downstream Control
Dynamic Range	Low – 10:1 Typical	Moderate – 500:1 typical, 1000:1 max	High – 1000:1 typical, 10,000:1 max
Types of Pumps	Will not work with all pumps	Works with any pump that can operate at process pressure	All
Speed of Response	Moderate	Fast	Fast
Initial Capital Cost	AC or SCR motor controller	Bypass valve and optional controller	Exhaust valve and controller
Extra Operating Costs	None	None	None
Susceptibility to Effluent Gases	None	None	Slight
Special Requirements	Pump controller and flow rates must be properly sized	Bypass valve and pump must be properly sized	None

Table 16. Pressure control techniques.

Automatic closed-loop control of the vacuum/pressure is accomplished by monitoring the process pressure using, most commonly, a capacitance manometer which provides a feedback signal for chamber pressure to a process control system. This system uses PID (Proportional Integral Derivative), a common combination of control algorithms or some other kind of control algorithm to generate corrections to the pressure control setpoints in one of the control schemes detailed above (1 to 3 above).

Table 16 compares the different approaches to vacuum/pressure control over a number of operational and cost parameters.

MKS Product Applications for Vacuum/Pressure Control

MKS's approach to vacuum/pressure control focuses on the equipment needed for vacuum isolation (next section) and downstream closed-loop control of each variable that affects a vacuum process (subsequent sections).

In the downstream control approach, an exhaust throttle valve is opened or closed, changing the conductance to the vacuum pump in order to achieve and maintain the desired process pressure—this separates the process variables of gas flow and process pressure, making them independent of one another within the limits of acceptable process conditions. Downstream pressure control provides high dynamic range, works well with all types of vacuum pumps, has a fast response time, is tolerant to most effluent gases, and has moderate initial costs.

MKS's product line for vacuum/pressure control allows the user to isolate and maintain closed-loop control of the process pressure through the use of a throttling control valve and digital PID or self-tuning pressure controllers. A typical pressure control system that uses MKS components works as follows:

1. A Baratron capacitance manometer senses pressure in a vacuum chamber
2. This pressure is compared to the desired set point pressure in the pressure controller
3. The pressure controller commands the control valve to open or close, changing the chamber pressure and bringing it to the desired process set point

The following products enable this approach to vacuum/pressure control.

MKS supplies detailed product configurations for reliable pressure management and downstream pressure control.

MKS Vacuum Isolation Valves

Figure 90. Vacuum Isolation Valve models.

Integral to any vacuum/pressure control scheme is the ability to isolate a process system from the vacuum source (pump). MKS produces a broad spectrum of application-specific vacuum isolation valves (Figure 90). The MKS family of valves includes: bellows sealed valves, ball valves, soft start dual-stage valves, UHV valves, and safety shut off valves. Valve actuation includes manual, pneumatic, and electromagnetic actuators. They are extremely reliable with lifetimes of up to 1,000,000 cycles under clean conditions. Specialty valves for harsh process conditions include heated valves that prevent by-product condensation and deposition, and corrosion resistant valves for applications such as metal etch.

MKS Downstream Valves and Pressure Controllers

MKS supplies downstream throttling control valves in a number of configurations, both as stand-alone units with integrated controllers and as units configured for remote control. Many MKS downstream pressure control valves are available in heated options.

Stand-alone units include T3Bi high speed valves (Figure 91), and "smart" exhaust throttle valves such as the 153D and 683B valves (Figure 92). T3Bi valves are high speed throttling valves that integrate all control, communication, and driver circuits within the throttle valve assembly, eliminating the need for a separate pressure control electronics module. The T3Bi self-tuning control algorithm and high-speed operation drives the system to set point quickly and with minimum overshoot, ensuring repeatability in process recipes. RoHS compliant, the T3Bi valve accepts Baratron manometer inputs and accommodates multiple communications protocols for ease of integration into local control schemes. T3Bi valves are programmable for pressure, position or set point limits. T3Bi valves are particularly well suited for "house exhaust" or atmospheric pressure throttling applications. The 153D and 683B valves are based on the standard MKS 253B exhaust throttle valve with add-on electronics that eliminate the need for separate control modules. The 153D and 683B valves can be configured for either pressure control or flapper position control. They exhibit linear valve transfer characteristic for smooth, linear pressure control. They are specifically designed for computer-controlled applications where a simple pressure control system is desired.

MKS also supplies conventional throttle valve and remote controller configurations in the 253B exhaust throttle (Figure 93) and 653B high speed exhaust throttle valve products with associated controllers such as the 651C digital/analog pressure controller (Figure 94). The MKS Type 253 is available in standard sizes and flange styles, and is compatible with all MKS throttle valve controllers. Type 653 valves are available in a variety of sizes and flange styles, and are compatible with MKS Type 651 and 1651 controllers.

MKS provides an excellent overview in the Downstream Pressure Controllers and Valves product selection guide.

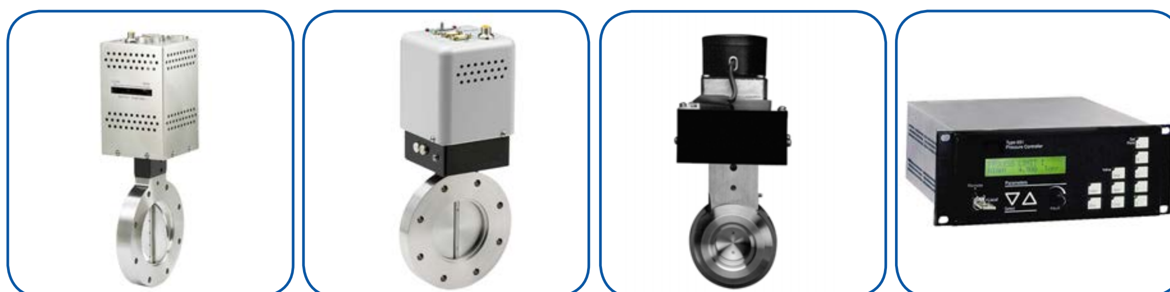


Figure 91.	*Figure 92.*	*Figure 93.*	*Figure 94.*
T3Bi Throttle Valve.	*683B "Smart" Throttle Valve.*	*253B Exhaust Throttle Valve.*	*651C Pressure Controller.*

MKS Upstream Valves and Pressure Controllers

MKS supplies upstream gas flow control valves that can be interfaced with appropriate MKS controllers and pressure sensing valves to control chamber pressure by regulating the gas flow into the chamber. Both conventional valve/controller and fully integrated valve/controller configurations are available.

Conventional upstream valve and valve controller combinations include the baseline 248D flow control valve, (Figure 95), the 148J all-metal control valve, and the 154A high flow valve which can be controlled using the MKS 651 and 1651 pressure controllers. These valves can be controlled using the MKS 250E pressure/flow control module or the 1249 solenoid control valve driver. The latter control valve driver is intended for use with PCs, PLCs or other controllers that provide the necessary PID control functions but do not have the required current output to drive a flow control valve.

MKS supplies integrated upstream pressure valves and controllers for closed-loop electronic pressure control. These are self-contained compact, closed-loop electronic control systems used for upstream or downstream pressure control. MKS's 640A and 641A absolute and gauge pressure controllers (Figure 96) contain a Baratron capacitance manometer, a normally-closed proportional control valve, and closed-loop control electronics. Pressure output and input control signals in these controllers are linear 0-5 VOC or 0-10 VDC. Two trip points are included in the 640/641, with LED status indicators, for use as simple on/off process limits. The 640 and 641 pressure controllers can be interfaced with a single- or four-channel power supply/readout or with the Type 647 multichannel gas flow and pressure controller.

In addition to single zone integrated pressure controllers, MKS offers the DPC dual-zone pressure controller (Figure 97) that integrates an inlet pneumatic shut-off valve, two independent channels of pressure control with mass flow metering, and a vacuum outlet. This was designed to reduce the overall cost of ownership of pressure control subsystems for backside wafer cooling, specifically for the latest two-zone electrostatic chucks.

MKS also offers the πPC series PC90 integrated closed-loop pressure controller and PC99 integrated closed-loop pressure controller (Figure 98) that includes a mass flow meter. The πPC pressure controller provides digital control in a web-enabled format. It is a self-contained, compact, closed-loop electronic pressure control system that is designed for a wide range of pressure and flow conditions. It contains a Baratron capacitance manometer, normally closed or normally open control valve, and closed-loop control electronics. The πPC is available in either an upstream or downstream pressure control configuration, making it well suited for controlling process chamber backpressure or process gas delivery pressure. With either digital (DeviceNet or RS-485) or analog I/O and its Ethernet setup/diagnostics capabilities, the πPC is easily integrated into most process tools. The πPC pressure controller is a complete, compact pressure control package that minimizes cost and space requirements.

Figure 95.
248D Flow Control Valve.

Figure 96.
640 Series
Pressure Controller.

Figure 97.
DPC Dual-zone
Pressure Controller.

Figure 98.
πPC PC99
Pressure Controller.

MKS Vacuum/Pressure Switches

MKS offers several different models of vacuum/pressure switches (Figure 99) for accurate and reliable protection of vacuum equipment, atmospheric switching, and vacuum/pressure processes. Designed for applications where a DC signal output is not required, these switches provide relay outputs that are readily interfaced with alarms, valve actuators, computers, process controllers, load locks and other protection devices.

Figure 99.
Vacuum/Pressure Switch.

MKS Vacuum/Pressure Calibrators

A transfer standard is a secondary device that is traceable to a national laboratory or other recognized standards body (e.g., National Institute of Standards and Technology, NIST). Transfer standards are used in both

laboratory and production environments. MKS's PVS6E vacuum calibration system provides NIST-traceable calibrations over pressures ranging from 10^{-5} to 1000 Torr and can be used to calibrate capacitance manometers, thermocouple gauges, Pirani gauges, convection enhanced Pirani gauges, other capacitance manometers and transmitters, and mechanical/dial gauges. Hot or cold cathode ionization gauges can be calibrated over the upper end of their range. Portable variants of this system (the PBTS1A and PBMS2B Portable Baratron Transfer Standards) with somewhat reduced functionality are also available. In addition, the SRG-3 spinning rotor gauge can be used as a traceable standard for calibration of pressure measurements.

E. Residual Gas Analysis

Residual gas analysis is the analysis of low levels of residual gas that remain in a vacuum chamber following pump down. Such gases may include residues from ambient air such as oxygen, nitrogen, water, along with process residues that may include reactants such as silane, organometallics, halides, etc., and reaction products such as ammonia, hydrogen, moisture, etc. Vacuum chambers can also contain contaminants due to system leaks and hydrocarbons that backstream into the process chamber from the pumping system. Reproducible process results from a vacuum process are only possible when the system is guaranteed to be leak-tight and when the kind and quantity of residual gases in the process system are known and accounted for in the process protocol. Residual gas analyzers (RGAs) are a critical analytical tool for gaining this knowledge.

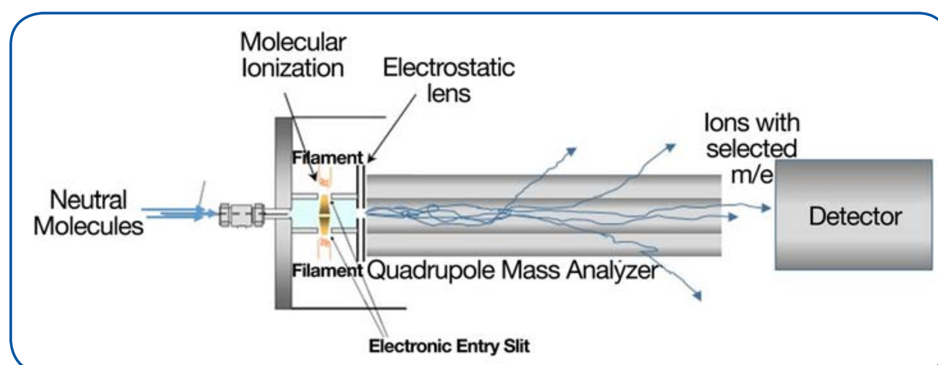


Figure 100. Functional components of a quadrupole mass spectrometer.

Modern RGAs utilize quadrupole mass spectrometry as their underlying operational principle. Quadrupole mass spectrometers (QMS) can monitor multiple gas-phase species in real-time. They have a wide dynamic range and can track gas-phase reactant, product and contaminant concentrations from ppb to % levels. A QMS detects and quantifies chemical species by first ionizing the molecules and/or atoms in a gas stream and then separating these ions by mass/charge (m/e) ratio. A quadrupole analyzer requires high vacuum for operation since ions must pass freely through the analyzer, affected by only the analyzer field. The operating pressure within a QMS is typically below 10^{-6} Torr and this is maintained using a turbomolecular pumped vacuum system, usually on-board the RGA. Depending on the pressure of the system being analyzed, the inlet to the QMS may be a simple vacuum connection (if the system being analyzed is under medium to high vacuum) or a differentially pumped capillary tube (to limit the volume of gas entering the QMS and avoid raising the internal pressure of the RGA). When the RGA is in operation, a gas sample stream from the chamber being analyzed enters the QMS at the ion source (from the left in Figure 100). Various methods may be used to ionize the sample, with electron-impact ionization being the most common. Electrons emitted by a hot filament knock electrons off the incoming molecules and atoms to form positive ions. The ions are separated according to their m/e ratio by a quadrupole mass filter and detected by a Faraday plate or secondary electron multiplier. Readers desiring a greater understanding of how the quadrupole mass analyzer works are directed to references such as [186] [171] and, especially, [190]. All of the gas components are analyzed within a single scan and the results of the analysis are displayed as a mass spectrum (plot of m/e

versus peak intensity). The ionization process produces both parent ions (the original molecule with a positive charge) and ionized fragments of the original neutral molecules; the fragmentation pattern acts as a fingerprint that identifies the various neutral species that were present in the gas stream sample. If the identities of the major components are known, specific mass peaks can be monitored and calibration coefficients applied to derive the relative composition for each species present. Owing to the fundamental nature of the species being measured, the QMS data is relatively simple to interpret. Furthermore, because all species within a given sample are analyzed, mass spectrometers can detect and highlight the presence of unexpected species or contaminants in a product gas stream.

Some newer RGAs employ an ion trap as a mass analyzer rather than a quadrupole. The theory of operation of these systems is quite complex and well beyond the scope of this introductory work. Refer to [190] pages 100-122 for an in-depth discussion of these systems.

A number of compact, accurate and simple to use RGA systems are commercially available. MKS's RGA product line offers a variety of levels of RGA utility, from simple partial pressure analyzers to complex mass spectrometers.

MKS Product Applications for Residual Gas Analysis and QMS Process Monitoring

MKS, Granville-Phillips® RGA Partial Pressure Analyzers

MKS, Granville-Phillips offers the 835 Vacuum Quality Monitor (VQM®) mass spectrometer (Figure 101) for the measurement of partial pressures in vacuum systems. This system uses an autoresonant ion trap mass spectrometer gauge rather than a quadrupole mass analyzer. When coupled with a total pressure gauge such as a Stabil-Ion Bayard-Alpert gauge, the VQM provides a measurement of the partial pressures of each gas component in the vacuum system. The VQM can detect gases such as H_2, H_2O, CO_2 and CO at pressures as low as 10^{-14} Torr. The MKS website provides a useful instructional video for this system.

Figure 101.
835 VQM® System.

MKS QMS Products

MKS offers a number of configurations within its Vision and Cirrus™ gas analysis product lines. These may be configured as RGAs, RGA Process Monitors, and Atmospheric Pressure Gas Analyzers.

Process Monitors

In-depth Application Notes for MKS RGA process monitors can be found on the MKS website (RGA vacuum process monitors.) MKS Vision and HPQ products utilize the TOOLweb® process-specific automated control and monitoring platform for RGAs and other sensors on semiconductor process tools.

MKS's Microvision 2 (Figure 102) "smart head" technology collects data on potentially damaging residual gases that negatively impact the process quality at millisecond speed. The electronics unit mounts directly onto the analyzer head, and connects to the system PC via Ethernet. It is designed as an application specific tool to monitor pumpdown, baseline, and leakback conditions in semiconductor and thin film processes.

The Vision 2000-C™ (Figure 103), Vision 2000-E™, and Vision 2000-P™ RGAs combine a closed ion source with an automated inlet to enable monitoring of the complete CVD, PVD or Etch process cycle, from base vacuum to process pressures of up to 700 Torr. These RGAs incorporate Microvision 2 "smart head" technology and are integrated with the Process Eye™ Professional control platform. Process Eye is designed for process monitoring applications where a flexible control platform is required to achieve automation and full integration with a process tool. Process Eye Professional uses recipes to define the way in which the RGA scans, displays and responds to the data acquired.

MKS produces the 300mm Resist-Torr RGA (Figure 104) for the detection of photoresist residues on wafers during the high pressure (~8 Torr) wafer degas step just prior to insertion of the wafer into a 300 mm PVD chamber. The 300 mm Resist-Torr is a fully automated monitor with built-in calibration for the calculation of the PR index that measures the photoresist contamination level in the degas chamber. The Resist-Torr employs a fast response capillary sample inlet to ensure vacuum integrity in the QMS. When integrated with Process Eye and TOOLweb RGA sensor control software, the 300mm Resist-Torr system provides completely automated operation and highly reliable photoresist detection.

The MKS HPQ3 and HPQ3S (Figure 105) are designed as in situ process monitoring tools that can operate well beyond the normal 1 x 10^{-4} mbar total pressure restriction of most RGAs without the need for differential pumping of the sample inlet. The lack of requirement for differential pumping significantly simplifies the operation of these tools and reduces their footprint in the process environment, making them particularly suitable as in situ monitors in a wide range of CVD, PVD and Etch applications, from leak detection to process monitoring.

Figure 102. Microvision 2 Residual Gas Analyzer

Figure 103. Vision 2000-C™ Residual Gas Analyzer.

Figure 104. 300mm Resist-Torr® Photoresist Detection Monitor.

Figure 105. HPQ3 High Pressure Residual Gas Analyzer.

Atmospheric Pressure Gas Analyzers

MKS's Cirrus line of atmospheric pressure gas analyzers are based on quadrupole mass analyzers in which the gas sampling system has been configured to permit the analysis of a sample gas at atmospheric pressure. These analyzers accomplish this by employing an inert, silica-lined capillary inlet that restricts the amount of gas sample that is allowed to pass to the quadrupole mass analyzer chamber (which is maintained under high vacuum using a turbomolecular pump/diaphragm pump combination).

Two categories of Cirrus gas analyzers are available, the Cirrus 2 product line and the Cirrus 3-XD.

The Cirrus 2 benchtop system (Figure 106) is a state-of-the-art QMS specifically designed for high sensitivity and long-term analyzer stability. It is capable of near real-time monitoring of a wide array of gases and gas mixtures with a dynamic sensitivity range that varies between parts per billion (ppb) and percentage levels. It can be configured to sample up to 16 gas streams, with special configurations available to sample corrosive gases or to produce high mass resolution spectra. As well, it can be supplied with an integrated optical sensor for continuous carbon monoxide (CO) sensing. The Cirrus 2 incorporates an Ethernet interface, allowing it to be controlled by either a local PC or through an organizational network. It is operated using the Process Eye Professional control platform which allows for fully automated operation and calibration and which incorporates and tracks data from other process sensors.

Figure 106. Cirrus™2 Benchtop Gas Analyzer.

Cirrus 3-XD atmospheric pressure gas analyzers (Figure 107) are designed for research and engineering applications requiring trace gas analyses that are beyond the limits of normal QMS systems. The addition of patented V-lens™ ion optics technology to the Cirrus QMS platform provides a unique analytical advantage for the Cirrus 3-XD that ensures gas-independent low baseline and stable, robust detection of gases at low ppb levels. This guarantees a high level of confidence in trace level analyses. The fast response capillary inlet design allows the system to capture 250 data points per second for near real-time analyses. The Cirrus 3-XD employs the Process Eye Professional control platform for fully automated operation and calibration. Available in either benchtop or rack-mounted configurations, the Cirrus 3-XD is ideal for in-line monitoring and analysis of trace contaminants in process gases, including solvent vapors, hydrocarbons, atmospheric and inorganic gas species (including corrosives), freons, and noble gases.

Figure 107. Cirrus™3-XD Atmospheric Pressure Gas Analyzer.

MKS Software for RGA, Process and Chamber Monitoring

Process Eye Professional (Figure 108) is a network-compatible application that leverages the features and functions of Microsoft® operating systems (32-bit and 64-bit Windows® XP, Vista, Server 2008 and Windows 7). It is designed for process monitoring applications requiring a flexible control platform for automation that is fully integrated with a process tool. Process Eye Professional uses recipes to define the way in which the RGA scans, displays data, and responds to the data acquired.

EasyView (Figure 109) is the latest RGA software platform for MKS's range of RGAs. Designed specifically for simplicity in installation and operation, EasyView is an interactive, ICON-driven package with many advanced data display and data storage capabilities. It is a network-compatible application that leverages the features and functions of Microsoft® operating systems (32-bit and 64-bit Windows® XP, Vista, Server 2008 and Windows 7).

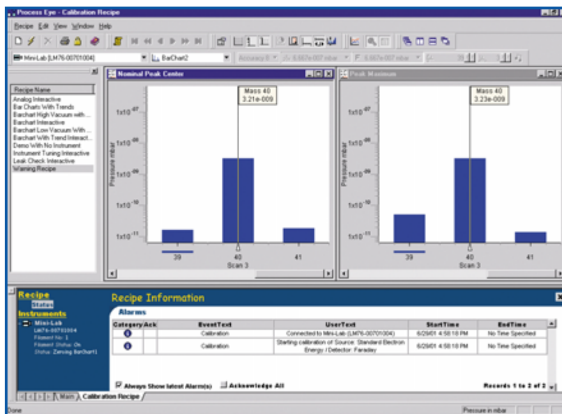


Figure 108. Process Eye™ Professional.

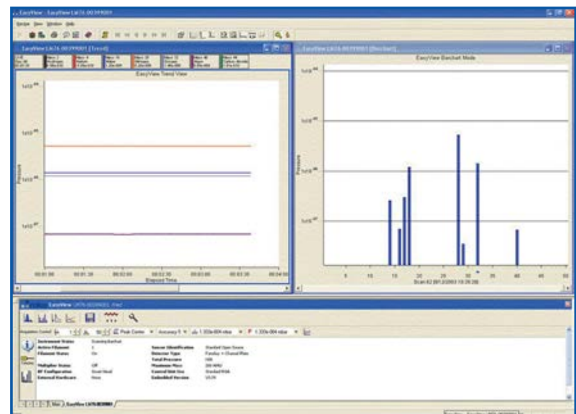


Figure 109. EasyView RGA.

TOOLweb RGA (Figure 110) is an automated platform for the control and monitoring of process-specific RGA and other general sensors on semiconductor tools. Comprised of sensor hardware, a tool connectivity module, and software for real-time control and web-based data review, TOOLweb RGA monitors the tool chamber environment both during and between periods of wafer processing and after a pump-down. It can function as either a stand-alone tool-level controller or as a component in a tool group or fab-wide installation. TOOLweb RGA enables complete reporting and alarming on both a per tool and cross tool basis with full tool and factory host integration capability.

Figure 110. TOOLweb® RGA.

II. Substrate Surface Cleaning

A. Piranha, SC1, SC2, RCA and DHF

Wet chemical cleaning and conditioning of wafer surfaces is a critical process step in most, if not all, semiconductor device fabrication schemes. The acidic cleans listed in the title of this section have a long history of use in semiconductor processing; forecasts call for their continued use into the foreseeable future. We have already discussed the chemical composition, actions and process applications of these wet cleaning methods in Section A.III.1 of this book and we will not repeat that discussion here. Readers interested in the underlying science of these cleans can reference the in-depth treatments of the fundamentals of silicon substrate surface cleaning that are available in the literature [47] [153] [154].

This section will briefly describe some of the equipment used for wet cleaning silicon substrates in modern semiconductor fabrication plants. Such equipment can range from very simple “wet bench” cleaning stations in older technologies to sophisticated automated single wafer cleaning tools employed in advanced technologies.

Reinhardt and Kern state that, “in the past, wafer cleaning and surface preparation were considered more of an art than a science.” [153]. This was especially true in the era when “wet benches” for simple batch immersion and rinsing dominated substrate cleaning. Scientific studies over the past three decades have revealed the importance of chemical concentration control over all points on a substrate surface. This is needed in order to guarantee the required material properties of the layers and interfaces in the often severe topographies of advanced devices. Batch processing can limit the degree to which such control can be achieved in high aspect ratio geometries, and single wafer approaches have been developed to solve such issues. Indeed, back end of line (BEOL) post-patterning cleaning, especially for cleaning Cu interconnect lines and high-k interlayer dielectrics, has had to move entirely away from batch wet bench processing to more controlled single wafer wet cleaning approaches. Furthermore, batch methods, while very convenient for wafer diameters up to 150mm and which can be made workable for 200mm substrates, become unwieldy when applied to substrates with diameters beyond 200mm. This too has contributed to a shift to single wafer wet cleaning technologies. Finally, batch methods are not as easily automated as are single wafer approaches. However, conventional wet benches will often be found in semiconductor fabs, either in less critical front end of line (FEOL) application areas or in R&D and process development applications.

A modern wet bench cleaning station is shown in Figure 111. Modern benches are configured with a FOUP (front opening universal pod) interface and wafer handling robotics to ensure low particulate substrate transfer into and out of the wet bench. Within the wet bench are liquid tanks in which a cassette containing substrates can be immersed in the desired cleaning solution. Cleaning solutions are typically pre- or in situ mixed within the wet bench station using ultra-high purity semiconductor grade reagents and ultrapure water (UPW). Once the substrates have been immersed in the

Figure 111. A modern wet bench cleaning station [191].

cleaning solution for a prescribed time, the chemical reaction is rapidly stopped and any chemical residues removed by a combination of high volume UPW spray rinsing and rapid overflow dump rinsing. In this latter step, the wafers are subjected to repeated cycles of a very fast hot or cold UPW immersion and draining with spray rinsing using ultrapure water. Once a cassette of substrates has been rinsed, it is quickly dried, typically using some kind of high speed spin drying. Megasonic agitation and/or isopropyl alcohol (IPA) may be employed during the rinse/dry cycle to enhance the rinsing and drying abilities of the system. Depending on the tool design, these steps may require multiple tanks and cassette spinners or all functions may be integrated within a single tank.

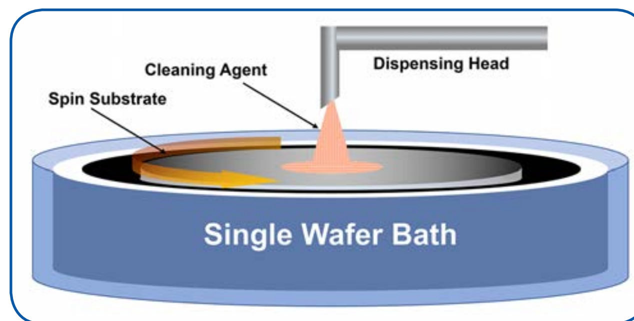


Figure 112. Automated single wafer cleaning station, basic principles [194] [195].

As noted above, advanced device designs and larger substrates cannot be cleaned using batch wet bench approaches. Over the past decade, single wafer cleaning systems have been developed that have leveraged improvements in economical spray as well as new and improved cleaning chemistries. Highly automated, single wafer treatment systems have been designed that employ multiple cleaning stations, each cleaning one wafer at a time, to maintain production-scale throughputs. These new single wafer systems incorporate significantly more precise control over localized surface chemistries than can be achieved using batch processing. Additionally, particle contamination is inherently more controllable in single wafer handling and processing systems; this is a critical aspect for low defect processing in modern equipment. Figure 112 shows a schematic of a typical automated single wafer cleaning arrangement in a commercial single wafer cleaning tool.

MKS Product Applications in Wet Substrate Surface Cleaning

While MKS does not offer a product line in either bench or single wafer substrate cleaning tools, MKS does supply equipment for producing certain chemicals used in cleaning solutions, specifically MKS's line of ozonated water delivery systems [192] [193]. DI water/ozone solutions (DIO_3) have been found to provide a clean, safe and highly effective replacement for Piranha and RCA SC-1 and SC-2 cleans in many aspects of surface cleaning.

Organics Removal and Photoresist Strip

As a strong oxidizer, ozone, O_3, rapidly reacts with most organic chemical compounds. O_3 reacts directly with hydrocarbons and also generates oxygen radicals that react even more vigorously with organics. Thus DIO_3 can be used as an effective clean that removes ambient organic molecules adsorbed on the wafer surface. More importantly, when coupled with megasonic agitation, DIO_3 is very effective in removing photoresist residues from a wafer surface.

SCROD Cleaning

While pure DIO_3 is not suitable for cleaning metals and particles from a wafer surface, cleaning protocols that combine the use of O_3 with HF and/or hydrochloric acid are highly effective for this purpose. Single-wafer spin cleaning with repetitive, alternating use of DIO_3 and dilute HF (DHF) is known as the SCROD cleaning method. SCROD cleaning oxidizes the substrate (through the action of the O_3) creating a self-limited oxide layer of about 1 nm thickness which the HF subsequently dissolves. The sequential oxidation and dissolution of the oxide layer efficiently removes both particles and adsorbed metals from the substrate surface. Originally developed by workers at Sony, SCROD cleaning has been shown to remove

87% of aluminum oxide particles, 97% of silicon nitride particles, and 99.5% of polystyrene latex particles from a substrate surface [196] [197]. This is nearly an order of magnitude more efficient particle removal than can be achieved using a standard SC-1 clean. Similarly, studies have shown that metal contamination levels on a substrate surface of less than 1 x 10^9 can be achieved using SCROD approaches [197] [198].

Advanced Reticle Cleaning

Traditional methods (RCA) for reticle cleaning degrade their optical properties and reticles can only be cleaned between 2 and 8 times before unacceptable degradation occurs due to surface roughening [199] [200]. Since ozone-based chemistries produce very little surface roughening, they are becoming preferred in reticle cleaning applications.

MKS LIQUOZON® Dissolved Ozone Delivery Systems

MKS offers the LIQUOZON® product line containing several different configurations for the generation and delivery of DIO_3:

LIQUOZON® PrimO$_3$ Ozonated Water Delivery System

The LIQUOZON® PrimO$_3$ system is designed specifically for application in multi-chamber single wafer cleaning tools. Configuration options that permit the system to be mounted in the subfloor and allow for DIO_3 reclaim make the unit particularly convenient and cost effective within the fab environment. The PrimO$_3$ can deliver up to 60 L/min of DIO_3, depending on the desired ozone concentration. At 2 L/min flow, ozone concentrations as high as 115 ppm can be maintained while at 60 liters/min ozone concentrations of 20 ppm are achievable. Figure 113 shows the LIQUOZON® PrimO$_3$ unit and Figure 114 depicts the DIO_3 delivery characteristics of the system.

Figure 113. LIQUOZON® PrimO$_3$.

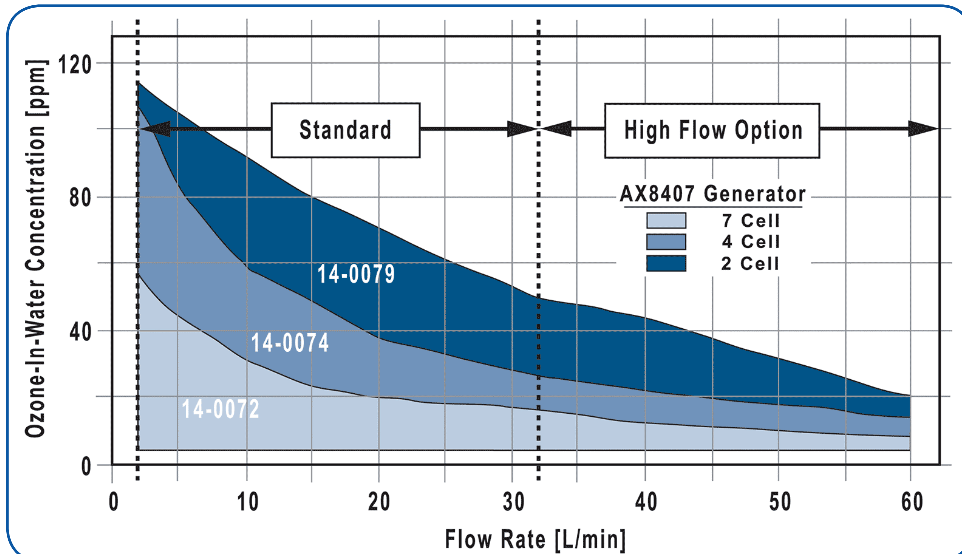


Figure 114. LIQUOZON® PrimO$_3$ DIO_3 delivery characteristics.

LIQUOZON® Single Ozonated Water Delivery System

The LIQUOZON Single DIO_3 delivery system is a highly compact unit specifically designed for low flow single chamber, single-wafer cleaning systems. The system delivers up to 95 ppm ozone DIO_3 at rates of 2 liters/min; at flow rates up to 20 liters/min ozone concentrations of 30 ppm are achievable (Figure 115). Overall, the unit delivers DIO_3 water with ozone concentrations between 5 and 95 ppm at flow rates between 0.5 and 20 L/min and at a pressure of up to 2.5 bar (gauge) (0.25 MPa). The small footprint of the Single Series makes it uniquely suited to single wafer applications, as well as many immersion applications.

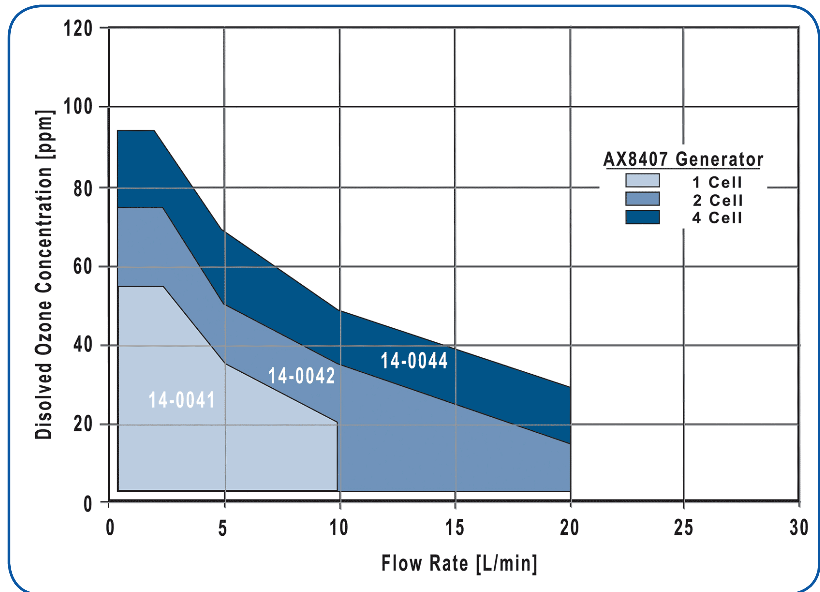


Figure 115. LIQUOZON® Single DIO_3 delivery characteristics.

LIQUOZON® Smart High Output Ozonated Water Delivery System

The LIQUOZON Smart system is available in two flow configurations. The basic LIQUOZON Smart system delivers up to 25 liters per minute while the high flow option can produce up to 50 liters per minute (Figure 116). The system is configured for multiple outlets and flow rates up to 80 liters per minute makes it suitable as a supply for multiple cleaning tools. Maximum dissolved ozone concentration is 80ppm. DIO_3 is generated in the LIQUOZON Smart system at a pressure of up to 2.5 bar (gauge) and flow rates of up to 50 L/min (80 L/min with multiple outlets). At 5 L/min flow, a dissolved ozone concentration of 114 ppm can be achieved while at 50 L/min flow, the ozone concentration has a maximum of 36 ppm.

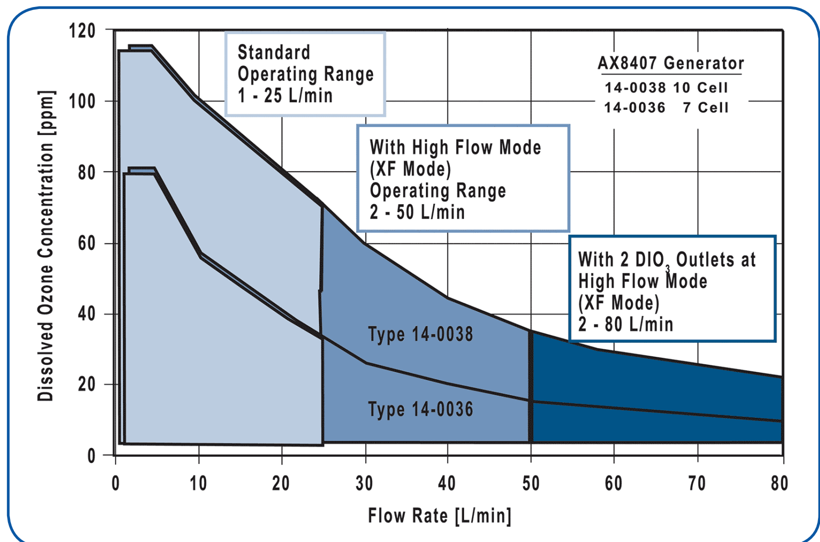


Figure 116. LIQUOZON® Smart delivery characteristics.

LIQUOZON® Stream High Output Ozonated Water Delivery System

The LIQUOZON Stream system is designed for use with multi-chamber single wafer cleaning tools having cleaning applications requiring up to 140 L/min DIO_3 flow and ozone concentrations of 25 - 115 ppm (Figure 117). The LIQUOZON Stream system has an integrated analyzer for dissolved ozone that can be used for accurate closed-loop control of the DIO_3 concentration. LIQUOZON Stream is specifically designed for the maintenance of stable DIO_3 concentrations, even with varying DIO_3 demands. It has an integrated booster pump for low pressure UPW supplies.

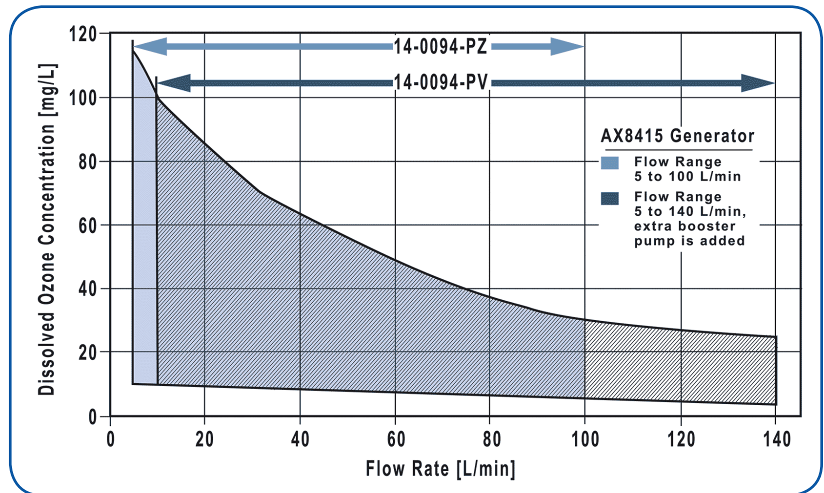


Figure 117. LIQUOZON® Stream DIO_3 delivery characteristics.

LIQUOZON® $HeliO_3$ Modular Ozonated Water Delivery System

LIQUOZON $HeliO_3$ (Figure 118) is an ozonated water delivery system specifically designed for solar applications, including cleaning, surface conditioning and oxide growth. The system is modular with at least two modules: a stand-alone gas module and a wet module that integrates with the customer tool (Figure 118 and Figure 119). It can be configured for closed-loop control and operation in either single pass or recirculation mode. It can also be configured with special safety features.

Figure 118. LIQUOZON® $HeliO_3$ Modular Ozonated Water Delivery System.

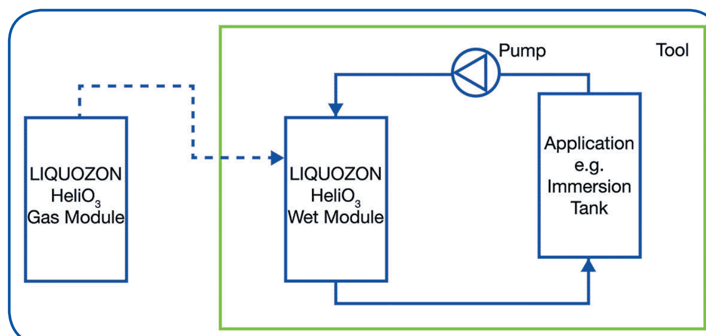


Figure 119. Typical application set-up for the LIQUOZON $HeliO_3$ system.

B. "Dry" Substrate Cleaning

"Dry" substrate cleaning refers to processes in which reactive, unstable ions and/or radicals are created in a plasma, directed to a substrate surface and there allowed to react with contaminants on the surface to produce volatile products that can be pumped away. In this way, contaminants such as organic material can be removed from the substrate surface.

The most common application for dry substrate cleaning is photoresist stripping and "descum" steps. Photoresist stripping involves the removal of photoresist residues by reaction with atomic oxygen radicals or some other highly reactive species to produce a volatile product that can be pumped away by the vacuum system. Photoresist descum is a similar but milder process carried out after photoresist patterning and development to remove residual photoresist left in a developed area. Such residues can negatively affect the uniformity in subsequent dry or wet etching steps. A descum step can also improve sidewall profiles.

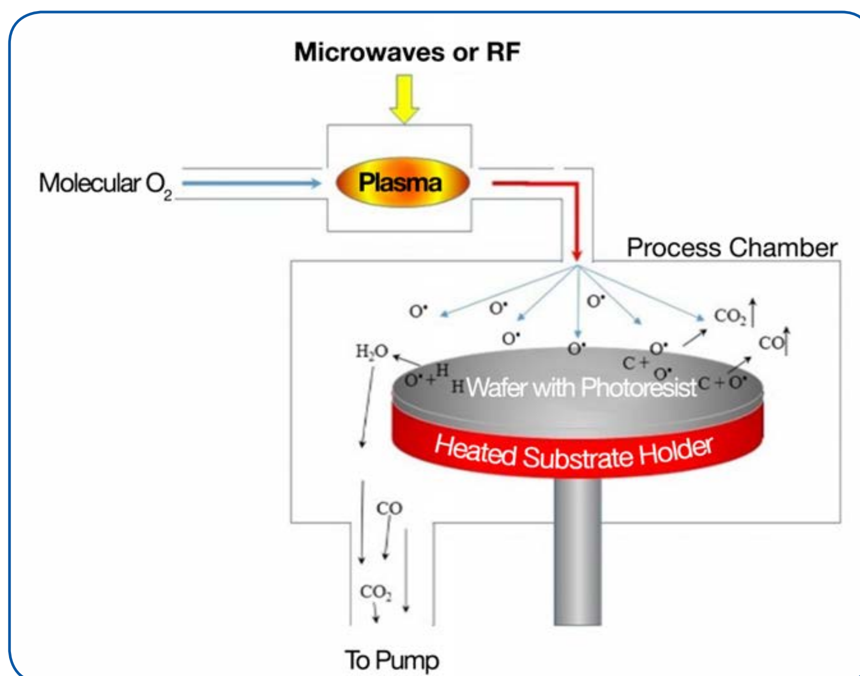
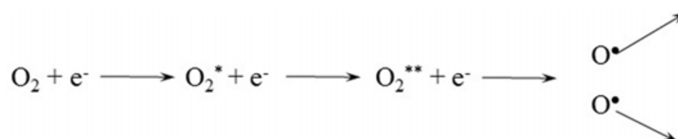


Figure 120. Components in a photoresist strip tool.

A representative chemical reaction sequence for the generation of oxygen radicals in a photoresist stripping process is shown in Figure 120 and in the chemical equation below. In this process, molecular oxygen gas is fed into a plasma chamber in which a microwave plasma is present. The plasma is generated under high vacuum conditions to ensure that the electrons, ions and other reactive species within it have a relatively long mean free path and thus long lifetimes. Highly energetic free electrons, present in the plasma, collide with the incoming, low energy oxygen molecules. If enough energy is transferred in a collision or if enough additive transfers occur, the internal energy of the oxygen molecule is raised to the point where the molecule is no longer stable and splits apart into two atomic oxygen radicals:



$$O_2 + e^- \longrightarrow O_2^* + e^- \longrightarrow O_2^{**} + e^- \longrightarrow O^\bullet + O^\bullet$$

Oxygen radicals are oxygen atoms with an unpaired electron. These extremely reactive species are long-lived under the material and vacuum conditions present in the system. They are pumped into the process chamber where they impinge on the heated surface of the substrate. There they react with organic hydrocarbon residues such as photoresists to generate carbon oxides and water. These volatile oxides are then pumped away by the vacuum system. Other chemistries may also be employed in cleaning processes, depending on the nature of the contaminants that need to be removed from the substrate surface.

The increase in etch and deposition steps, new materials, and new structures used in 2.5D and 3D packaging have made cleaning processes like photoresist strip and descum increasingly important for high device yield. The varying levels of cleanliness requirements and the different materials employed in the manufacturing process have made the availability of multiple cleaning options in a product line increasingly important to high yield.

Surface activation, an important process tied to cleaning, prepares the surface for the next process step ensuring good quality adhesion resulting in high quality die. This can also be accomplished using plasma techniques. Acceptable cleaning processes for semiconductor applications must achieve the following challenging objectives:

- Contamination free surfaces for highly sensitive devices
- Surface preparation and activation for better adhesion
- High throughput, ensuring minimal cycle time for higher overall productivity
- Ecologically friendly alternatives to concentrated sulfuric acid wet clean or others

MKS Product Applications in Dry Substrate Surface Cleaning

MKS offers RF and microwave plasma alternatives to wet substrate surface cleaning. These alternatives are compatible with multiple process gases, ensuring the best clean based on material chemistries. MKS's RF and Microwave Plasma products are an economical and environmentally sustainable alternative to wet cleans, avoiding the use of acids and solvents that require special storage and disposal.

The MKS R*evolution® III (Figure 121) is an integrated remote plasma source that provides extremely clean reactive gas species for surface cleaning applications. It integrates a quartz vacuum chamber, an RF power supply and all necessary controls into a compact, self-contained unit easily installed on a tool's process chamber. The extreme cleanliness of the reactive gas supplied by the R*evolution® III is ensured by its advanced design that employs a low-field toroidal RF plasma source. RF power is inductively coupled through ferrite cores into plasma that is confined within a toroidal quartz chamber. The source operates efficiently over an extremely wide range of gas flows and pressures, and generates reactive species in O_2, N_2, H_2, H_2/N_2, and H_2/He gasses. The use of quartz as the material of construction of the plasma chamber significantly reduces losses of atomic gases such as O, H, and N through wall recombination. Surface recombination rates of atomic gases on quartz are 100-1000 times lower than on most metals and dielectrics resulting in higher reactive gas output for the R*evolution III.

Figure 121.
*R*evolution® III*
Remote Plasma Source.

Toroidal RF sources are uniquely suited for high throughput photoresist stripping in advanced device fabrication. Their design produces high plasma densities and high concentrations of reactive species while the low electric fields result in minimal ion bombardment of the chamber walls and excellent gas stream purity. ICP mass spectrometry (ICP-MS) and x-ray photoelectron spectroscopy (XPS) analysis of wafers processed in a photoresist strip tool using a remote toroidal RF source under conditions producing strip rates 2-3 times greater than in the same system equipped with a microwave source show extremely low (~1 x 10^{10} cm^{-2}) particulate, quartz and metal contamination.

Applications for the R*evolution III include photoresist strip, wafer pre-clean, gate nitridation, and oxidation.

The MKS AX7610 Downstream Plasma Source (Figure 122) is a general duty microwave plasma source for use in remote plasma applications such as photoresist strip and passivation, surface modification, chamber cleaning and reactive gas generation. It can be configured with either a quartz tube for cleaning applications requiring atomic oxygen, hydrogen and nitrogen or a sapphire tube compatible with more corrosive reactive gas generation from species such as CF_4 and NF_3.

The MKS AX7610 is designed to be integrated into a microwave plasma subsystem comprised of power supplies, microwave magnetron heads, matching systems, etc. as shown in Figure 123. MKS offers a variety of microwave power generators and system accessories.

Figure 122. AX7610 Downstream Plasma Source.

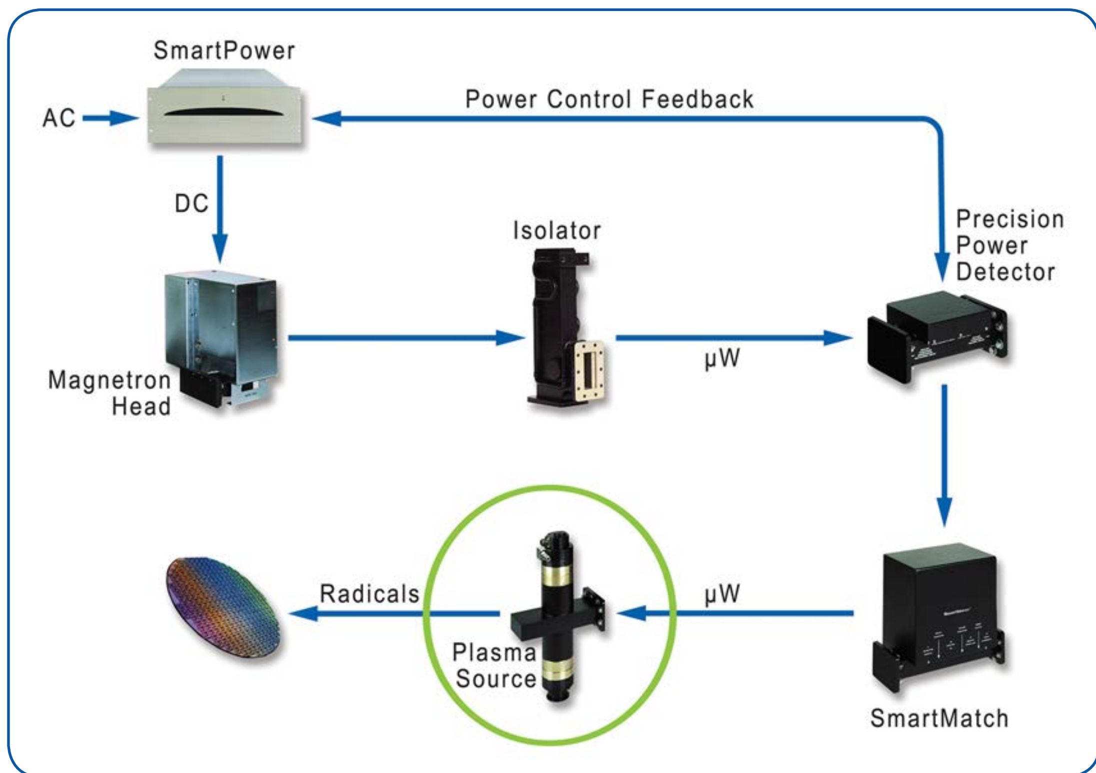


Figure 123. Microwave Plasma Subsystem.

III. Ion Implantation

The primary means of doping silicon for semiconductor device manufacturing have been ion implantation, gas source doping, and solid source doping. Ion implantation is by far the most important mode of introducing dopant atoms into silicon substrates and it will be the only method discussed here. Readers interested in older methods such as gas and solid source doping are referred to standard texts for further information. We will only provide a brief discussion of the basic aspects of ion implantation here. Those interested in developing an in-depth understanding of the issues associated with ion implantation are directed to texts on VLSI and Ultra Large Scale Integration (ULSI) Technologies [32] [96] and to available monographs on ion implantation [201] [202] [203].

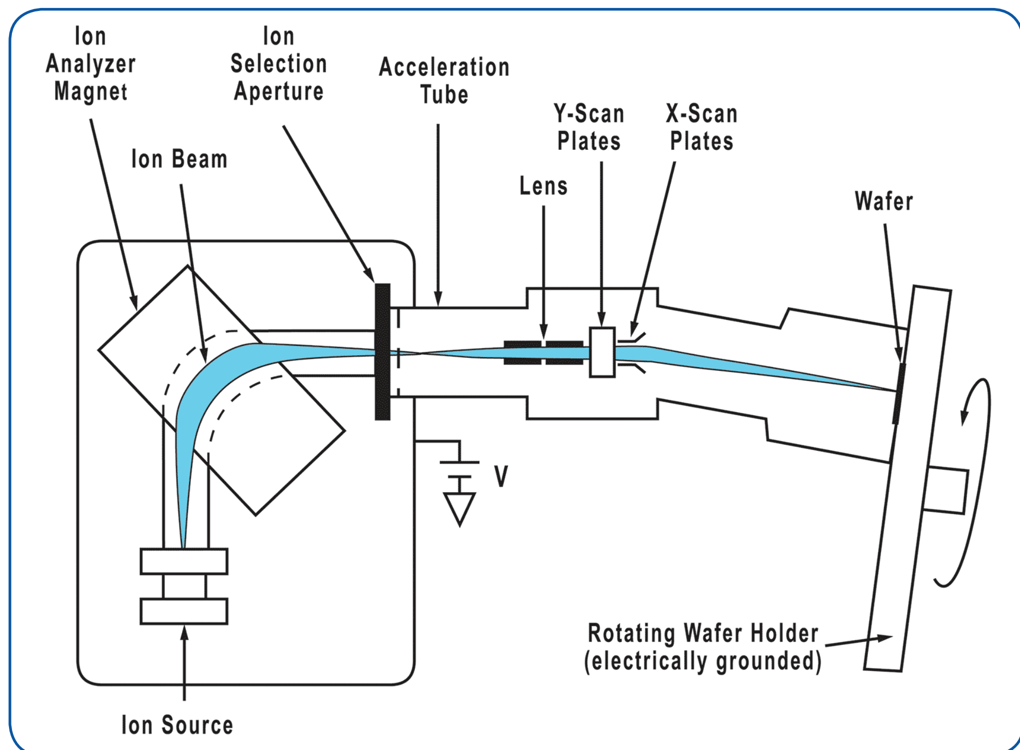


Figure 124. Components of an ion implanter [204].

Ion implantation owes its importance to the fact that it allows precise control over the depth of penetration of dopant atoms into the silicon. In the ion implantation process, dopant atoms are first ionized in an ion source. The source typically contains a plasma, generated by either RF or microwave radiation, in which the dopant atoms are ionized by electron impact. Figure 124 shows a schematic of an ion implanter (from [204]). Ion implanters must be maintained under high vacuum to permit linear free travel of the ions without occurrence of dispersion due to collisions with ambient gas molecules. Ions are extracted from the ion source using electromagnetic fields and the ion beam created by the extraction is directed into a mass analyzer where the beam is focused and bent through a right angle. The radius of the bend is determined by a combination of electromagnetic field characteristics and the mass to charge (m/e) ratio of the ions. This allows ions of a particular mass to be selected to exit the mass analyzer using a movable aperture (or an electromagnetic lens). In this way, only the desired dopant atom is selected from the different ions that may originate from the ion source. The beam of selected ions is then accelerated to high energies, ranging anywhere from sub-keV to MeV values and the high-energy ion beam is steered using electromagnetic fields so that it impinges on the semiconductor surface.

When a dopant ion strikes the substrate surface, it penetrates into the substrate crystal matrix to a depth that is proportional to its energy and angle of incidence. It doesn't penetrate the substrate in a linear manner but rather takes a "drunkard's walk" path through the crystal as its direction is scattered by collisions with substrate atoms (Figure 125). This results in a concentration profile for dopant atoms vs. penetration into the substrate that has a Gaussian distribution similar to that shown in Figure 126. Since dopant atoms do not necessarily substitute into lattice sites during the implantation process (they can come to rest in interstitial positions), post-implant annealing is necessary to "activate" the dopant by substituting the dopant atom into crystal lattice sites. The anneal will also act to repair any damage done to the silicon crystal matrix by collisions with the high-energy dopant ions and to broaden the dopant distribution profile somewhat. Once substituted into the lattice, the dopant will act as either a donor or acceptor depending on its electronic structure.

Ion implantation is normally used for n- and p-well formation in the substrate silicon and source/drain formation. The CMOS process flow discussed in Section A briefly discusses the use of ion implantation for this purpose.

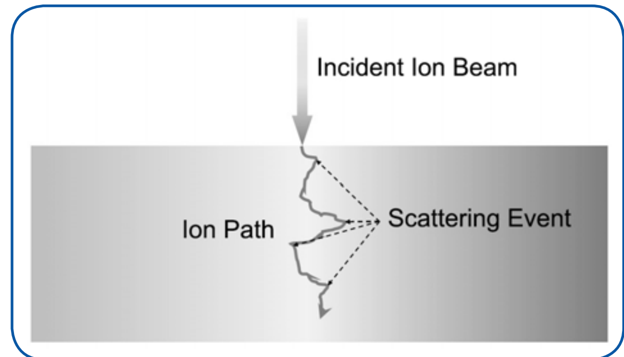


Figure 125. Representative ion penetration path into a silicon substrate.

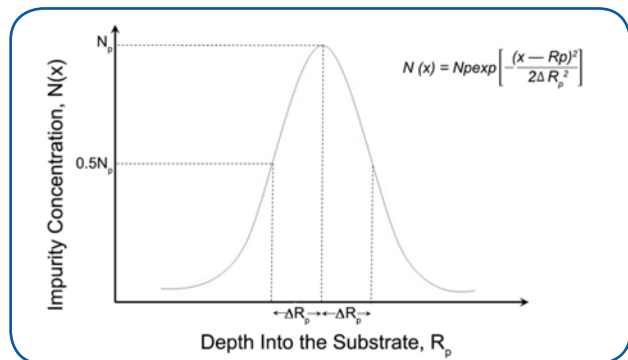


Figure 126. Representative dopant profile in a substrate that has undergone ion implantation.

MKS Product Applications in Ion Implantation

The success of ion implantation processes is heavily dependent on the quality of the vacuum within ion implanters. As a consequence, most of MKS's product lines associated with high vacuum applications in vacuum generation, control and monitoring are of relevance in the ion implant market. Detailed discussion of vacuum generation, control and monitoring is provided in Section B.I. These include:

- MKS Medium and UHV Vacuum Components and Fittings
- MKS Heater Jackets for Vacuum Components
- MKS Product Applications for Vacuum Pressure Measurement
- MKS Product Applications for Residual Gas Analysis and QMS Process Monitoring

MKS Products for Ion Source and Beamline Cleaning

MKS ASTRON® Paragon® (Figure 127) remote plasma generators can be used for F* generation in cleaning applications for ion sources and beamlines. Application notes on the use of reactive gas generators in this application are available on the AVS User Group website and from the IEEE [205].

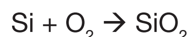
Figure 127. ASTRON® Paragon® Remote Plasma Source.

IV. Thermal Processing

The primary high temperature (thermal) processes employed in semiconductor processing are thermal oxidation, diffusion and annealing. These processes are normally performed at temperatures in excess of 900°C and at atmospheric pressure.

A. Oxidation

Thermal oxide (silicon dioxide) is a silicon dioxide film produced by the oxidation of substrate silicon, usually at temperatures in excess of 1000°C. Thermal oxides can be grown using a “dry” oxidation process:

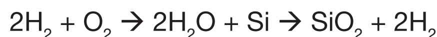


$$Si + O_2 \rightarrow SiO_2$$

Or a “wet” oxidation process:

$$Si + 2H_2O \rightarrow SiO_2 + 2H_2$$

Dry oxidations are typically performed at 900°C – 1200°C at high oxygen pressures. Dry oxidations exhibit the lowest oxide growth rate of the thermal oxidation processes used in semiconductor device manufacture, typically around 14 – 25 nm/hr. Because of this, dry oxidation processes are normally used only for processes that require silicon dioxide film thicknesses of less than 100 nm. Dry oxidation processes tend to produce silicon dioxide films with the highest quality electrical and material characteristics. “Wet” oxidations can be performed using either entrained water or in situ generated steam produced by the reaction of hydrogen and water:



$$2H_2 + O_2 \rightarrow 2H_2O + Si \rightarrow SiO_2 + 2H_2$$

Historically, wet oxidation was used in the LOCOS (LOCal Oxidation of Silicon) process for the production of electrically isolating “field oxide” in older device designs. The use of TEOS oxides and new flowable CVD films has largely supplanted this process in advanced device fabrication.

Silicon dioxide has been, arguably, the most important material of the different thin films employed for semiconductor device fabrication for more than a half century. Indeed, the fact that silicon forms a stable and adherent oxide with good electrical properties is probably the reason why. As a consequence, many studies are available in the literature that deal with the growth and characterization of thermal silicon dioxide thin films and the interested reader is referred to these works and references contained therein for an in-depth understanding of such aspects as the oxide growth mechanism and interface characteristics [105] [108] [109] [110].

Practically, thermal silicon oxides can be grown using a variety of equipment configurations and the best approach depends on the particular device requirements. In this section we will describe those equipment configurations that would most likely be encountered in the semiconductor fab.

Wafer Surface Preparation

Thermal oxides are used primarily for gate oxides in silicon semiconductor devices. In this application, it is critical that the silicon dioxide/silicon interface be as perfect as possible with a minimum of atomic scale defects such as “dangling” chemical bonds and non-stoichiometric hydroxyl contaminants, as discussed in Section A, Chapter III.A. To ensure this, substrates normally undergo a surface preparation step immediately prior to oxidation. This step typically involves the removal of any native oxide followed by hydrogen passivation of the surface using a dilute HF-last process. Alternatively, silicon surfaces can be passivated using a remote plasma-based process.

Batch Thermal Oxidation

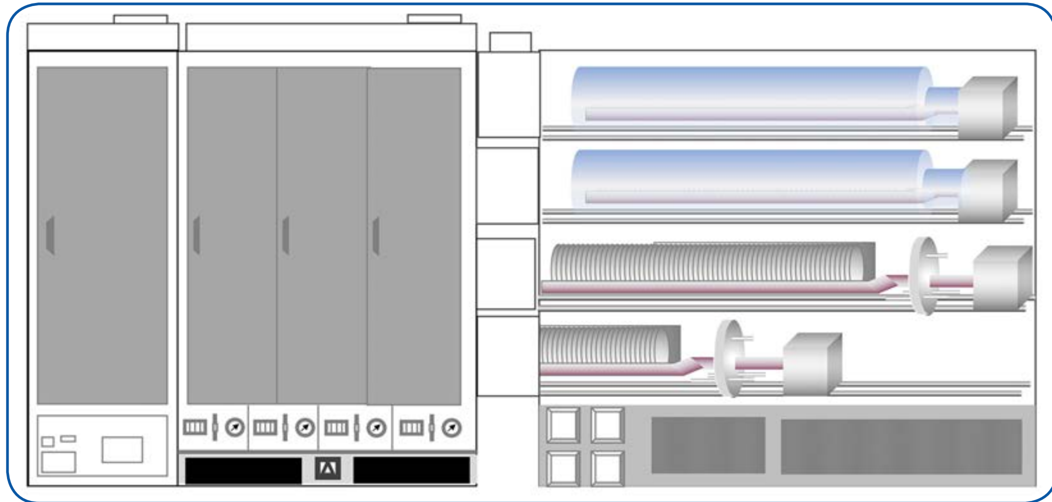


Figure 128. Horizontal hot-wall furnace load station.

Both dry and wet thermal oxide films can be produced using batch processing and resistively heated tube furnaces. These furnaces can be configured as either "horizontal hot-wall" or "vertical hot-wall" furnaces. Horizontal and vertical furnace systems are ubiquitous in older semiconductor fabs. Any reader who has visited an older fab will be familiar with these installations. Figure 128 shows an image of a conventional "load station" of a horizontal hot wall system. Unfortunately, it is extremely difficult to automate wafer handling in horizontal systems and ultra-low particulate processing demands such automation. Furthermore, vertical furnace processing has been found to yield better oxide film uniformities across the substrate surface. For these reasons, any batch processing that is found in more modern fab environments is usually conducted in furnaces configured for vertical processing. Such configurations can be readily adapted to robotic, cassette-to-cassette wafer handling and FOUP interfaces. For this reason, horizontal systems will not be considered further in our discussion.

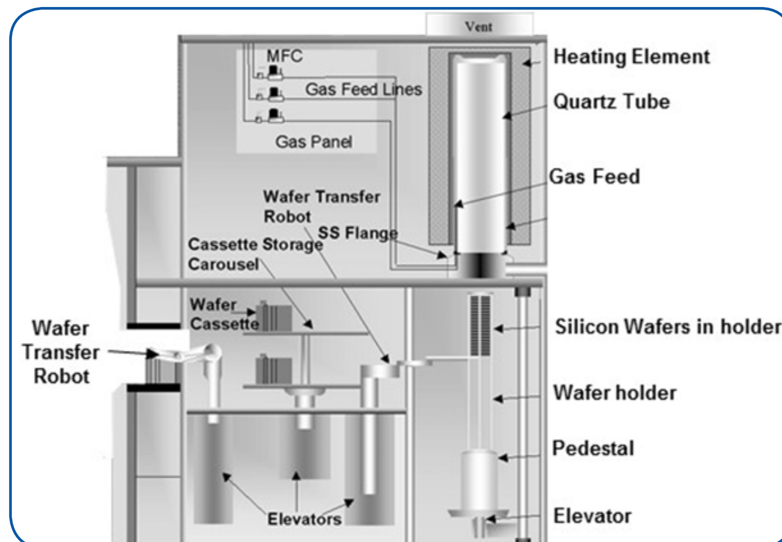


Figure 129. Vertical thermal batch furnace.

Figure 129 shows a schematic diagram of the internal components in a verticle furnace. Wafers of up to 300 mm diameter can be processed in vertical furnace environments. In a typical "dry" oxidation process, wafers are robotically transferred from a FOUP pod to a quartz wafer holder within the furnace. This wafer holder, commonly called a "boat" or "cassette," is used to hold the wafers during processing. These boats maximize the number of wafers processed for a given spacing between wafers since a minimum spacing between wafers is required to maintain acceptable uniformity of the oxide film across the wafer. A typical batch will contain 50 – 100 wafers with a number of baffle (non-process) wafers located at the top and bottom of the load. The quartz boat is mounted on a mechanical elevator. Once the boat is fully loaded, the elevator raises the load into the hot environment within the quartz process tube, sealing the tube against atmosphere. The tube is usually hot during the load step, but at a lower temperature than during the oxidation process. The ambient atmosphere within the wafer handling compartment and process tube is an inert gas in order to eliminate any potential for uncontrolled oxide growth as the wafers enter the hot zone of the furnace. After the load is in the hot zone and the tube sealed, the temperature is raised to process temperature, again under a blanket of inert gas that prevents any oxide formation. Process temperatures can range from 1000 to 1250°C. Temperatures and temperature profiles within the furnace must be precisely controlled, usually to within tolerances of ±0.5°C variance from setpoint or less. Once the whole system has equilibrated at the desired process temperature, the oxidant gas is introduced and the oxidation allowed to proceed for a fixed time, forming the oxide film on the silicon substrates. The oxidant gas can be (most often) high purity oxygen or a mixture of pre-combusted hydrogen and oxygen (steam + oxygen), depending on the device requirements. After the oxidation has proceeded for the prescribed time, oxidant gas flow is shut off, the tube is purged with inert gas and the temperature of the system lowered to the unload setpoint. The elevator then unloads the system by lowering the boat containing the wafers into the inert gas-filled wafer handling compartment to a position accessible by the wafer handling robotics. The load is allowed to cool to ambient temperature and the robotics transfer the processed wafers from the wafer boat to waiting FOUP pods.

Steam oxidation using a vertical batch furnace proceeds similar to dry oxidations but uses pyrogenic steam as the oxidizing gas. Pyrogenic steam is generated outside of the furnace in a quartz chamber where hydrogen and oxygen are combusted and the product steam is fed to the process tube.

Single Wafer Thermal Oxidation Processes

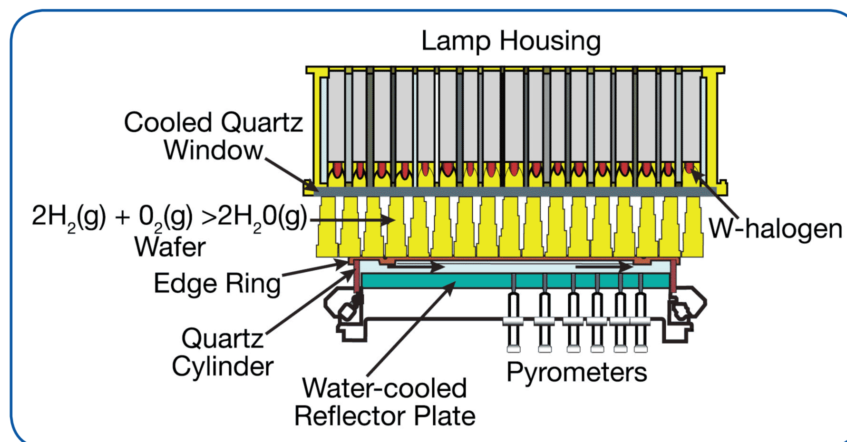


Figure 130. Rapid thermal oxidation chamber showing wafer in chamber and multi-zone heating lamp array [206].

Conventional, furnace-based thermal oxidation processes can be problematic for advanced device fabrication. In addition to particle control issues in the large batch process tools, many advanced device structures cannot tolerate the temperature burden experienced by the substrate during the load, equilibrate, process and unload process cycle. The reason is the temperatures and process durations required for thermal oxidation, dopant atoms migrate out of the device design parameters, changing concentration profiles and thus the electrical properties of the transistors being fabricated. There are only two ways to avoid this problem: either lower the temperature or reduce the time that the substrate spends at high temperatures. Since the time at temperature required for a given silicon oxidation process and film thickness is more or less fixed, process engineers have sought ways to reduce the ancillary time at temperature budget in the thermal oxidation process. This has been achieved through the use of rapid thermal oxidation (RTO) tools. These are single wafer tools in which the wafer is rapidly heated to oxidation temperature using an array of lamps in an optical system. When using RTO tools, temperature ramp up and ramp down times can be drastically reduced (Figure 130). Additionally, since the substrate spends only a relatively short time at oxidation temperatures, it is possible to perform the oxidation at somewhat higher temperatures reducing the thermal budget even further. Figure 131 compares the thermal budget for a conventional vertical hot-wall batch reactor with that for a rapid thermal oxidation tool.

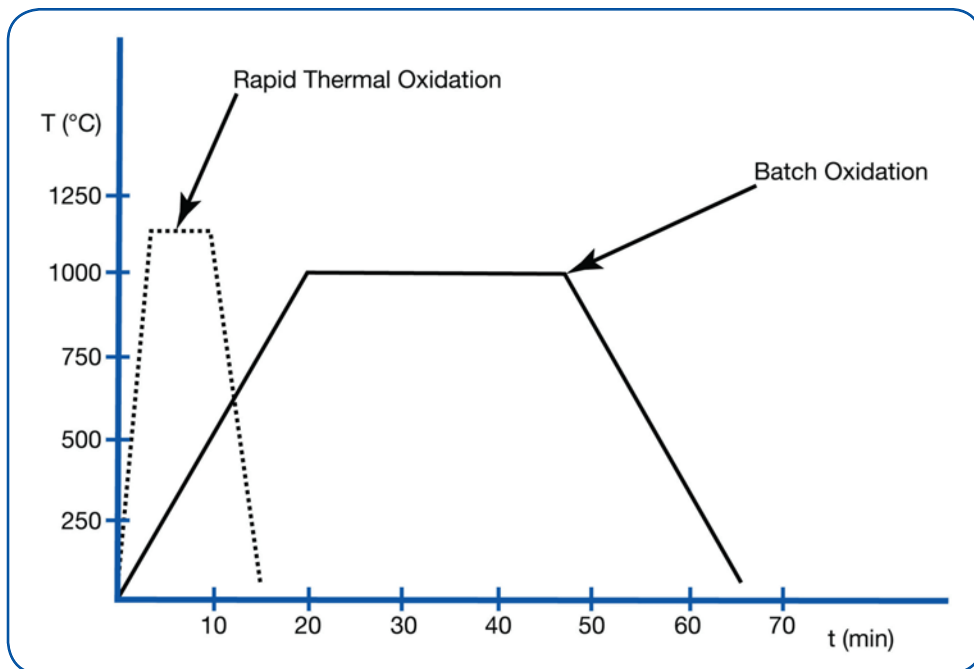


Figure 131. Thermal budget for RTO versus batch hot wall oxidation.

As with vertical batch systems, single wafer RTO systems are configured to allow cassette-to-cassette robotic wafer handling throughout the process. Typically, the RTO tool is incorporated into a multi-chamber cluster tool with 4 or more process stations. Different stations in the tool will perform different functions. For example, a cluster tool dedicated for gate stack formation might have individual stations for oxidation, nitridation, anneal, and cool down. Figure 132 shows a typical cluster tool configured for the production of gate stacks. Wafers are brought to the tool in a FOUP pod and interfaced with the robotic handler (bottom two stations in the cluster tool). The robot transfers wafers from the FOUP pod to the starting gate oxide station where the substrate undergoes rapid thermal oxidation. It then undergoes, sequentially, nitridation, anneal, and polysilicon deposition in dedicated cluster stations, forming the device gate stack on the substrate. After a cool down phase, the robotics transfer the finished wafer to the outgoing FOUP station.

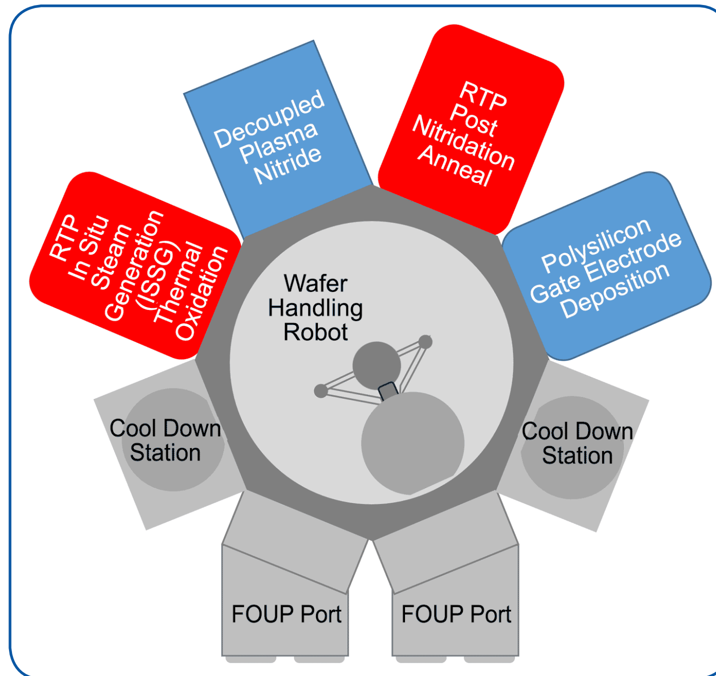


Figure 132. Integrated gate stack cluster tool.

B. Diffusion

Diffusion is the movement of impurity atoms in silicon (or another semiconductor) at high temperature. Basically, diffusion can be described as a "smoothing out" of any localized high concentration of impurity atoms that is driven by the motion of the atoms in the material and by their mobility through the dominant atomic matrix (Figure 133). The impurity concentration gradient, dC/dx, is the driving force for diffusion. Diffusion can be mathematically expressed using Fick's first and second laws of diffusion. The diffusion of dopant atoms in silicon is reasonably predictable using these laws along with information on the temperature and time at temperature for a substrate. We will not discuss the quantitative relationships for diffusion here. For more details, see the excellent discussion of this topic in Wolf and Tauber's text [32] and references therein.

Diffusion is normally carried out in a horizontal hot-wall tube furnace that is essentially identical to the batch oxidation furnaces described in the previous section.

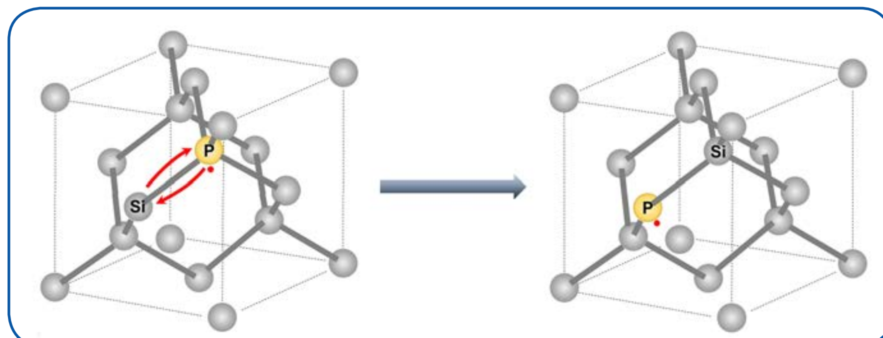


Figure 133. Dopant atom diffusion in the silicon crystal lattice.

mks

C. Annealing

Conventional, high-temperature annealing is used in device manufacturing to relieve stress in silicon; to activate ion-implanted dopants and to reduce structural defects and stress; and to reduce interfacial charge at the silicon-silicon dioxide interface. There are also lower temperature annealing processes that may not be referred to as such. Some other heat treatment processes are occasionally referred to as annealing. These include glass densification, polysilicon and doped polysilicon anneal, silicide formation and contact sintering. The normal temperature range for annealing processes lies between 900 and 1100°C (when processes such as polysilicon anneal and silicide formation are included, the lower limit is extended to about 700°C). It is normally performed in an inert ambient such as nitrogen or a reducing ambient such as forming gas (N_2/H_2). Wolf and Tauber [32] provide an excellent discussion on the importance of thermal annealing in device fabrication.

Historically, thermal annealing processes were most often performed using conventional resistively-heated tube furnaces. However, as device designs shrank late in the last century, it was found that the time at temperature profiles necessary for dopant activation and crystal damage repair in such approaches resulted in undesirable dopant diffusion. Since atom mobility within the silicon matrix is negligible at temperatures much below 900°C, temperature reduction was not an available option for solving this problem. As with the thermal oxidation processes discussed above, most of this undesirable dopant migration occurred during the heat-up and cool-down phases in these processes. A typical furnace anneal process for dopant activation and crystal damage repair might require 30 minutes at 900°C. Under such conditions, dopant atoms, especially the small B atom, suffer excessive diffusion. Studies showed that dopant activation and crystal repair in the same substrate could be achieved with a 1050°C for 20 seconds (or less) without excessive dopant diffusion which led to the obvious problem resolution of rapid thermal annealing (RTA). As with rapid thermal oxidation, in RTA the silicon wafer is heated to high temperature in seconds by a bank of heat lamps. Cooling is also rapid since the thermal mass of the entire system is small.

D. Rapid Thermal Processing

Thermal budgets have become so important in modern processing that single wafer rapid thermal processing approaches are now ubiquitous in the industry. These processes are commonly lumped together under the label Rapid Thermal Processing or RTP and they include annealing, oxidation, and thin film deposition processes.

MKS Product Applications in Thermal Processing

While MKS does not directly produce thermal processing equipment for the semiconductor industry, it acts as an OEM supplier to other organizations that produce this equipment. Here we list some of the key equipment components supplied by MKS to these producers.

Gas Generation and Flow/Delivery Products

All thermal processes require precise control of the gas ambient. Thermal oxidations require an inert ambient during heat up and cool down coupled with sharply delineated periods when oxidant gas is present at the process temperature. Diffusion and annealing processes must be performed under either an inert gas ambient or a reducing ambient such as forming gas. MKS gas generation and flow control products are commonly used across the semiconductor industry for this purpose. The discussion below includes some, but not all, MKS Gas generation and Flow/Delivery Products, focusing primarily on those products that are expected to find use in thermal processing applications.

Native Oxide Removal and Surface Preparation

Modern device structures demand precise, atomically clean and defect-free interfaces between gate insulators and substrates. This can only be achieved when ill-defined native oxides are not present during

the initial growth phase of thermal oxides. To achieve this, it is necessary to etch the native oxide away and to passivate the surface with hydrogen. A number of MKS products are relevant to this surface preparation process.

The Paragon® H* (Figure 134) is a remote RF plasma source specifically created and optimized to produce a high output of hydrogen radicals. Atomic hydrogen etches native oxide from the substrate surface and passivates the surface. The use of remote plasma for hydrogen atom generation eliminates any direct exposure of the substrate to plasma, ensuring low defect surface passivation. It is designed to deliver up to 5 L/min hydrogen at pressures of up to 3 Torr, delivery parameters that are well suited to plasma-based silicon surface passivation processes. MKS supplies an Application Note on Toroidal Plasma Source Operation with Hydrogen that is focused on the use of remote toroidal plasma sources for hydrogen generation.

Figure 134.
ASTRON® Paragon®
Remote Plasma Source.

Microwave Plasma Subsystems are also suitable for atomic hydrogen production in surface passivation applications.

Plasma-based surface passivation technologies are all performed under medium to high vacuum conditions. MKS Vacuum Measurement and Control, detailed in Section B, Chapter I, all find application in these technologies.

Thermal and Pressure-based Mass Flow Controllers and Meters

Mass Flow Controllers (MFCs) and Meters (MFMs) from MKS are designed to meet a broad variety of processing applications. Most MKS mass flow control and metering products offer both analog (0-5 VDC; 4-20 mA) and digital (Devicenet, Profibus®, EtherCAT®, RS485) I/O with embedded Modbus and an Ethernet user interface.

G-Series MFCs and MFMs are broadly applicable instruments that provide cost-effective, high performance measurement and control of gas flows between 5 cc/min (sccm) and 250 L/min (slm) with an accuracy of 1% of setpoint. They have multi-range, multi-gas capability and are available in either elastomer or metal-sealed configurations (Figure 135).

ALTA MFCs are especially suitable for mass flow control of high flows of hydrogen in pyrogenic oxidation and annealing applications. Hydrogen normally presents a notoriously difficult flow control problem.

I-Series MFCs are metal-sealed multi-gas/multi-range controllers that are also suitable for hydrogen service in pyrogenic oxidation and annealing applications. The I-Series MFCs are IP66-rated for use in harsh industrial environments.

Figure 135.
G-Series Thermal
Mass Flow Controllers.

This page intentionally left blank.

V. Thin Film Deposition

The thin films that are used to fabricate microelectronic devices are all formed using some kind of deposition technology where the term refers to the formation of a deposit on a substrate. In semiconductor device manufacture the following deposition technologies (along with their commonly used acronyms) are:

- Low Pressure Chemical Vapor Deposition – LPCVD
- Plasma Enhanced Chemical Vapor Deposition – PECVD
- Sub-Atmospheric Pressure Chemical Vapor Deposition – SACVD
- Atmospheric Pressure Chemical Vapor Deposition – APCVD
- Atomic Layer Deposition – ALD
- Physical Vapor Deposition – PVD
- Ultra-High Vacuum Chemical Vapor Deposition – UHV-CVD
- Diamond-Like Carbon – DLC
- Commercial Film – C-F
- Epitaxial Deposition – Epi

Since epitaxy was discussed in some detail in Part A, the process will not be discussed further in this section. APCVD processes other than epitaxy can be understood within the context of the discussion on CVD and thin film formation below; these processes will not be described in detail. Rather, we will briefly describe thin film deposition principles and representative CVD processes followed by a brief description of PVD processing.

Typical films produced by CVD processes include:

- Epitaxial Silicon
- Epitaxial Compound Semiconductors
- Polycrystalline Silicon
- Silicon Dioxide (including P- and B-doped Oxides)
- Aluminum Oxide
- Low-k Dielectrics (SiO-F, DLC, amorphous C-F)
- Silicon Nitride
- Titanium Nitride
- Tantalum Nitride
- Transition Metal Oxides (i.e. TiO_2, ZrO_2, Hf_2O_5, Ta_2O_5);
- Metals (Al, W, Cu)
- Silicides (WSi_x, $CoSi_x$)

A. Chemical Vapor Deposition and Thin Film Formation

Chemical vapor deposition processes can be defined as: any process in which a thin solid film is formed on a substrate by the surface-mediated reaction of adsorbed precursors from the gas phase. The reactive nature of CVD processes distinguishes them from physical processes such as evaporation and sputtering that are employed in PVD. The term "surface-mediated" refers to the fact that the solid film is formed by a heterogeneous reaction occurring at the substrate surface. The chemical compounds that react on the surface can be either the original reagent chemicals fed to the system or short-lived intermediate species created in the high temperature gas phase. As a general rule, most CVD processes and equipment are designed so as to avoid the formation of such gas phase intermediate species, since they can self-react to produce gas phase particulates. Historically, most CVD processes were carried out using batch reactor approaches in either horizontal, hot-wall tube reactors under vacuum or in continuous reactors at atmospheric pressure. As device sizes shrank, however, the need for tighter control over the chemical processes in CVD reactions became ever more critical, leading to a greater prevalence today of single wafer/cluster tool equipment configurations in which thin films can be grown.

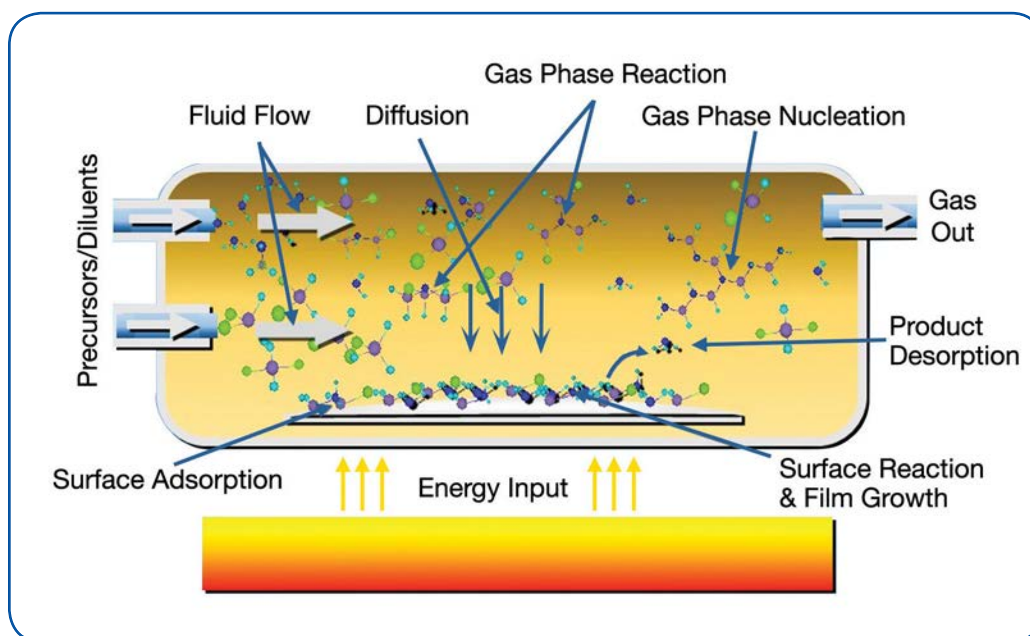


Figure 136. Aspects of a CVD process.

Figure 136 shows a schematic that can help in understanding the different processes in a CVD reactor. Chemical vapor deposition processes can be broken down into a number of discrete steps: First the precursor chemicals must be fed into the CVD reactor. Once in the reactor, precursor molecules must be transported to the substrate surface, usually through a combination of fluid transport and diffusion. Once on the surface, the precursor molecule must remain there long enough to react. After reaction occurs, the product thin film atom must remain on the surface while the by-product molecules must desorb from the substrate surface to make room for more incoming precursor molecules.

For now, we will not consider the pressure of the reaction nor the kind of reactor configuration, since these can vary widely and are often process-specific. The main common characteristics among different types of CVD processes is that, in all cases, the reactions or critical aspects of the reactions are thermally driven, and this requires that energy be added to the process. This energy can be in the form of heat to the substrate, plasma energy to the reactants or a combination of both.

The precursors used for CVD reactions must be of extremely high purity since any impurity will end up incorporated into the deposited film. Such impurities introduce uncontrolled changes in the material properties of the films that are detrimental to device performance. As a corollary to the statement on purity, CVD precursors must also be stable under storage conditions since any decomposition produces impurities that will be fed to the process. Ideally, CVD precursors only react under the temperature and pressure conditions that exist within the CVD process.

Precursors/Diluents

Typical precursors for CVD processes include:

- Oxygen
- Halides: H_2SiCl_2, $HSiCl_3$, $TiCl_4$, WF_6, etc.
- Hydrides: SiH_4, GeH_4, $AlH_3(NMe_3)_2$, NH_3, etc.
- Organometallics: $AlMe_3$, $Ti(CH_2tBu)_4$, etc.
- Metal Alkoxides: TEOS, Tetrakis Dimethylamino Titanium (TDMAT), $Ti(OiPr)_4$, etc.
- Metal Dialkylamides: $Ti(NMe_2)_4$, etc.
- Metal Diketonates: $Cu(acac)_2$
- Metal Carbonyls: $Ni(CO)_4$

The first step in the actual CVD process is the controlled introduction of the precursors and any required diluent gases into the reaction chamber; this introduction is depicted on the left side of the chamber in Figure 136. To do this, the precursors must have sufficient vapor pressure to produce a stable, controllable flow to the process chamber. High pressure gases such as silane, hydrogen, ammonia, etc. meet this criterion and are easily delivered to the process using MFCs. Many liquids (e.g. halides such as dichloro- and trichlorosilane, titanium tetrachloride, tungsten hexafluoride, tantalum pentachloride; organometallic compounds such as TEOS, trimethylphosphate (TMP), aluminum alkyls, tetrakis(diethylamido) hafnium, tetrakis(dimethylamido)titanium, etc.) also have sufficient vapor pressure to produce stable, controllable flows using MFCs. Potential precursors having lower volatility, including volatile solids, present a problem for delivery to CVD reactors since most methods employed to date suffer from repeatability problems. In practice, a substrate is loaded into the process chamber and heated to the required process temperature under inert gas flow. Once the substrate is at temperature, the precursor/diluent gas mixture is introduced to the process chamber.

Fluid Flow/Diffusion

Once in the chamber, the precursor molecule is transported to the substrate surface by fluid flow effects, diffusion or a combination of both phenomena. These factors control the rate of arrival of precursor molecules at the substrate surface; this rate combined with the residence time/reactivity of the precursor on the surface, determines the rate of growth of the CVD film. In very low pressure CVD reactions (i.e., UHV-CVD), the precursor mean free path is of the same dimensional order as the size of the CVD process chamber and precursor molecules flow to the substrate surface without the interference of hydrodynamic boundary layer effects or gas phase reactions. Under these conditions, the rate of film growth is dependent only on the chemical reaction rate on the substrate surface provided that sufficient precursor is fed to the reactor inlet. Under intermediate pressures such as those in conventional LPCVD and PECVD processes (0.1 – 10 Torr), gas phase collisions and boundary layer effects more strongly influence precursor delivery to the substrate surface. LPCVD processes, while they are still primarily rate-limited by surface reaction rates, can be affected by mass transfer rates (diffusion rate) of precursor through the boundary layer to the substrate surface. LPCVD process conditions are designed with the assumption that the rate of film growth

will not be limited by the rate of diffusion of precursor through the boundary layer rather than by the surface reaction rate, however, many LPCVD processes end up operating in a mixed regime where boundary layer effects play a role. The allowance for boundary layer effects plays a significant role in the physical configuration of most LPCVD reactors. A similar situation exists in PECVD reactors where mass transfer issues can influence the process even more strongly since the highly reactive chemical species produced in plasma result in much higher surface reaction rates. Additionally, LPCVD and PECVD reactors must be designed with care for surface-to-volume ratios and gas phase recirculation cells to avoid gas phase nucleation and particle contamination. APCVD processes operate at pressures of 760 Torr and, under these conditions, gas phase collisions and surface boundary layer effects dominate. Fluid flow patterns must be carefully managed within APCVD reactors to ensure that equal concentrations of reactants reach all parts of the substrate surface. As with LPCVD/PECVD processes, surface-to-volume ratios and gas recirculation zones are present and, in fact, are an even more critically important consideration for APCVD reactor designs.

Gas Phase Reaction and Nucleation

Depending on the chemical nature of the precursor as well as the temperature and pressure conditions within a CVD process, gas phase reactions may occur. All CVD reactors and processes are specifically designed to minimize or eliminate such reactions since they can lead to the nucleation of particulates in the gas phase and this can drastically reduce device yield. The reader should be aware, however, that some processes such as TEOS/ozone SACVD actually depend on the gas phase formation of reactive intermediate compounds in order to achieve beneficial process and product characteristics. This issue is highly process-specific and it is beyond the intended scope of this discussion.

Surface Adsorption and Reaction

While it may not be the controlling factor for growth rate in all forms of CVD, the adsorption and reaction of the precursor(s) on the substrate surface are obvious pre-requisites for thin film growth and it is worthwhile to take a moment to understand the processes that occur on the surface during thin film deposition. The factors that control the surface reaction are:

- Precursor flux to the surface
  - Controlled by fluid flow and/or diffusion (see above)
- Availability of surface sites that will adsorb precursor molecules and retain them long enough for reaction to occur
  - Controlled by product desorption rate and affected by adsorbed impurities
- Substrate surface temperature
  - Affects both residence time and reactivity of the precursor molecule

Once the precursor reaches the surface, it must adsorb and remain there long enough to react either by decomposition or with a co-reactant similarly adsorbed on the surface. The concentration of adsorbed precursors on a substrate surface is governed by the rate of arrival of precursor molecules at the surface; the proportion of incident molecules that stay on the surface; and the density of sites available for the adsorption of precursor molecules. Precursor arrival rates are also called the precursor flux and the probability of an incident molecule actually sticking to the surface once it arrives there is known as the "sticking coefficient". The rate of adsorption of precursor molecules is directly proportional to both factors and to the density of free adsorption sites on the surface. If reaction by-products remained adsorbed, they would block fresh precursor from adsorbing and thus slow the rate. Similarly, if impurities are present they may occupy surface sites limiting the adsorption of fresh precursor and slowing film growth rates.

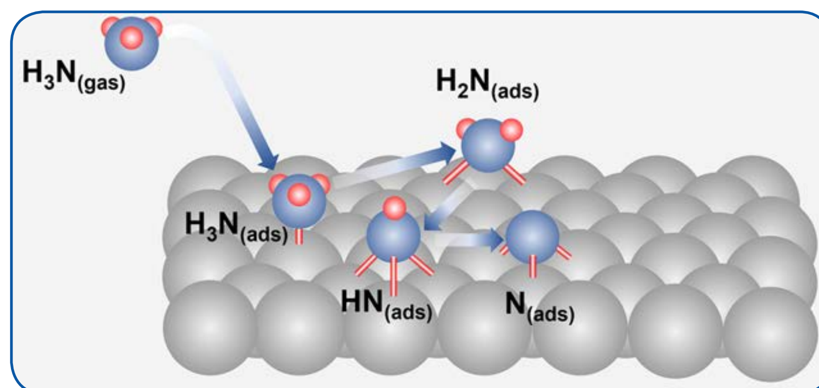


Figure 137. Decomposition of ammonia on a FCC metal surface [207].

Molecular precursors, once adsorbed on the substrate surface, undergo decomposition and/or reactions that are promoted by their interaction with the surface. A typical surface decomposition is represented by the adsorption and decomposition of ammonia, NH_3, on a metal surface, shown in Figure 137. In this case, the ammonia molecule undergoes a step-wise decomposition to produce a nitrogen atom on the surface and hydrogen molecules which desorb into the gas phase. The hydrogen desorption rate is dependent on the temperature and the process pressure. As noted above, this rate impacts the overall rate of the deposition process since the sites occupied by hydrogen or other by-product atoms cannot adsorb incoming precursor molecules. Silane and other molecular precursors used in semiconductor CVD adsorb and decompose on silicon and other surfaces in an analogous manner.

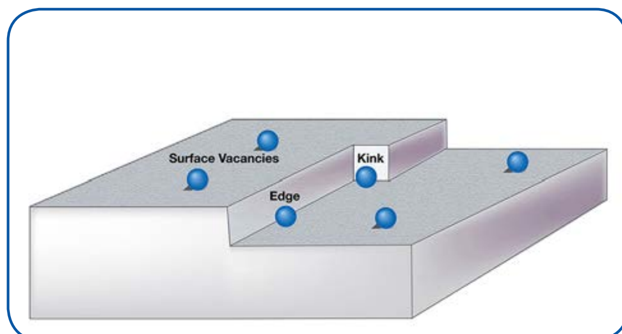


Figure 138. Vacancy, edge and kink sites on a substrate surface.

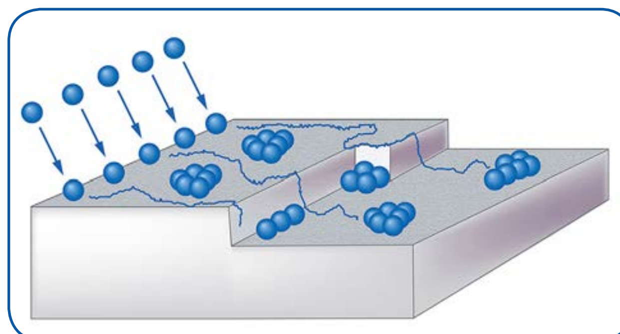


Figure 139. Adatom surface migration and nucleation.

The elemental adatoms (atoms that are adsorbed on a surface) that are created in the surface reactions are not stationary; in the case of a single crystal substrate such as a silicon wafer, they migrate across the surface until they encounter a high-energy surface site. High energy sites on a single crystal surface are sites such as atomic vacancies in a crystal face, lattice edges where a crystal plane ends and sites where a lattice edge experiences a “kink” (Figure 138). On non-crystalline surfaces such as silicon dioxide these well-defined high energy sites are not present, however, other types of surface sites act to trap adatoms. Once it encounters such a site, the adatom tends to remain stationary and to act as a starting point for surface nucleation of the desired thin film (Figure 139). Nuclei grow on the surface by the addition of migrating adatoms until they finally coalesce into a coherent film layer and the process begins again on this layer (obviously, this is an idealized view and actual growth will be “messier” with growth occuring over multiple layers and levels on a real surface). Figure 140 shows a highly-idealized picture of the different stages of film growth that follow the adsorption of atomic precursors on a substrate surface. These stages occur in all CVD and PVD thin film deposition processes.

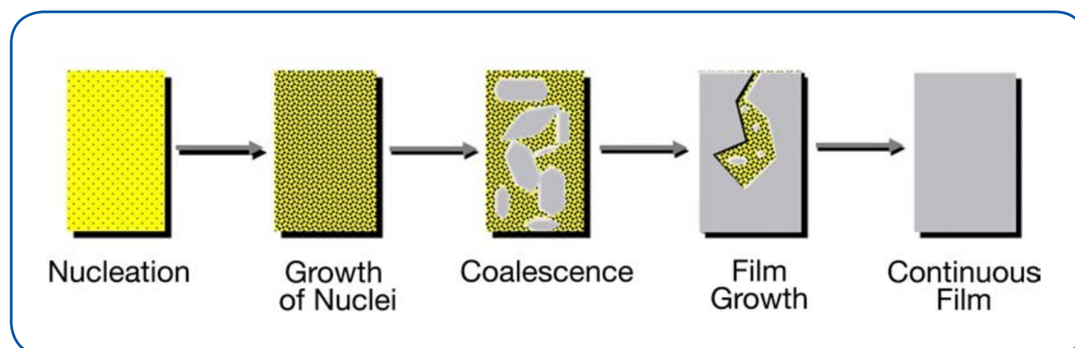


Figure 140. The stages of thin film growth.

Energy

The surface reactions that produce the thin films in these deposition processes require energy to promote the chemical reaction and to enable the surface migration of adatoms which produces nucleation and continuous film formation. In most deposition processes, thermal energy fulfills either all or a portion of the energy requirements of the process. Usually, this thermal energy is supplied using conventional resistive heaters, inductive heating or infrared lamps; the choice usually depends on the device requirements/ limitations and the equipment configuration in use. Energy can also be supplied to the system in the form of high energy chemical species produced using a plasma. The plasma source is usually either a radio frequency (RF) or microwave power supply. Historically, plasmas that supplied reactive chemicals for the deposition have been created adjacent to substrate surfaces; however, many modern processes employ remote plasma sources from which the reactive species can be extracted and fed to the thin film process without incurring direct exposure of the substrate to the plasma. The reason for this lies in the fact that the free ions and electrons in a plasma can bombard the growing film, producing defects and adding impurities that are detrimental to its material properties.

Transport of Product Gases

In order for the chemical reactions that produce the thin films on the substrate surface to proceed at an acceptable rate, it is necessary to remove the product gasses from the deposition environment. This is usually accomplished using mechanical vacuum pumps or blowers. Obviously, pumping speeds, gas handling capacity and materials compatibilities of the pumping system must be matched to the inlet flows, reaction characteristics and physical/chemical properties of the product gases.

1. Low Pressure Chemical Vapor Deposition (LPCVD) Systems

LPCVD deposition systems typically operate at pressures that range from 0.1 to 10 Torr. The reader will recall that this is considered a medium vacuum application. Reactor configurations that have been used for LPCVD thin film processes include resistance heated tubular hot-wall reactors, vertical flow batch reactors and single-wafer reactors. Throughout the latter part of the 20th Century, much LPCVD processing was performed in horizontal hot-wall tube reactors, a schematic of which is shown in Figure 141.

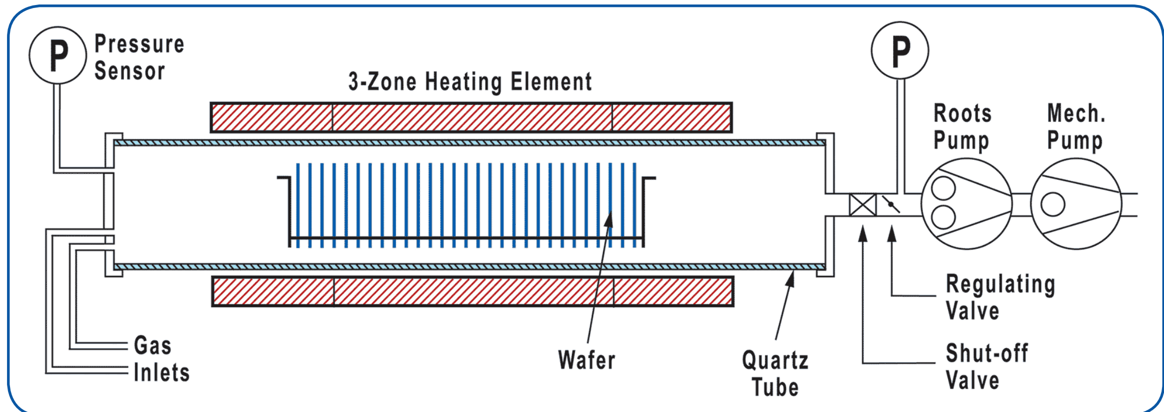


Figure 141. A horizontal hot-wall LPCVD reactor [32].

These reactors could process 100 or more wafers simultaneously and were very successful in depositing LPCVD silicon dioxide, silicon nitride and polysilicon thin films. LPCVD reactions for these films require temperatures that range from 425°C (silicon dioxide, low temperature oxide, LTO) to 740°C (silicon nitride) with occasional processes running at greater than 800°C (silicon dioxide, high temperature oxide, HTO). These process temperatures are achieved using resistive electrical heaters. The precursor gas in this equipment configuration is fed into the reaction chamber using either an inlet at the front of the tube or long injector tubes which lie on the bottom of the tube and run the length of the boat of wafers. However, both of these gas inlet configurations present problems for the user. If the precursor is simply fed in at the front of the tube, it will react on the wafer and equipment surfaces as it passes down the tube, depleting the precursor concentration in the gas phase. This results in higher precursor concentrations and therefore thicker deposited films on wafers located at the front of the tube. As a consequence, the film thickness profile for a LPCVD process in which the precursor is input at the door and a constant temperature is maintained over the wafer load looks like that shown in Figure 142. This depletion problem could be overcome through the use of temperature ramps over the wafer load (increasing temperatures towards the pump end of the load) or through the use of gas injectors in lower temperature processes. However, as device designs grew smaller in the '90s and early 2000s these solutions failed as they led to unacceptable variations in film properties (between films deposited at the cooler temperatures at the door end of the load and those deposited at the higher temperature pump end) or to reproducibility and costly maintenance issues in the case of injector use. Additionally, the relatively high time-at-temperature burden of batch hot-wall processes could no longer be tolerated when processing devices at smaller design rules. Finally, certain hot-wall LPCVD processes required "caged boat" wafer carriers (doped TEOS oxide, in situ doped polysilicon) that produced particle levels that were unacceptable for advanced device processing. For these and other reasons, horizontal hot wall systems have fallen into disuse in most fab environments.

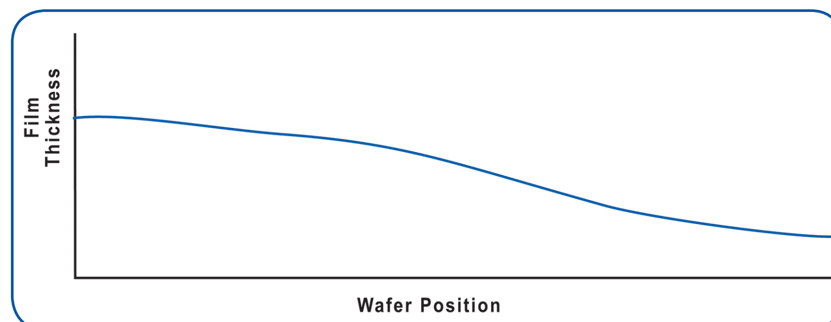


Figure 142. Film thickness profile over a wafer load in a constant temperature LPCVD process without precursor injectors.

The failings of conventional LPCVD equipment in advanced device processing led to an exploration of several alternative equipment configurations for LPCVD processes. Cold-wall, inductively coupled systems and cold-wall, vertical flow small batch systems experienced some acceptance and can still be found in use today. However, modern fabs have largely migrated to the use of single wafer cluster tools for CVD and other processing needs owing to demonstrated advantages in wafer handling, effective particle and process control and process integration.

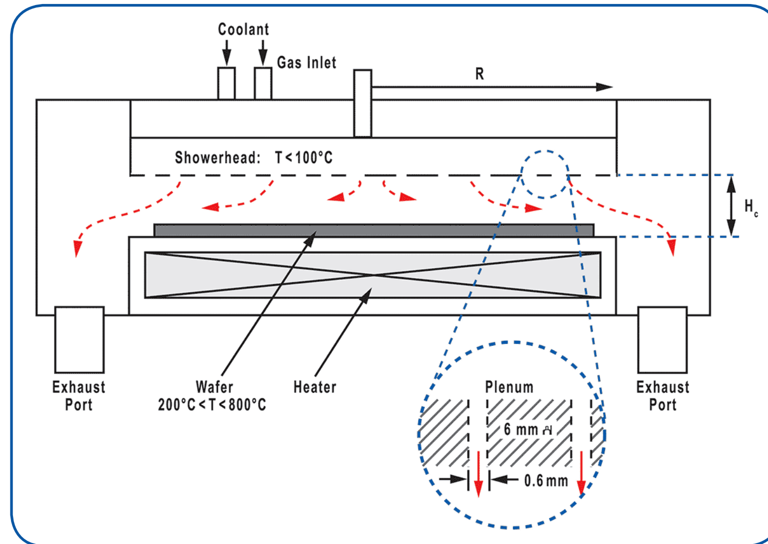


Figure 143. A typical single wafer CVD process chamber configuration.

Single wafer cluster tools process a single wafer at a time in small process chambers that can be designed for optimal performance in a variety of areas. Gas recirculation and surface-to-volume characteristics within single wafer chambers can be much more precisely tuned than in other equipment configurations, eliminating any potential for gas phase nucleation and in-process particle generation. Fluid flow and diffusion is much more effectively controlled in these configurations which results in more efficient use of precursors and higher deposition rates. The latter characteristics have allowed device manufacturers to significantly reduce the time-at-temperature burden in CVD processes, a key requirement in the manufacture of devices with nanoscale geometries. These and other more process-specific advantages have resulted in the wide-spread use of such cluster tools within the industry.

The internal configuration of a single wafer process chamber within the cluster tool depends on the intended process to be performed within the chamber. Figure 143 shows a schematic of a representative single wafer CVD chamber. Precursor is fed to the chamber using a showerhead arrangement that ensures a uniform concentration of precursor over the entire wafer face. The showerhead and chamber walls are cooled to temperatures below which deposition should not occur in order to minimize any potential for particle production by the spalling of deposits on these surfaces. The substrate is heated using resistive or optical heating of the wafer chuck. The gas exits the chamber via exhaust ports below the wafer. The geometries within the chamber are precisely designed to optimize process performance for zero particle generation, even film uniformity and desirable film properties.

Figure 144 shows a schematic for a cluster tool configuration. Wafers are introduced into a transfer station using a load-lock arrangement and FOUP protocols. Once the transfer station has been pumped to vacuum conditions, a robotic wafer handler (in the central chamber) picks up the wafer and transfers it to a process chamber. The process chamber is sealed and the wafer undergoes the intended process, in this discussion thin film deposition by CVD. When the process is complete, the wafer is extracted from the process module and transferred either to another process module (i.e., gate stack processing, annealing,

etc) or to a cooling module then to the transfer station that is brought back to atmospheric pressure so that the processed wafer can be removed from the system.

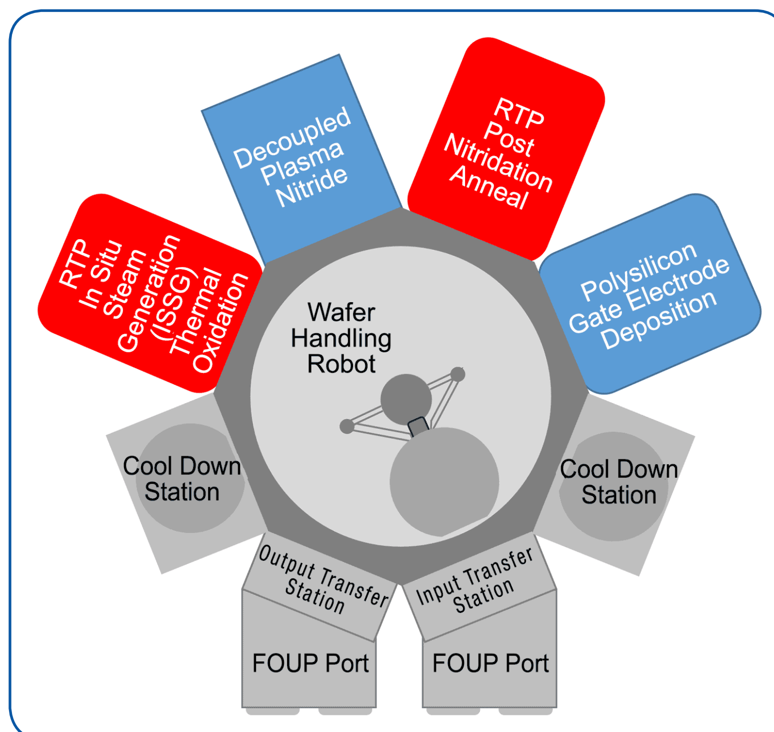


Figure 144. Schematic of a modern cluster tool.

2. Plasma Enhanced CVD (PECVD) Systems

Plasma enhanced CVD systems, like LPCVD systems, began as batch processors for loads of up to 100 wafers at a time. The key advantages sought in the use of PECVD vs. LPCVD were the ability to reduce process temperatures while maintaining or increasing deposition rates. As device geometries grew ever smaller, limiting time-at-temperature became more important in maintaining the material properties and electrical characteristics of the components already in place on partially fabricated devices. As PECVD processes matured, other advantages such as the ability to manipulate thin film material and conformational properties became apparent. The early commercial configurations for PECVD processing were derived from LPCVD technology of the time, with the PECVD process performed in an evacuated (2-10 Torr) hot-wall tube reactor environment (Figure 145). These systems exhibited similar failings to hot wall LPCVD and modern PECVD processing migrated to single wafer cluster tool environments with the advent of VLSI and ULSI processing.

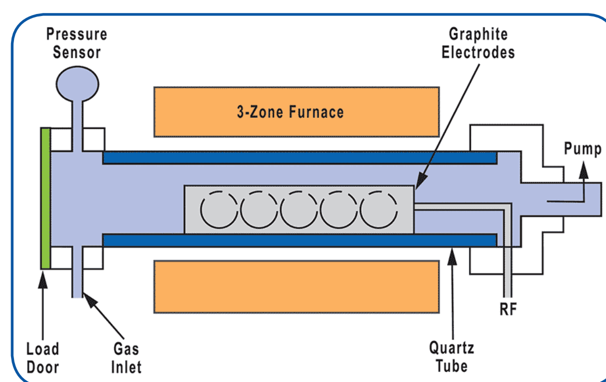


Figure 145. Hot wall PECVD system.

Outwardly, single wafer process chambers for PECVD look somewhat similar to those for LPCVD. Figure 146 shows a schematic that illustrates the characteristics of a single wafer plasma chamber. As with single wafer LPCVD chambers, the precursor gas is fed to the chamber using a showerhead arrangement to ensure uniformity of precursor concentration over the wafer face. Direct exposure RF (radio frequency) PECVD systems typically employ the showerhead as an electrode for the introduction of RF energy to create the plasma. Precursor entering the plasma undergoes electron-molecule collisions, producing high energy excited molecules and molecular fragments that adsorb on the substrate surface and deposit the film. Like LPCVD systems, the wafer sits on a heated platen and the by-product gases are exhausted through ports below the wafer level.

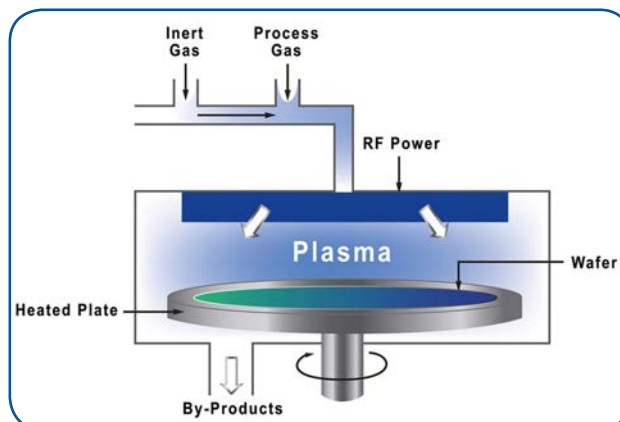


Figure 146. Single wafer plasma chamber.

Methods such as High Density Plasma (HDP)-CVD use remote inductively coupled plasma or electron cyclotron resonance chambers to generate high concentrations of reactive species and ions. In some cases, HDP-CVD reactors are designed with a microwave plasma source for reactive species generation and an ability to bias the substrate for controlled ion bombardment such as is shown in Figure 147. Other PECVD reactor configurations do not expose the substrate to direct plasma exposure so as to avoid ion and electron bombardment. In these configurations, remote plasma generation is combined with ion screening so that only neutral excited species such as molecular and molecular-fragment radicals can reach the substrate.

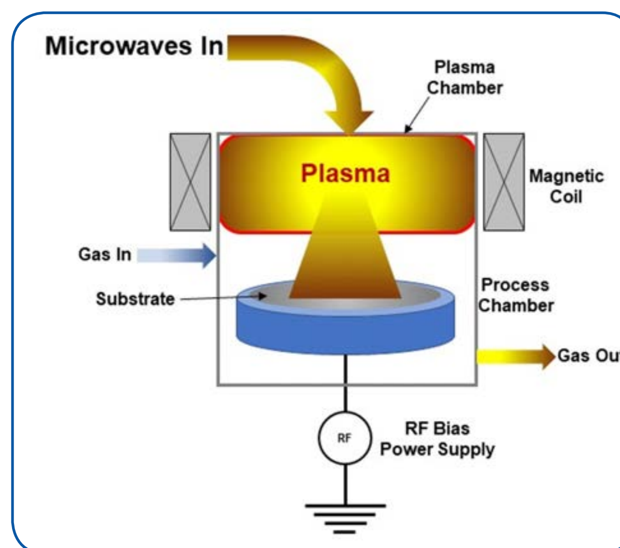


Figure 147. HDP-CVD system [208].

3. Sub-Atmospheric Pressure CVD Systems

The primary application for SACVD equipment is in the TEOS/ozone silicon dioxide CVD process. TEOS/O_3 processes fulfill the need for reduced time-at-temperature in interlevel dielectric processing for ULSI. Additionally, TEOS/O_3 has self-planarizing characteristics superior to other silicon dioxide thin film deposition process.

B. Atomic Layer Deposition (ALD)

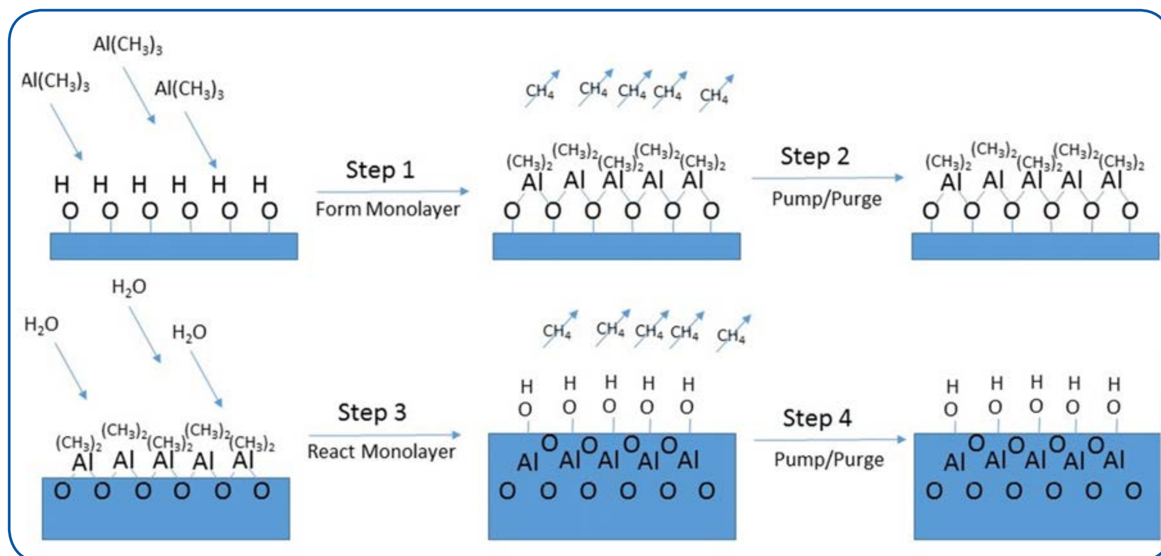


Figure 148. Steps in an ALD process.

Atomic layer deposition is similar to LPCVD except that the chemical process is broken down into steps that isolate different adsorption and reaction steps to have self-limiting reactions. The process employs separate pulses of precursors and reactants that pass sequentially through the process chamber. Figure 148 illustrates the ALD process. With the substrate in the process chamber and under high vacuum, an initial precursor is introduced into the chamber. The molecular character of the precursor is such that it will form a chemically-bound monolayer on the substrate surface (Step 1). Any layers beyond the monolayer are only bound by physisorption forces which are weak enough to allow any precursor other than that in the monolayer to be pumped away under high vacuum. Once the monolayer is present on the substrate, the chamber is re-evacuated and purged to remove any excess precursor (Step 2). Next, a reactant is introduced to the process chamber (Step 3). It reacts with the monolayer material to form the desired compound on the substrate surface (Step 4). By-products of this reaction are pumped away.

In terms of representative chemistry, consider an ALD process for aluminum oxide. A pulse of aluminum alkyl compound, in this case trimethylaluminum, is introduced to the process chamber. The untreated substrate has been prepared prior to the ALD process so that it has a well-ordered covering of hydroxyls on the surface and the aluminum alkyl reacts with hydroxyls that coat the surface forming an Al-O bond and losing a CH_4 group through reaction between the CH_3 ligand and the surface OH group (Step 1). CH_4 is a gas; it is pumped away and any residual in the chamber is removed using a fast inert gas purge (Step 2). The surface is now coated with Al-CH_3 and the reactant, in this case water, is introduced to the chamber (Step 3). It reacts with the Al-CH_3 bonds, generating more CH_4 gas, a bridging Al-O-Al and an Al-OH bond. The residual CH_4 is removed from the system using pump/purge (Step 4). The bridged Al-O-Al becomes part of the growing film and the Al-OH at the film surface presents a new, hydroxyl coated surface that is ready for the ALD process to start all over again.

ALD processing thus requires a very demanding and precise combination of effective precursor delivery and control with process and tool monitoring. The ALD process consists of many cycles of short cycle-time steps employing multiple precursors delivered as very small, tightly controlled gas pulses. The key advantage of ALD processing is the fact that it produces perfectly uniform films over large area substrates and perfect three dimensional conformality in the film. As well, the controlled monolayer-by-monolayer growth allows the user precise control over film thickness. With the proper

selection of precursor and reactants, time-at-temperature burdens can be kept very low in ALD processing. The primary disadvantage of the process is its relatively slow deposition rate, but this is not proving to be a serious impediment to its use for devices having nanometer scale feature sizes.

ALD processes can be either thermally driven or plasma enhanced. Plasma enhanced ALD processes using direct plasma, remote plasma, and combined direct/remote plasma have been demonstrated. They are normally carried out using single-wafer cluster tool equipment configurations. Figure 149 shows schematic depictions of different kinds of ALD reactors.

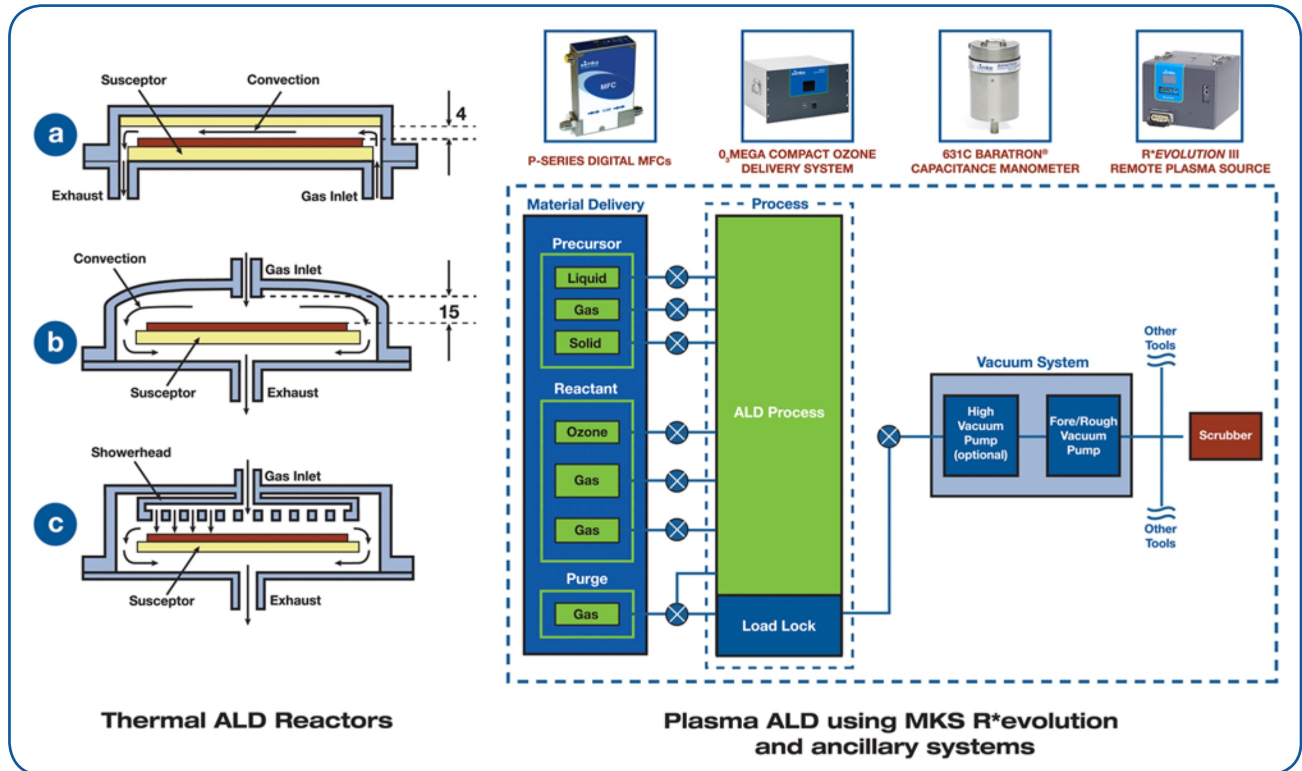


Figure 149. Different kinds of single wafer ALD reactor configurations [209] [210].

C. Physical Vapor Deposition (PVD) Systems

Physical Vapor Deposition (PVD) processes deposit thin films by using physical processes such as evaporation or sputtering. These processes vaporize material from a solid source under high temperature/ high vacuum or plasma and redeposit the material on the substrate surface. Thermal evaporation and sputtering are the primary methods of PVD. Thermal evaporation vaporizes the source material by heating it in vacuum while sputtering uses a plasma to produce ion bombardment (normally argon ions) that creates a vapor from the source material. Films that can be deposited by PVD include most metals and dielectric materials.

MKS Product Applications in Thin Film Deposition

Figure 150 shows a typical single wafer chamber thin film deposition system and all of the ancillary MKS components involved in the operation of the system. This generic diagram is applicable to single wafer LPCVD, PECVD, SACVD, ALD and PVD. Specific MKS product applications are noted in the following pages.

Products for Vacuum Production, Measurement and Control

- See Section B.I.

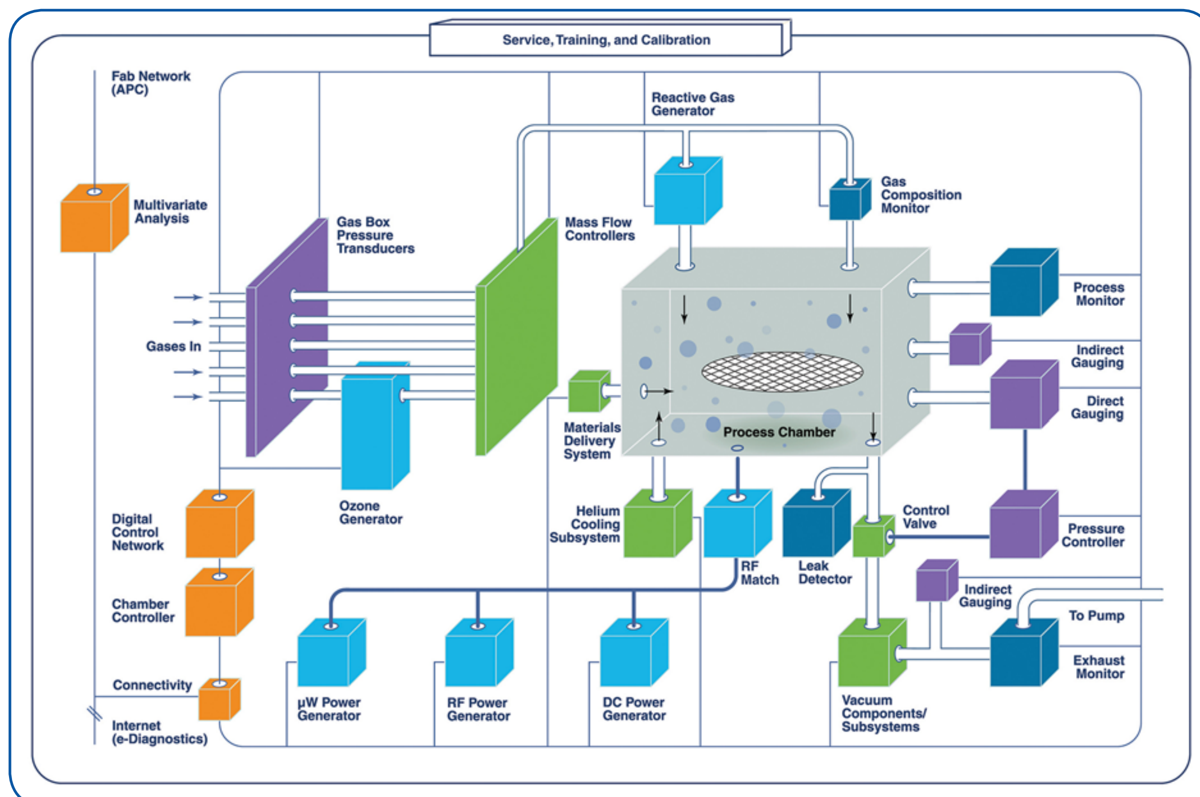


Figure 150. Ancillary systems in thin film deposition equipment.

Products for Precursor Delivery (except ALD)

Thermal and Pressure-based Mass Flow Controllers and Meters

Mass Flow Controllers (MFCs) and Meters (MFMs) are designed to meet a broad variety of processing applications. Most MKS mass flow control and metering products offer both analog (0-5 VDC; 4-20 mA) and digital (Devicenet, Profibus, EtherCAT, RS485) I/O with embedded Modbus and an Ethernet user interface.

G-Series MFCs and MFMs are broadly applicable instruments that provide cost-effective, high performance measurement and control of gas flows between 5 cc/min (sccm) and 250 L/min (slm) with an accuracy of 1% of setpoint. They have multi-range, multi-gas capability and are available in either elastomer or metal sealed configurations.

ALTA MFCs are especially suitable for mass flow control of high flows of hydrogen in pyrogenic oxidation and annealing applications. Hydrogen normally presents a notoriously difficult flow control problem.

I-Series MFCs are also suitable for hydrogen service in pyrogenic oxidation and annealing applications. The I-Series MFCs are IP66-rated for use in harsh industrial environments.

MKS ozone gas delivery products are suitable for a variety of purposes including TEOS/O_3 SACVD. They incorporate field proven, high concentration, ultra-clean ozone generation technology as well as integrated ozone concentration monitor, flow control, and power distribution. Safety monitors, status indicator, and ozone destruct are available on some models.

The O_3MEGA® generator is a compact, integrated ozone solution that incorporates flow control for both O_2 and dopant gas species, as well as an electronic pressure controller (Figure 151). Designed for maximum flexibility, O_3MEGA subsystems are the smallest, most complete ozone delivery systems available.

Also available is the SEMOZON® line of ozone gas generators which includes the SEMOZON AX8575, a fully integrated, high output ozone gas delivery system that can be configured as a multi-channel system delivering ozone for up to 4 channels supporting multiple chambers or multiple tools. It has an optional in-rack chiller for ultra-high concentrations Figure 152.

Products for Flow Ratio Control

MKS offers a line of digitally controlled, web-enabled gas flow ratio measurement and control technology for multi-channel gas flow distribution. The DELTA™ II mass flow controller divides and controls mixed process gas flows to either multiple chambers or zones within a process chamber at ratios specified by the user to maximize process uniformity and repeatability. The DELTA II is used in a variety of flow splitting applications, including etch, strip and CVD (Figure 153).

Plasma Generators for PECVD

MKS RF Generators deliver reliable solid state power for today's thin film deposition applications. MKS RF generators insure consistent, repeatable power delivery in the most critical applications.

The SurePower 13.56MHz product line offers the highest power density available, up to 13kW in a 3U full rack package. SurePower uses patented intrinsic power amplifier protection and embedded V-I sensor to achieve superior reliability and power accuracy.

The MKS KEINOS™ line of plasma generators are designed for pulsed duty applications in environments that experience fast impedance changes such as PECVD (Figure 154). The KEINOS line of generators delivers up to 13 kW of power with pulsing to 20 KHz, multiple set point pulsing, pulse shaping, and frequency tuning. MKS also produces a line of pulsed DC power generators for use in PECVD applications.

Figure 151.
O_3Mega Ozone Generator

Figure 152.
SEMOZON® AX8575
Ozone Delivery System.

Figure 153.
Delta™ II
Flow Ratio Controller.

Figure 154.
KEINOS™
Plasma Generators.

Reactive Gas Generators

MKS offers a variety of plasma-based reactive gas generators designed for both general use and for specific processes (Figure 155). These generators provide reliable sources of active NF_3 and other fluorine-based radical species as well as reactive species derived from O_2, N_2, H_2 and H_2O. These generators find application in deposition technologies in which remote sources of plasma species are needed and in maintenance functions such as the generation of reactive species for CVD chamber cleaning.

The AX7610 is a general duty microwave plasma source for use in remote plasma applications and reactive gas generation. It is designed to be integrated into a microwave plasma subsystem comprised of power supplies, microwave magnetron heads, matching systems, etc. MKS offers a variety of microwave power generators, magnetron heads and system accessories.

Plasma Sources for PVD Thin Film Deposition Reactors

MKS offers a line of RF and DC power generators for applications that include PVD and direct plasma PECVD applications. Product frequencies range from DC to high RF frequencies with power levels up to 100 kW.

Products for ALD Systems

MKS offers the high performance P-Series MFCs and meters for applications in ALD processing (Figure 156). These are metal sealed, pressure insensitive MFC/MFMs with full scales up to 250 slm. Key for precise gas pulse production in ALD applications, they exhibit less than 750 millisecond settling time and are multi-gas, multi-range units with both analog and digital I/O and embedded Modbus and Ethernet user interfaces.

Ozone generation products from MKS are suitable for a variety of purposes including TEOS/O3 SACVD. They incorporate field proven, high concentration, ultra-clean ozone generation technology as well as integrated ozone concentration monitor, flow control, and power distribution. Safety monitors, status indicator, and ozone destruct are available on some models.

The Paragon H* (Figure 155) and R*evolution (Figure 157) remote plasma sources find use in ALD as an ultra-clean source of reactive gas for wafer pre-cleaning to promote film adhesion as well as for generating reactive precursors.

The MKS T3Bi throttle valve, an intelligent, high speed exhaust valve, provides fast downstream vacuum/pressure control for ALD applications. This throttle valve is designed for fast recovery from flow and pressure variations, a pre-requisite for vacuum control in ALD systems (Figure 158).

Figure 155.
ASTRON® Paragon®
Remote Plasma Source.

Figure 156.
P-Series Mass Flow
Controllers and Meters.

Figure 157.
*R*evolution*
Reactive Gas Generator.

Figure 158.
T3Bi Throttle Valve.

Gas Composition and Process Monitors (for CVD, ALD, PVD)

MKS gas composition monitors are based on Fourier-transform infrared spectroscopy (FTIR) and quadrupole mass spectrometer (QMS) technologies. The MultiGas™ 2030 (Figure 159) is an operator-friendly FTIR-based analyzer capable of ppb to ppm sensitivity for multiple gas species process monitoring applications. It has permanently stored calibration spectra, simplifying installation and calibration needs. The Vision 2000-C™ (Figure 160) gas analyzer is an application-specific RGA designed for continuous in situ monitoring of chemical vapor deposition (CVD) and etching processes during chamber clean, passivation and deposition. The Vision 2000-C is designed to detect subtle changes in low concentration species and high mass species decay with respect to time.

MKS also offers the Process Sense™ non-dispersive infrared (NDIR) analyzer for use in monitoring CVD and etch chamber cleaning processes (Figure 161). The Process Sense™ endpoint sensor is a small, low cost partial-pressure analyzer specifically designed to determine the completion of plasma chamber cleaning of deposition chambers. Process Sense is based on near infrared absorption.

Effluent Management Systems

Downstream effluents in some thin film manufacturing processes can lead to yield loss and excessive equipment downtime due to clogged forelines, damaged pumps and valves, and contaminated transducers. Clogged lines simultaneously reduce flow conductance and raise particle counts. Effluent problems can also create serious safety issues. MKS has developed a series of traps and effluent management solutions and vacuum line heaters to help maintain clear piping.

Automation and Control Systems

MKS offers a Programmable Automation Platform (PAC) that can be easily configured to perform standard control and automation tasks (Figure 162). The advanced analytics feature of the Automation Platform provides a smarter automation solution. It can be configured to connect, monitor, control, analyze, and optimize processes—fully delivering on the concept of "learned" automation. The Automation Platform supports the MKS SenseLink™ QM application for real-time multivariate analytics, incorporates knowledge and solutions in design of experiments (DOE) and process optimization, and applies it to controls and automation. In addition to supporting advanced control algorithms (both MKS and third party), the Automation Platform easily integrates with other MKS devices and instruments including pressure gauges, mass flow controllers, valves, gas analysis solutions, etc.

MKS Pressure Measurement

MKS Baratron capacitance manometers are widely used for pressure in thin film deposition.

Figure 159.
MultiGas™ 2030.

Figure 160.
Vision 2000-C™
Residual Gas Analyzer.

Figure 161.
Process Sense™
Endpoint Sensor.

Figure 162.
Automation Platform.

VI. Photolithography

A. Basic Principles and Procedures

The word lithography comes from the Greek words for stone, *lithos,* and to write, *graphia*. It is defined in the Merriam-Webster dictionary as: “a method of printing from a flat surface (such as a smooth stone or a metal plate) that has been prepared so that the ink will only stick to the design that will be printed.” In semiconductor device manufacturing, the stone is the silicon wafer and the ink is the deposition, lithography and etch process that creates the desired feature. Since lithography for device fabrication involves the use of optical exposure to create the pattern, semiconductor lithography is commonly called “photolithography.” An illustration of the photolithography process, in this case to define shallow trench isolation features, consists of the following steps:

1. Substrate cleaning and preparation
2. Form layers of thermal oxide and deposit a layer of silicon nitride on the clean substrate
3. Deposit a carbon hard mask followed by a layer of anti-reflective material
4. Deposit a layer of photoresist
5. Pre-bake the photoresist
6. Align the substrate/photoresist and photomask and expose the photoresist using UV radiation
7. Post exposure bake
8. Develop the pattern in the photoresist and hard bake to remove remaining solvent
9. Perform etch to open dielectric anti-reflective coating (DARC) and hard mask pattern and remove photoresist and DARC
10. Perform etch to open trenches in substrate and remove hard mask
11. Clean surface

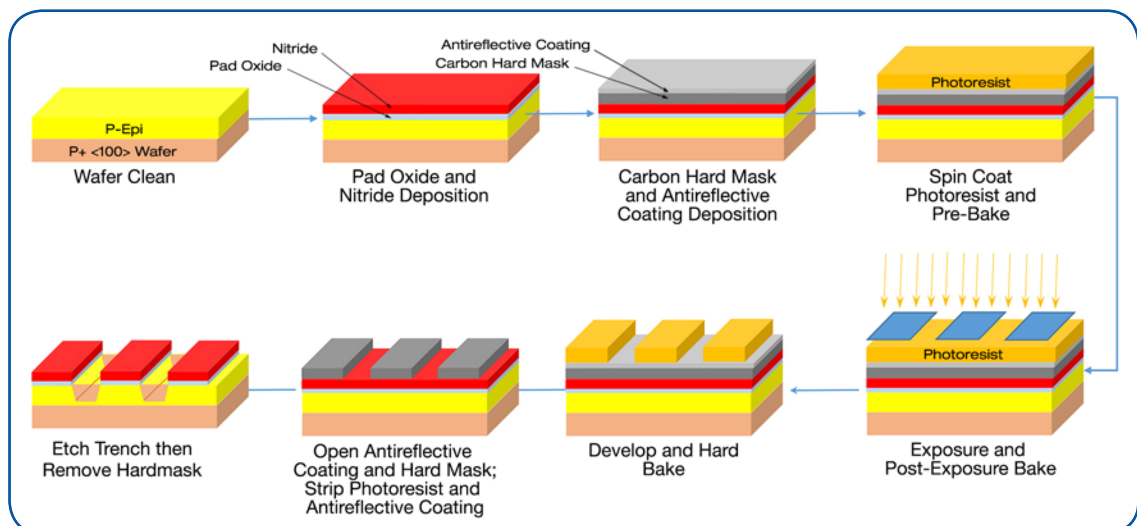


Figure 163. A schematic representation of a semiconductor device patterning process.

mks

Figure 163 shows a schematic of the basic steps in patterning features for silicon device fabrication; we will discuss each step in some detail. The reader is also referred to the section on CMOS Process Flow in Part A of this work to help place the photolithography process in context with device fabrication.

Step 1: The substrate to be patterned must first be cleaned and have its surface treated to improve photoresist adhesion. Fresh unprocessed wafers (as is the case in n- and p-well formation) typically undergo an RCA clean (see Part A). More structured substrates may require a simple degrease, rinse and dry or the removal of residual photoresist from earlier steps using a remote plasma oxygen clean or a dissolved ozone clean, DIO_3. Once the surface is clean, a photoresist adhesion promoter is often employed to improve the wetting characteristics of the photoresist (sometimes called "wafer priming"). The primer makes the surface more compatible at the molecular level with the organic photoresist. In wafer priming, the substrate is first dried (15 minutes at 80-90°C), then exposed to the primer. Primers are usually polar molecules based on siloxane chemistries (e.g. 1,1,1,3,3,3-hexamethyldisilazane, HMDS, $(CH_3)_3SiNHSi(CH_3)_3$). Exposure of the substrate to the primer can be done in either the liquid or vapor phases.

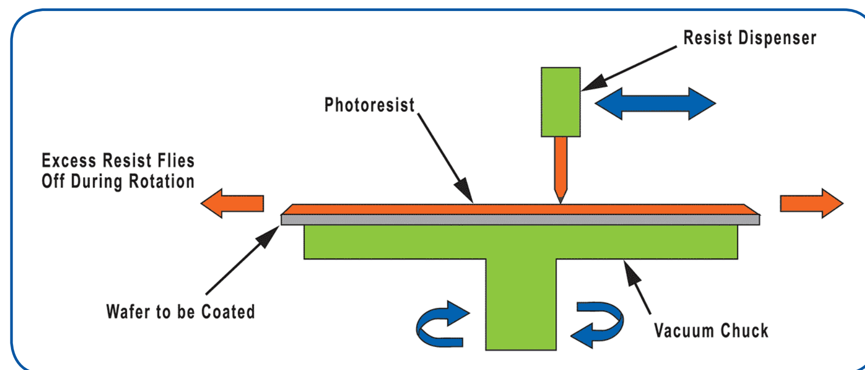


Figure 164. Schematic for a spin coater showing major components and actions.

Step 2: The clean substrate has a thermal "pad" oxide formed on its surface followed by the deposition of a layer of CVD silicon nitride.

Step 3: Using PECVD, a layer of amorphous carbon hard mask followed by a layer of DARC are deposited on the substrate. A hard mask is employed to improve line edge roughness and the integrity of the structure during etch.

Step 4: The substrate is placed on a vacuum chuck in a photoresist spinner. It is then spun at a rate of 3000-6000 rpm. A viscous, liquid photoresist solution is dispensed at the center of the substrate and centrifugal force spreads the photoresist over the entire substrate surface at a uniform thickness. Figure 164 shows the principle components and actions in the photoresist spin coating process. The final thickness of the resist is determined by the resist viscosity and the reciprocal of the spinner rotational speed. Resist layer thicknesses vary according to process needs; they can be as thick as 1-2 μm.

Photoresists are chemical compounds that change their structure when exposed to ultraviolet (UV) light; this, in turn, changes the solubility of the photoresist in the exposed region. Two types of photoresists are used, either positive resists or negative resists. A positive photoresist becomes more soluble when exposed to UV light while a negative resist becomes insoluble in the exposed areas.

Step 5: The pre-bake step (sometimes called a "soft bake") evaporates the photoresist solution solvent from the photoresist layer on the substrate. This is typically done using microwave or IR heating in a production environment, or a convection oven or hot plate in development labs. It is important that pre-bake conditions are optimized to a given device process requirement.

Step 6: Following the pre-bake step, the photoresist-coated substrate is exposed to a pattern of UV light. In modern ULSI technology this exposure is accomplished using projection technology and a step-and-repeat or scanner equipment configuration in which small sections of the substrate are patterned many times across the surface of the substrate. Older technologies for photoresist exposure include contact printing and proximity printing. Photolithography is an immensely complicated and exacting process and we will only introduce the basics of modern methods here. See more detailed discussions of semiconductor photolithography technologies that can be found in standard texts by Wolf and Tauber [32] and by Sze [96].

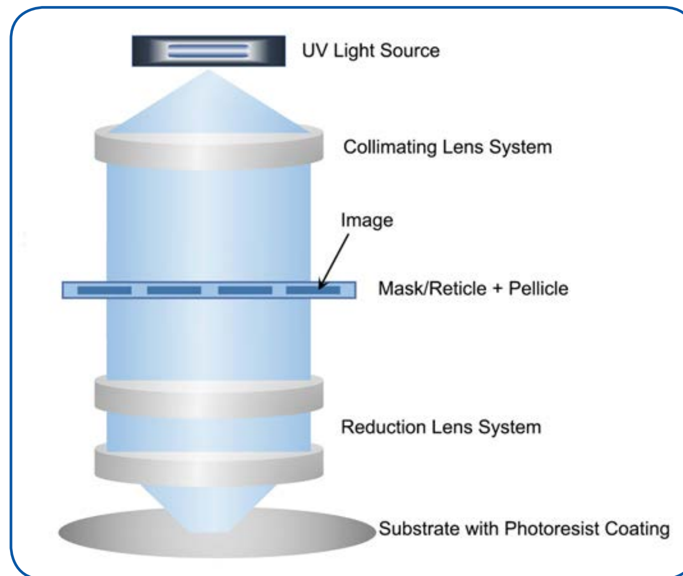


Figure 165. Projection photolithography components [211].

The basic components of projection lithography are shown in Figure 165. In projection photolithography, the substrate and mask are first aligned using registration marks on the mask and wafer. For exposure, UV light is passed through a lens system to produce a collimated (parallel) beam. The collimated beam passes through the photomask (Figure 166) which contains the pattern to be printed on the substrate. In projection lithography, these masks are called reticles, and they are designed to pattern one or more die or a section of a die, depending on die size and pattern complexity. In this method, the light beam containing the patterning information is passed through another lens system that reduces the pattern size to that required in the microelectronics device. Once a die or array exposure is completed, the stepper shifts the substrate (and/or mask) to align the system to the next section of the substrate to be patterned and the process is repeated.

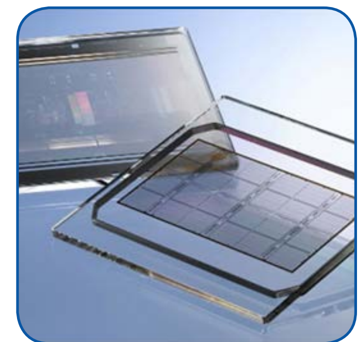


Figure 166. Photomasks (reticles) [213] (Photo provided by Toppan Printing Co., Ltd.)

Step 7: Following exposure of the entire substrate, the patterned photoresists undergo another bake (a “post exposure bake”). This reduces certain optical artifacts in the photoresist pattern (e.g., standing wave effects [212]) and promotes some chemical reactions important to changing the solubility of the developed resist. Additionally, it removes any trace amounts of solvent remaining in the resist.

Step 8: Following exposure, the photoresist pattern is developed by removing the more soluble portions (unexposed regions in a positive resist, exposed regions in a negative resist). Most common photoresists use aqueous (water-based) developers that are sprayed onto a rotating wafer in spin developer equipment much like that used in photoresist coating processes. Typically, the substrate is both rinsed and dried in the spin developer. After the development step, only the desired pattern of photoresist remains on the substrate surface. The photoresist pattern then undergoes a further baking procedure (known as a "hard bake") which hardens the final resist image so that it can withstand subsequent processes such as etching or ion implant. This is typically a high temperature bake, with temperatures of up to 150°C used in the process.

Step 9: The substrate with patterned photoresist undergoes sequential etch processes to replicate the photoresist pattern in the DARC and hard mask coatings. The photoresist is then stripped using an O_2 plasma ashing process, followed by DARC removal in a fluorocarbon etch. The substrate is subjected to a piranha clean to remove any organic residues.

Step 10: This step is the ultimate goal of the patterning process – pattern transfer to the underlying substrate. In ion implantation, incoming ions will not reach the substrate in those areas where hardened photoresist remains. Likewise, in an etch process, only those areas without photoresist will undergo material loss. In patterning for film deposition, those areas with photoresist will have the deposited film removed when the photoresist is removed.

Step 11: The final step is removal of the hard mask using an O_2 plasma strip and a final cleaning step that removes any remaining organic residues.

Note that precise temperature control is required in all of the pre- and post-exposure steps.

1. Deep UV (193 nm) Technology

Deep ultraviolet (DUV) technology for photolithography is exclusively based on projection optics. We will limit our discussion to this approach. Those interested in a detailed discussion of modern semiconductor photolithography technology are referred to available monographs [214] [215].

The physics of light are a strong determinant for the ultimate resolution achievable in a given photolithography process (along with other factors related to substrate and resist properties and design methodologies). Two physical rules determine the minimum feature size that can be created by a photolithography process:

- The smallest linewidth, W, that can be printed is determined by the wavelength, λ, of the exposing light and by the numerical aperture (NA) of the projection optics according to the Rayleigh Equation:

$$W = k_1 \cdot \frac{\lambda}{NA}$$

k_1 is a factor that accounts for the processing characteristics such as quality of the resist and the use of resolution enhancement techniques like off-axis illumination. k_1 has a theoretical minimum of 0.25, although values below 0.3 are considered too difficult or expensive for common use. The numerical aperture of a projection system, NA, is a measure of the optical system's ability to collect and focus the light from the source. Figure 167 shows the relation between the numerical aperture of a lens system and other relevant parameters in the system. In terms of smaller feature sizes, a larger numerical aperture is desirable since it reduces the minimum feature size achievable in the photolithography system. The maximum value for NA of a lens with air operating as the imaging medium is 1.0, although values greater than 0.95 are not normally found. The consequence of these considerations is that feature sizes below a quarter of the exposure

wavelength are theoretically impossible and that practically, feature sizes of slightly less than a third of the exposure wavelength are achievable. This means that, for optical considerations alone, 193 nm systems can process, in practical terms, devices with minimum feature sizes down to about 40 nm. This can be extended with design methods and additional multiple patterning techniques that are widely used at 22nm feature sizes and below (see below).

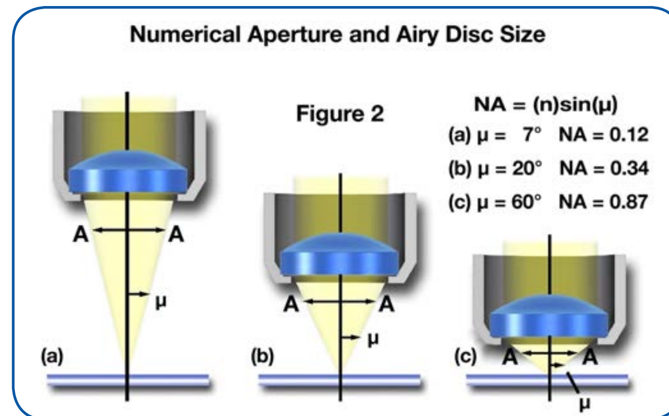


Figure 167. The relationship between numerical aperture (NA), the half-angle of the light cone, and the refractive index of the imaging medium between the lens and the substrate (reprinted with permission of Molecular Expressions.com at Florida State University [216]).

- The depth of focus (DOF) is also determined by the wavelength of the exposing light and the numerical aperture according to:

$$DOF = k_2 \cdot \frac{\lambda}{(NA)^2}$$

k_2 is related to k_1 and has a minimum value of 0.5 µm with conventional resist technologies. Using this representative value for k_2 and a maximum NA of 0.95 shows that resist thicknesses are limited to about 100 nm for 193 nm systems for conventional system characteristics (i.e., conventional resists, etc.). The use of systems with higher k_2 values (systems employing advanced resists such as Polymethyl Methacrylate (PMMA) and other system modifications) allows thicker resists to be used.

These rules highlight the key parameters that have had to be adjusted as photolithography technology has been adapted to produce smaller and smaller feature sizes. The Rayleigh Equation shows that to reduce the feature size you must either reduce the wavelength of the exposing light or increase the numerical aperture of the projection optics. Depth of focus (the vertical distance over which the image remains in focus) must be sufficient to ensure accuracy and precision in the feature size through the entire thickness of a resist. Simplistically, the feature size achievable in a lithography process depends on the relationship shown in the Rayleigh Equation while the process yield is critically dependent on the projection system's DOF shown in the above equation.

Figure 168 shows the historical progression of IC feature sizes and the wavelength of the photolithography light source required to achieve these feature sizes. Up until recently, photolithography equipment designers have focused primarily on reductions in the wavelength of the light used for exposure as the primary method for achieving smaller feature sizes. This is due to the fact that reductions in the wavelength of the light source resulted in less of a reduction in DOF than did increases in the numerical

aperture. Using this approach, lithography equipment has progressed from use of the G-line of the mercury arc lamp in the 1970s and early '80's (λ = 436 nm; useful down to feature sizes of about 450 nm) to the I-line of the mercury lamp in the mid-1980's (λ = 365 nm; useful down to feature sizes of about 380 nm) to the KrF excimer laser in the '90s (λ = 248 nm; useful down to feature sizes of about 250 nm) to the DUV technology which employs an ArF excimer laser light source (λ = 193 nm).

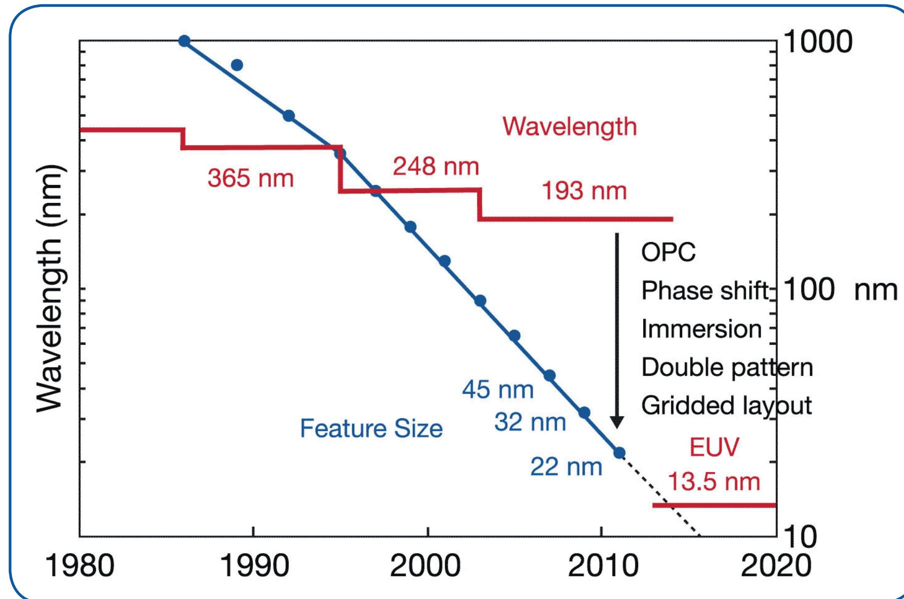


Figure 168. Historical progression of IC feature size and photolithography technologies.

DUV technology employs ArF excimer lasers and this should limit it to feature sizes of 65 nm and greater when air (NA = 1) is the medium between the optical system and the substrate. Obviously, then, semiconductor manufacturers require a new light source with a wavelength less than 193 nm to achieve the needed reductions in feature size for the 45, 32 and 22 nm technology nodes. Unfortunately, commercial light sources at wavelengths below 193 nm have not been available in the past (we will discuss the extreme ultraviolet, EUV source later). In order to get around this problem, device manufacturers have found a number of ways to adapt 193 nm technology to produce feature sizes below 65 nm. By creative use of different combinations of optical proximity correction (OPC – removal of optical distortions), phase shift, immersion lithography, and multiple patterning, manufacturers have proven out 193 nm lithography for the production of feature sizes significantly below conventional expectations (0.3· (λ/NA) = 61 nm for conventional, single-exposure 193 nm dry lithography). Intel, for example, employed 193 nm dry lithography and double patterning to achieve its 45 nm patterning technology. The company then adapted this technology to include immersion lithography and achieved 32 nm patterning technology.

OPC techniques compensate for image distortions that occur when printing feature sizes smaller than the wavelength of the exposing light. Typically, these distortions occur as shortening of line ends, corner rounding or changes to linewidths. OPC corrections are made at the mask level and involve changes to the mask image such as the addition of serifs at design corners and augmentation to linewidths. Figure 169 shows a simple example of how OPC corrections are made at the mask level.

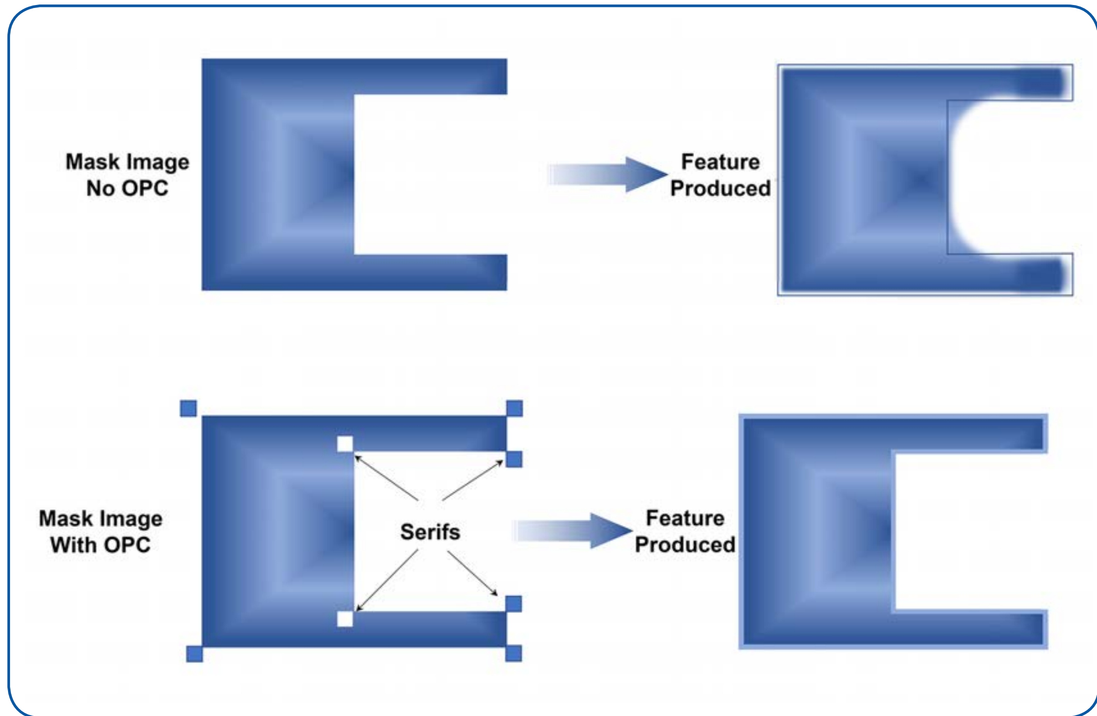


Figure 169. Representative optical proximity corrections to a photomask [217].

Phase shift [218] is a technique that enhances edge contrast in the image being patterned, removing defects that occur due to diffraction limitations at sub-wavelength patterning. Phase shift masks do this by changing the thickness of different sections of the pattern on the mask which changes the phase of light passing through (Figure 170). Reference [214] provides a useful, brief description of how phase shifting works.

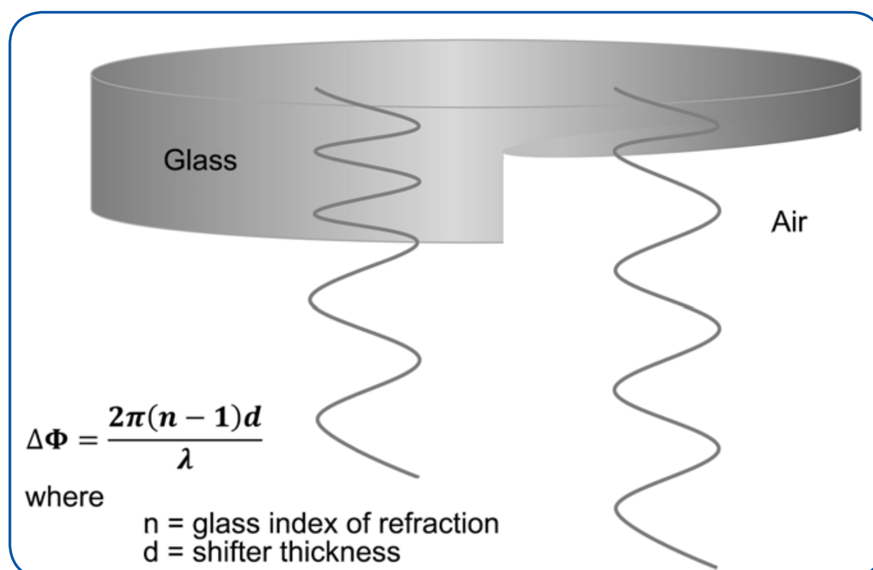


Figure 170. Shifting the phase of light by using different mask thicknesses [219].

Immersion lithography, Figure 171, bypasses the feature size limitations of dry lithography by changing the medium between the optical system and the substrate from air to water. This increases the value of NA beyond 1.0; water has a refractive index of 1.44. Consideration of the Rayleigh Equation shows that the use of water as a medium reduces the minimum single-exposure feature size to about 40 nm when using 193 nm light. Water has come into standard use as the medium in immersion lithography systems. The use of immersion techniques both increases the amount of light that can reach the resist (increasing the resolution) and changes the phase of the light so that it improves DOF. For these and other reasons, single exposure immersion lithography is the dominant approach for patterning in advanced device fabrication processes at design nodes down to 45 nm.

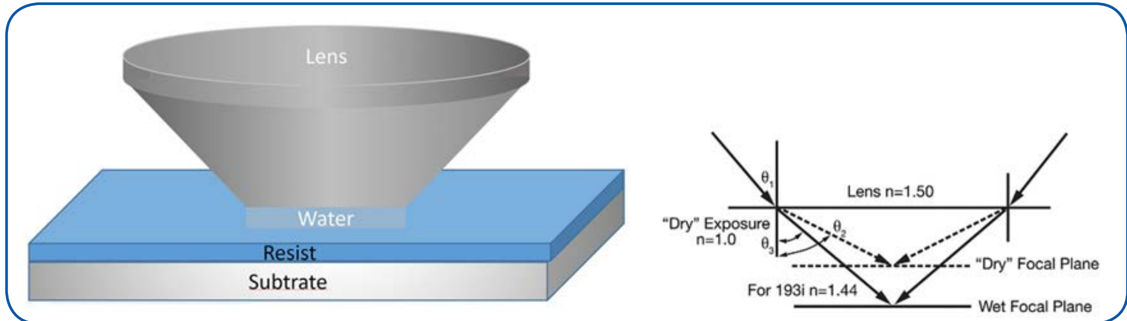


Figure 171. Layout and optical characteristics of immersion lithography [220].

Beyond the 45 nm node, the combination of 193 nm immersion lithography with enhanced techniques such as multiple (i.e., double, triple, quadruple) patterning provides patterning technology until the advent of cost-effective Extreme Ultraviolet Lithography (EUVL). Quadruple patterning uses multiple, shifted exposures such as the process shown in Figure 172, effectively "cheating" the feature size limits imposed by the wavelength of the light source. Even with EUV, multiple patterning will continue to be used. (Quadruple patterning has provided a solution for patterning as low as 5 nm.)

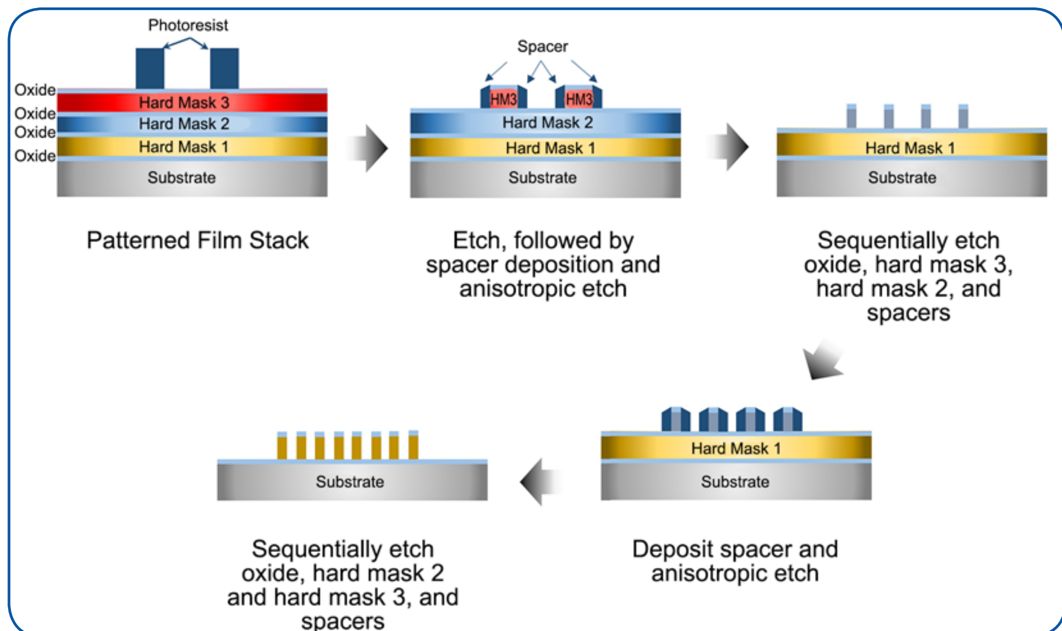


Figure 172. Multiple patterning technique [221].

The optical system in the 193 nm photolithography tool is known as a catadioptric system – the term simply means that it combines both lenses (refractive) and mirrors (reflective) as elements for directing and conditioning the light from the laser. This type of system is advantageous since it accommodates a relatively broad bandwidth of the source laser light with limited chromatic aberration. Refractive elements in the optical system are fabricated from either synthetic fused silica or calcium fluoride, materials that have low absorption of 193 nm light. Photomasks (reticles) in these systems are typically made from fused silica with chrome patterns. Figure 173 shows a schematic of the exposure mechanism and relative motions of the reticle and wafer in a step and scan exposure system along with the manner by which water immersion is maintained between the objective lens and the wafer. In the step and scan process, a slit of light is scanned across one or more die patterned on the reticle. The light passes through the reticle reproducing the part of the pattern on the reticle that is illuminated on the wafer, albeit at much reduced feature size owing to passage through the reduction lens. Simultaneous (and highly precise, accurate and repeatable) movement of both the reticle and the wafer is used to produce the full image of the die on the wafer. Once a die has been patterned, the process "steps" over to the next die area to be patterned. Overlay (the relative position of one patterning layer to another), CD (critical dimension size), and throughput drive the requirement for highly precise motion control in the reticle and wafer stage translations employed in step and scan systems. Typically, overlay tolerances of 15% of the CD are required in 193 nm technologies. Throughput requirements (up to 200 wafers/hr) result in a maximum processing time less than 20 s per wafer. This means that relatively high velocities and accelerations occur in the reticle and wafer translation operations. Motion control systems in these lithography tools must be able to achieve these velocities and accelerations with no impact on vibration levels of the reticle or wafer since this would impact the achievable CDs.

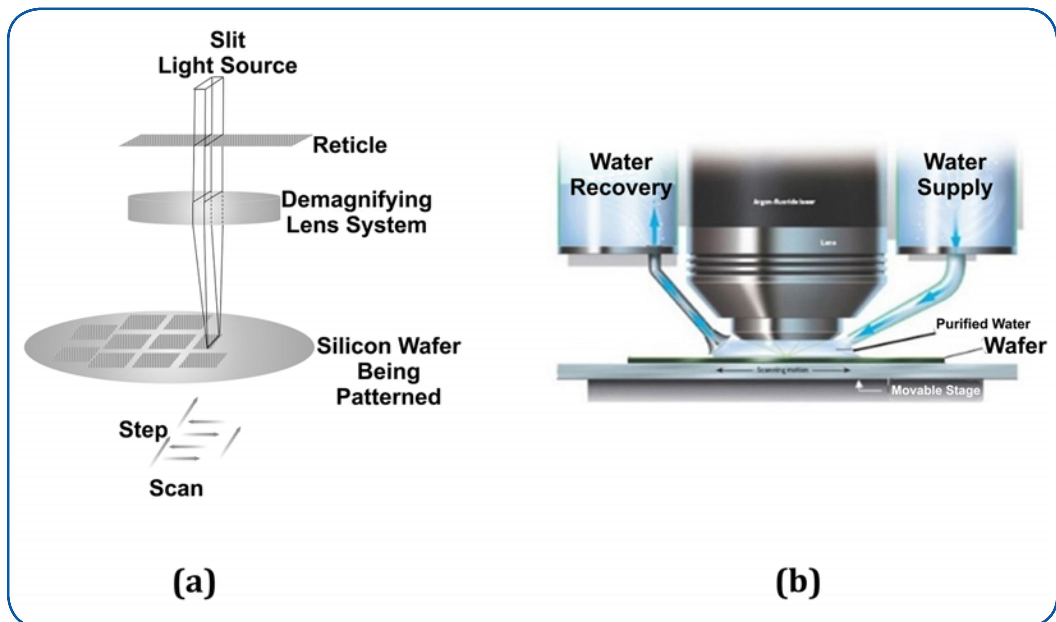


Figure 173. (a) Configuration and relative motions in a step and scan exposure tool [222]; (b) water immersion arrangement [223] (Courtesy of Lachina www.lachina.com).

k_1		NA								
		193 nm			193 nm water immersion				$193i^+$	
		0.75	0.85	0.93	1.07	1.2	1.3	1.35	1.55	1.80
Half-Pitch Size	130 nm	.505	.573	.620						
	90 nm	.350	.396	.429	.499	.560				
	65 nm		.286	.310	.360	.404	.438			
	45 nm					.280	.303	.315		
	32 nm								.257	.298

Table 17. k_1, NA and feature sizes for representative 193 nm patterning processes.

Table 17 provides examples of successful application of 193 nm lithography in IC manufacturing. Note: $193i^+$ refers to extended 193nm lithography aimed at avoiding the use of EUV lithography.

2. Extreme Ultraviolet (EUV) Lithography at 13.5 nm

EUV lithography [224] is being developed to fulfill single-exposure patterning requirements at feature sizes below 22 nm (Figure 174). Unique to this technology is the nature of the light source. There is no readily available conventional light source having a wavelength below 157 nm (the wavelength of light from a F_2 excimer laser). The F_2 excimer laser source has not gained broad use for lithography at small feature sizes. This is probably due to a number of factors, including difficulties with the required CaF_2 optics (increased mask registration errors, birefringence) and the success of the creative extension of 193 nm patterning technology that allowed it to function down to the single-exposure limit of 157 nm lithography (32 nm). Instead, lithography equipment developers have turned to an entirely new way of generating light at extreme UV wavelengths for use in lithography at sub-32 nm feature sizes.

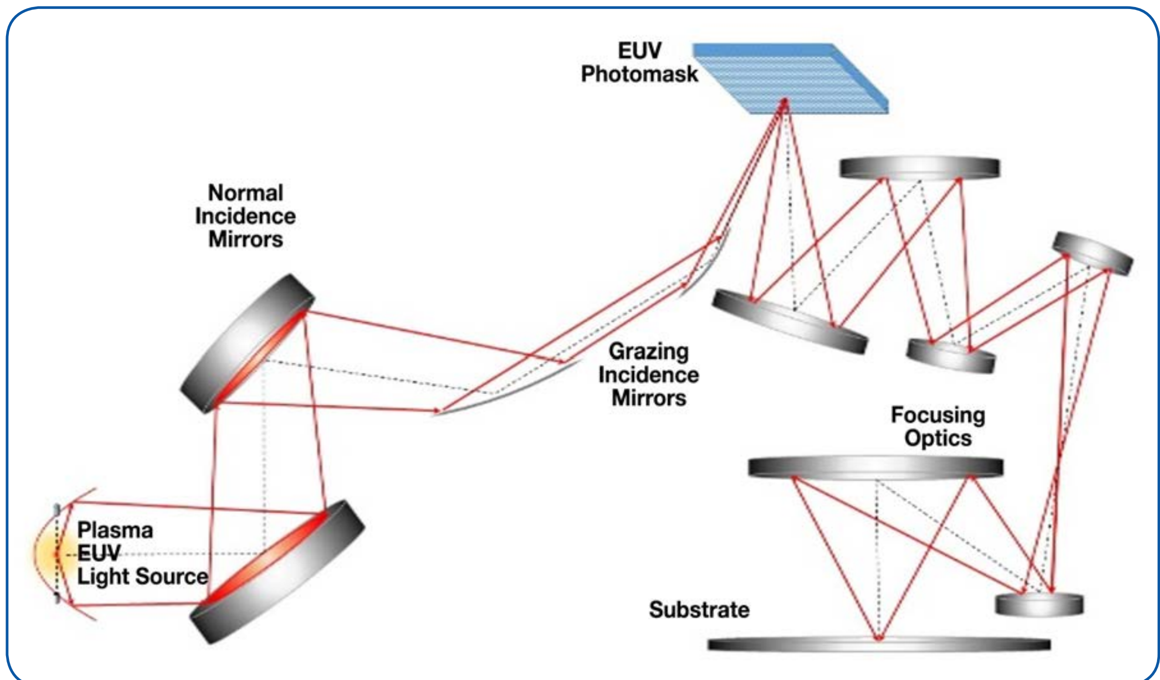


Figure 174. EUV lithography schematic [225].

Light sources of 13.5 nm consist of a high-power CO_2 laser, laser beam transport and focusing optics and a light source vessel. Within this system, 13.5 nm light is generated using laser pulsed plasma (LPP) excitation of tin molecules by a CO_2 laser. The 13.5 nm light is generated inside of a vacuum vessel where droplets of tin pass through the focal point of pulses of high intensity light from the CO_2 laser. As the small (~30 µm diameter) droplets of molten tin encounter the pulse of high intensity laser light, they are vaporized and the tin atoms undergo electronic excitation and ionization, creating a hot plasma with electron temperatures of 10's of electron volts (1 eV = 11,605 K). Electron-ion recombination and ion de-excitation within the plasma emit photons at 13.5 nm. Conversion efficiencies (CEs) (light energy produced divided by energy input to the system) for this mode of light generation are relatively low, with CEs of 3-5% reported for some Sn-based EUV sources. The light emitted by the plasma is collected using an ellipsoidal collector that reflects it to an intermediate focus at the front of the light source. The mirror is unique in that it is not a conventional polished optical mirror, but rather an engineered structure made up of thin layers of molybdenum and silicon designed to reflect the greatest amount of 13.5 nm light (this type of reflector is known as a Bragg reflector). From the intermediate focus, the light is transported to the optical train in the photolithography system using solely reflective optical components. The entire optical system is maintained under high vacuum since 13.5 nm light is strongly absorbed by all solids, liquids, and gases. The mirrors are Bragg reflectors, similar to the collector in the light source that can reflect up to 70% of the incident light. They are critical system components that must have extremely low surface roughness (a few atoms) and highly precise flatness and curvature. Since reflection is not 100%, the number of reflectors between the source and the substrate has a significant impact on the energy of the EUV beam at the substrate surface.

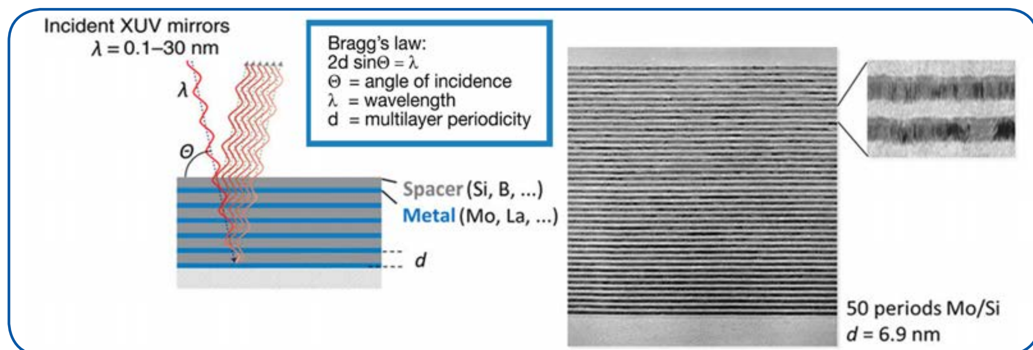


Figure 175. The principle employed in a Bragg reflector; cross-sectional TEM of multilayer EUV mirror grown by e-beam evaporation and ion beam sputter deposition. Courtesy: F. Bijkerk, University of Twente. [226].

Since the optics are all reflective, the masks used in EUV scanners must also be reflective since no optical materials are transparent to EUV. In theory, transmission masks could be used if the substrate was less than 100 nm thick, but this is not a practical solution. EUV masks are fabricated on very low thermal expansion substrates using similar multilayer Bragg reflector technology as the other reflective components in the system. The multilayer reflector has 40-50 pairs of Mo and Si thin film layers each and a total thickness of about 300 nm topped by a capping layer of about 11 nm of Si. A buffer layer, typically SiO_2, is deposited on top of the stack, followed by an absorber of Al, Cr, Ta or W material with a thickness of about 100nm. This absorber layer and the buffer layer are patterned to produce the reflective mask. Figure 175 shows a TEM image of the multilayer stack in a EUV mask while Figure 176 provides a simplified process flow for EUV mask fabrication. Since there are no optical materials that are transparent at 13.5 nm, any pellicles employed with EUV masks must be extremely thin (pellicles are the enclosures used to protect the photomask). TSMC has described a 50 nm thick pellicle membrane supplied by ASML that has a claimed transmission of 85% of the EUV light (useful for lower power EUV) [225]. Great care is required in maintaining these masks in a defect-free state and no defects are permitted in a completed mask. Transport of masks to and from a lithography tool is performed using specialized dual pod containers to ensure mask integrity. There are currently no actinic inspection techniques for EUV masks (i.e. inspection using a wavelength of light that is similar to the feature sizes on the mask)

Photoresist materials for EUV continue to be developed and improved. It is expected that photoresist formulations will have some unique requirements as compared with conventional resist technology. Chemically amplified resists that served the industry well since the introduction of 248 nm exposure wavelengths appear to have certain inherent resolution/sensitivity trade-offs that will force the development of more advanced resist formulations for EUV technology. Currently, hybrid approaches such as those that employ a sacrificial spin-on carbon (SOC) film and a resist layer containing organometallic $R_2M(OR)_2$ molecular linkages is receiving considerable attention [227].

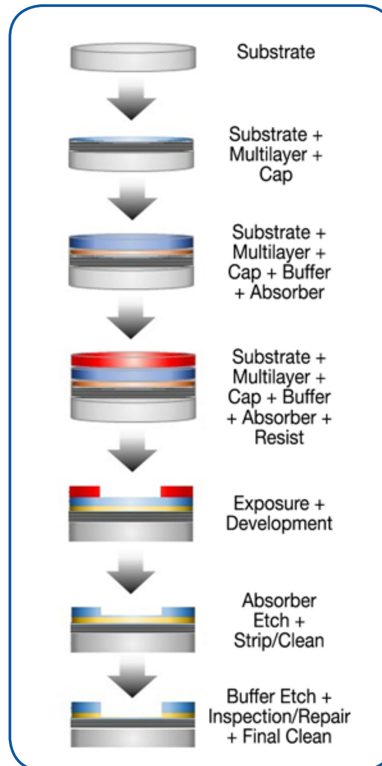


Figure 176. EUV mask fabrication step.

MKS Product Applications in Photolithography

Integrated Solutions Business

MKS Integrated Solutions is a well-established provider of modules and subassemblies for lithography steppers.

Optics

MKS, through its newly acquired Light & Motion Division (formerly Newport Corporation), offers a line of the highest quality optics, including lenses, mirrors, filters, beam splitters, polarizers, optical systems, windows, and many other components and systems. These high performance optic components cover UV, ViS, NIR, and IR wavelengths (Figure 177) and are suitable for photolithography applications—some of which are detailed at the Optics Application Portals.

Figure 177. Specialty Optics.

Motion Control

Precise, accurate and repeatable motion control is a necessity for modern step and scan photolithography tools. MKS offers Precision Motion Control Products with a broad array of general duty and customized motion solutions applicable in wafer patterning and other applications under the Newport® brand. Newport has developed an extensive catalog of manual positioning and motion control standard and custom products for applications ranging from industrial to nanopositioning (Figure 178**)**.

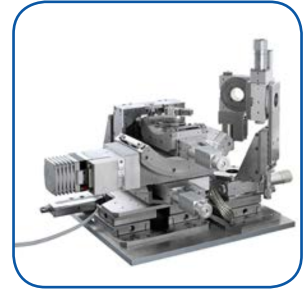


Figure 178. Nanopositioning Stage Assembly

Products for Semiconductor Lithography: The MKS Newport® product line offers a series of high performance air bearing stage solutions (Figure 179) suitable for use in semiconductor lithography applications and customized tools for automated manufacturing and process control. These are extremely rigid structures suitable for use with wafers up to 300 mm diameter. Very high accelerations (up to 5G in some models) and velocities (400 mm/sec to 1000 mm/sec) are achievable while simultaneously retaining high repeatability (25 – 50 nm) and accuracy (0.2 – 0.4 µm).

Figure 179. DynamYX® 300 Wafer Positioning Stage.

Lasers and Light Sources: Many precision positioning applications employ lasers and other light sources for measurement and calibration purposes in the application. MKS Spectra-Physics® product lines offer a broad selection of light sources (Figure 180) suitable for these and other applications.

Figure 180. Spirit® Laser Light Source.

Surface Cleaning Applications

Organics Removal and Photoresist Strip

As a strong oxidizer, ozone, O_3, rapidly reacts with most organic chemical compounds. O_3 reacts directly with hydrocarbons and also generates oxygen radicals that react even more vigorously with organics. Thus ozonated DI water can be used as an effective clean that removes ambient organic molecules adsorbed on the wafer surface. More importantly, when coupled with megasonic agitation, DIO_3 is very effective in removing photoresist residues from a wafer surface.

Advanced Reticle Cleaning

Traditional methods (RCA) for reticle cleaning degrade the optical properties and reticles can only be cleaned between 2 and 8 times before unacceptable degradation occurs due to surface roughening [199] [200]. Since ozone based chemistries produce very little surface roughening, DIO3 cleans are becoming preferred in reticle cleaning applications.

LIQUOZON® Dissolved Ozone Delivery Systems (see Section B.I): MKS offers the LIQUOZON® product line containing a number of different configurations for the generation and delivery of DIO_3. A slightly more detailed description of these products can be found in section B.II.

LIQUOZON® PrimO3 Ozonated Water Delivery System

LIQUOZON® Single Ozonated Water Delivery System

LIQUOZON® Smart High Output Ozonated Water Delivery System

LIQUOZON® Stream High Output Ozonated Water Delivery System

R*evolution® III Remote Plasma Source (see Section B.II): The R*evolution® III's Low-Field Toroidal technology is a stand-alone reactive gas source that produces ultra-clean, ion-free atomic neutrals and radicals for use in, among other applications, advanced photoresist removal processes. The very high flow of radical species that can be achieved in the R*evolution® III (up to 6 slm) produces photoresist strip rates twice as fast as conventional microwave systems.

Type AX7610 Downstream Plasma Source (see Section B.II): The AX7610 is a conventional microwave plasma source that can be used in less demanding photoresist strip processes. It is a conventional microwave source chamber with a quartz or sapphire plasma tube and must be incorporated into a larger system that includes a microwave power generator, waveguide components, and an intelligent matching system.

VII. Etch Technology

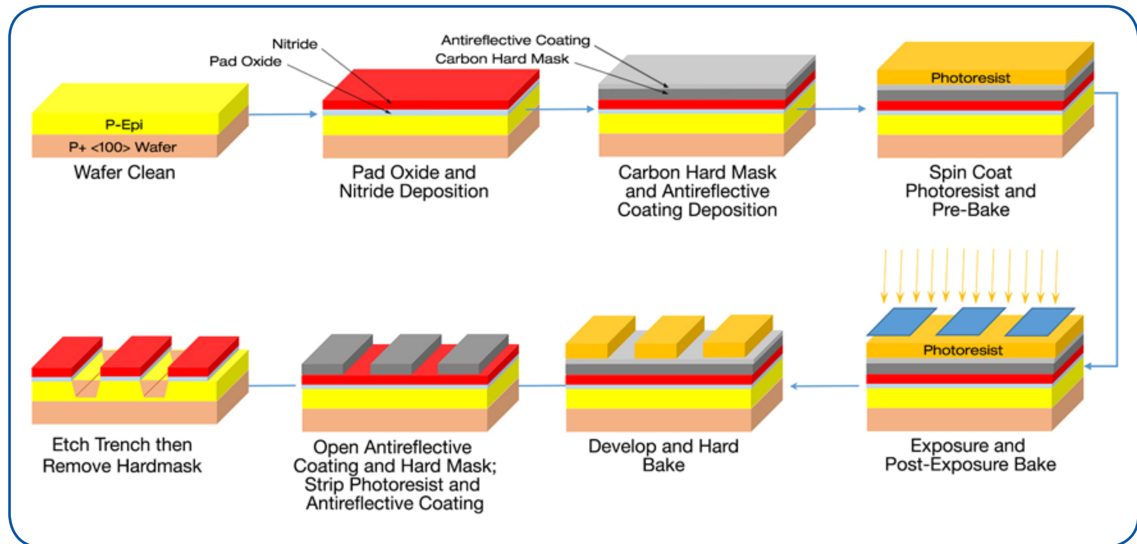


Figure 181. Etching to create a pattern on a substrate [228].

In semiconductor device fabrication, etching refers to any technology that will selectively remove material from a thin film on a substrate (with or without prior structures on its surface) and by this removal create a pattern of that material on the substrate. The pattern is defined by a mask that is resistant to the etching process, the creation of which was described in detail in Section B.VI, Photolithography. Once the mask is in place, etching of the material that is not protected by the mask can occur, by either wet chemical or by "dry" physical methods. Figure 181 shows a schematic representation of this process.

Historically, wet chemical methods played a significant role in etching for pattern definition, up until the advent of VLSI and ULSI technology. However, as device feature sizes were reduced and surface topographies grew more critical, wet chemical etching gave way to dry etching technologies. This shift was due, primarily, to the isotropic nature of wet etching. Wet etching produces material removal in all directions, as shown in Figure 182, which results in a discrepancy between the feature size defined by the mask and that which is replicated on the substrate. VLSI and ULSI designs demand much more precise mask to pattern feature size correlation than was needed at larger feature sizes. In addition, aspect ratios (depth to width ratios) in advanced devices increased and achieving these ratios required an ability to anisotropically etch material using *directional* etching technologies. Figure 183 provides a schematic to help in understanding isotropic vs. anisotropic feature generation and directional etching. The final blow to wet etching's utility in advanced processing may have been the fact that many of the newer materials being used for device fabrication did not have accessible wet chemistries that could be employed for etching. These issues combined to relegate wet etch technologies to nearly exclusive use for cleaning rather than in etching applications. Only devices that have relatively large feature sizes (such as some MEMS structures) continue to employ wet methods for etching. Surface cleaning has been discussed in detail in Sections A.III and B.II.

Anisotropic etching uses a suite of technologies cumulatively known as "dry" etch. These technologies are universally used for etching in VLSI and ULSI device fabrication and they will be the only methods discussed in detail in this Section. Dry etching can remove material through physical means such as ion impact accompanied by ejection of material from the substrate or by chemical reactions that convert substrate material to volatile reaction products that can be pumped away. Dry etching technologies include the following commonly used methods (whether the etch process occurs through chemical etching, physical etching, or a combination as noted in parenthesis):

mks

- Isotropic Radial Etching (Chemical)
- Reactive Ion Etching (Chemical/Physical)
- Sputter Etching (Physical)
- Ion Milling (Physical)
- Ion Beam Assisted Etching (Physical)
- Reactive Ion Beam Etching (Chemical/Physical)

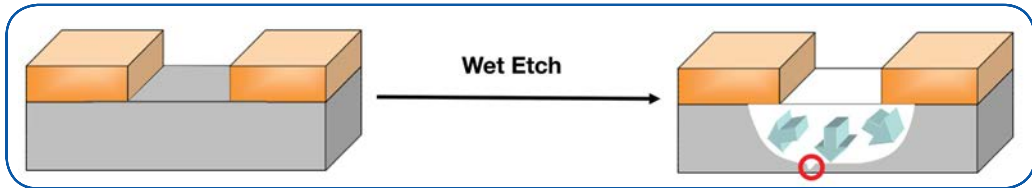


Figure 182. The isotropic character of wet etch processes.

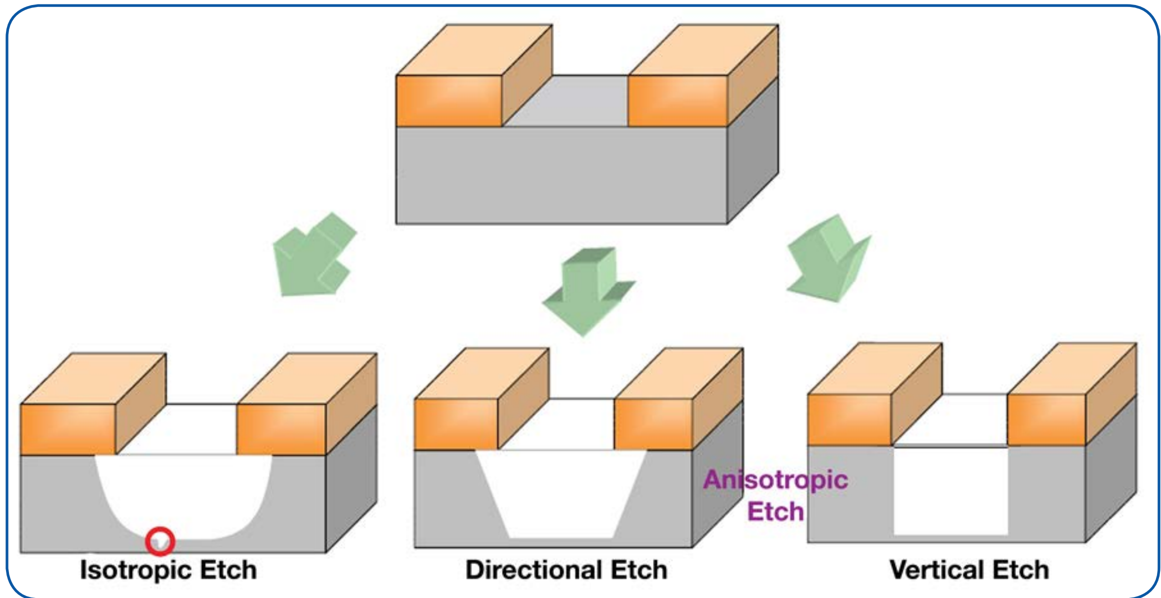


Figure 183. Degrees of directional etching.

All dry etching technologies are conducted under vacuum conditions with the pressure dictating, to some extent, the nature of the etch phenomenon.

Table 18 is taken from Wolf and Tauber [32] and shows the relative pressure regime and generalized characteristics for the different etch methods. While there are a number of specific variants on the equipment and process characteristics used for etching, we will limit our discussion to a brief description of the process basics and the three primary etching methodologies identified in Table 18.

Pressure	Etch Method	Characteristics
< 100 mTorr	**Physical Sputtering and Ion Beam Milling**	• Physical momentum transfer • Directional etch possible • Poor selectivity • Radiation damage possible
~ 100 mTorr	**Reactive Ion Etch**	• Physical (ion) and chemical • Directional • More selective than sputtering
Higher Pressure	**Isotropic Radial Etching**	• Chemical (faster by 10 – 10,000x) • Isotropic • More selective • Less prone to radiation damage

↑ Higher Excitation Energy

Table 18. Pressure regimes and characteristics for different etch methods.

A. Basic Processes

In-depth discussions of plasma etching fundamentals are available in a number of texts (Wolf and Tauber, Sze [32] [96]) and the interested reader is referred to these sources. Here we provide only the briefest description of the basic fundamentals of plasma generation. Within a plasma etching process, a number of physical phenomena are at work. When a strong electrical field is created in a plasma chamber using either electrodes (in the case of a DC potential or RF excitation) or a waveguide (in the case of microwaves), the field accelerates any available free electrons raising their internal energy (there are always a few free electrons in any environment resulting from cosmic rays, etc.). Free electrons collide with atoms or molecules in the gas phase and, if the electron transfers enough energy to the atom/molecule in the collision, an ionization event will occur producing a positive ion and another free electron. Collisions that transfer insufficient energy for ionization can nevertheless transfer sufficient energy to create a stable but reactive neutral species (i.e., molecular radicals). When sufficient energy is fed to the system, a stable, gas-phase plasma containing free electrons, positive ions and reactive neutrals is produced.

In plasma etching processes, the atomic and molecular ions and/or reactive neutrals from a plasma can be used to remove material from the substrate by either physical or chemical pathways, or by mechanisms that employ both. Purely physical etching (Figure 184) is accomplished by using strong electric fields to accelerate positive atomic ions (usually an ion of a heavy inert element such as argon) towards the substrate. This acceleration imparts energy to the ions and when they impact the substrate surface, their internal energy is transferred to the atoms in the substrate. If sufficient energy is transferred, a substrate atom will be ejected into the gas phase to be pumped away by the vacuum system. The incident ion is neutralized in the collision and, since it is a gas, it desorbs into the gas phase to be re-ionized or pumped out of the system.

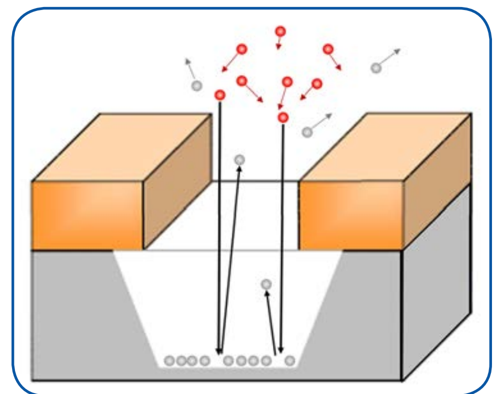
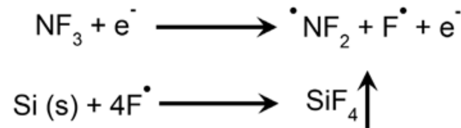


Figure 184. Physical etching of a substrate.

Chemical etching differs from physical etching in that it employs a chemical reaction between reactive neutral species created within the plasma and the substrate material. The most common type of chemical etching involves halide chemistries in which chlorine or fluorine atoms are the active agent in the etching process. A representative chemistry for etch processes is the use of NF_3 for silicon etching. The chemical reaction sequence in this etch process is:



$$NF_3 + e^- \longrightarrow {}^{\bullet}NF_2 + F^{\bullet} + e^-$$

$$Si\,(s) + 4F^{\bullet} \longrightarrow SiF_4\uparrow$$

NF_3 is dissociated in the plasma to produce highly reactive atomic fluorine radicals. These radicals react with silicon in the substrate to produce silicon tetrafluoride, SiF_4, which is a volatile gas that can be pumped away. In this manner silicon is etched from the substrate. Chemical etching, like wet etching, is an isotropic process without directionality (Figure 185). The reason for this is that the sticking coefficient of reactive neutrals is relatively low, so most impacts with the substrate surface do not result in etching, but rather with simple desorption of the reactive neutral back into the gas phase. This phenomenon results in an evening-out of the etch process within the feature being etched and ultimately isotropic character in the etch.

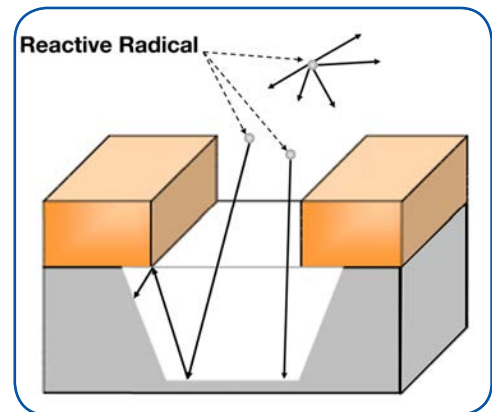


Figure 185. The source of isotropic character in plasma etching [208].

Most of the etching technologies used in modern device fabrication incorporate aspects of both physical and chemical etching. In processes such as reactive ion etching (RIE), directional etching is achieved by biasing the substrate so that ionic species from the plasma are accelerated towards the substrate surface. There they interact with the surface and the reactive neutrals to produce volatile products that can be pumped away (Figure 186). The ion energy in RIE is much lower than that employed for physical etching technologies and ion bombardment effects are negligible. The transfer of ion energies to the surface can enhance the directionality through improved adsorption of reactant species on the bombarded surface (incoming ions create high energy defects where adsorption and reaction preferentially occur) and through enhanced by-product desorption (incoming ion energies are transferred to reaction products causing them to desorb from the surface).

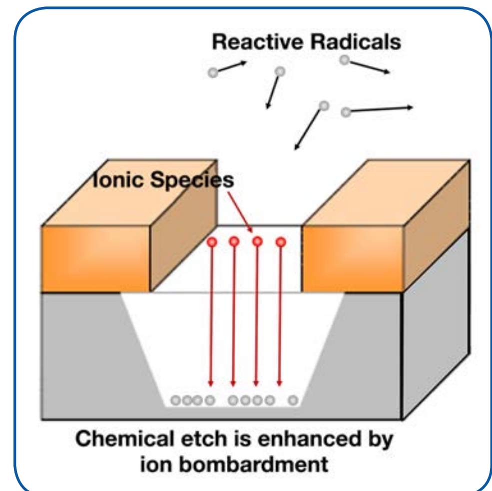


Figure 186. Chemical etch enhanced by low energy ion bombardment [208].

B. Isotropic Radical Etching

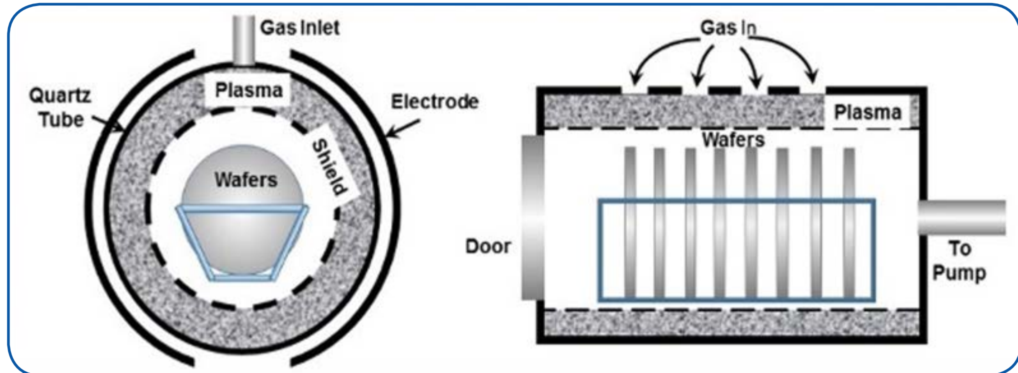


Figure 187. Schematic barrel etcher components.

High pressure plasma etching processes typically operate around 0.2 – 2 Torr. These are glow discharge methods that etch material via a chemical rather than a physical mechanism. The process pressures employed in plasma etching result in very low mean free paths for the ions generated in the plasma preventing any contribution from ion bombardment in the etching process. Thus, while plasma etching exhibits relatively good material selectivity, the purely chemical nature of the etching action results in primarily isotropic rather than anisotropic etching characteristics; this makes plasma etching of limited use in VLSI and ULSI device fabrication. So-called "barrel etchers" that use 13.56 MHz RF excitation are representative of this method. Barrel etchers are older technologies that were commonly used for photoresist strip, isotropic silicon nitride removal, silicon etching for solar cells, and plasma cleaning. Figure 187 shows a schematic of a barrel etcher. The etching process in these systems is fast and causes minimal substrate damage. As is shown in Figure 187, the plasma in a barrel etcher is created in an annulus between the shield and the RF electrodes which surrounds a quartz boat holding the substrate to be etched. A metallic shield protects the substrate from direct exposure to the plasma. Reactive neutrals diffuse through this shield to the substrate surface where they chemically react, removing material in the etching process.

More modern plasma etchers employ a parallel plate configuration that can be easily incorporated into mini-batch or single wafer process tools. Most configurations include the option of applying the RF bias to either the showerhead top electrode or to the substrate holder or both. This gives the user the choice of using simple, isotropic plasma etching (only the showerhead electrode is powered) or directional reactive ion etching (both the top and bottom electrodes are powered and the bias on the substrate attracts reactive ions – see below).

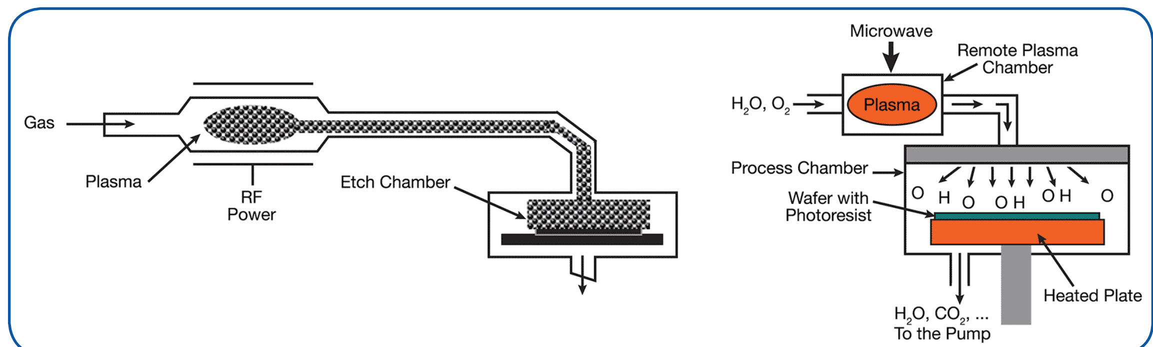


Figure 188. Plasma etch using remote sources [208].

mks

The MKS remote plasma sources that are used in applications such as photoresist removal fall into the category of plasma etchers. In these configurations, the plasma chamber is removed from the substrate as shown in Figure 188. Remote positioning of the source ensures that only reactive neutrals reach the substrate. Ionic species and high energy electrons are destroyed before they can reach the substrate. This type of configuration guarantees minimal damage of the substrate due to ion and electron impact.

C. Reactive Ion Etching

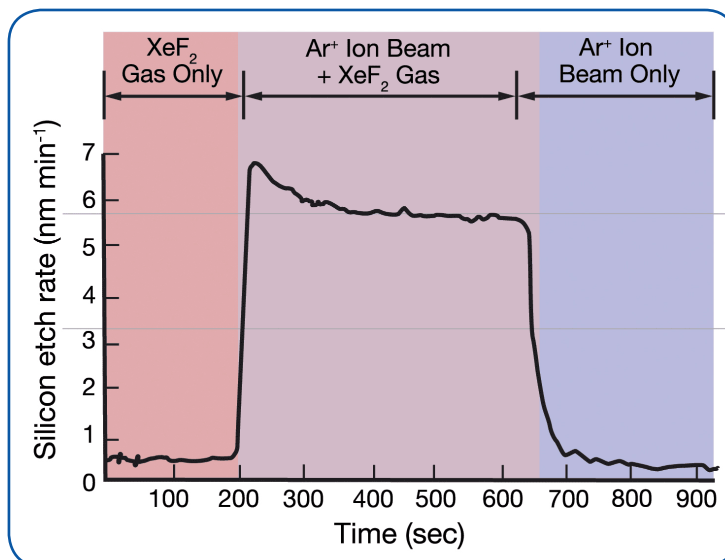


Figure 189. Synergistic effect of ion bombardment in the presence of reactive etchant – principle behind reactive ion etching [229].

Reactive Ion Etching (RIE) uses a combination of chemical and physical reactions to remove material from a substrate; it is the simplest process that is capable of directional etching. A highly anisotropic etching process can be achieved in RIE through the application of energetic ion bombardment of the substrate during the plasma chemical etch. The RIE process thus provides the benefits of highly anisotropic etching due to the directionality of the ions bombarding the substrate surface as they get accelerated towards the negatively biased substrate, combined with high etch rates due to the chemical activity of the reactive species concurrently impinging on the substrate surfaces. The synergistic effect of ion bombardment on increased etch rates in the presence of chemically active species was first demonstrated and explained by Coburn and Winters [229], where they showed the significantly higher silicon etch rate in the presence of both Ar+ ion beam and XeF2 gas compared to either the Ar+ ion beam or the XeF2 gas only as illustrated in Figure 189. In the RIE process, the ions carry sufficient energy to break the chemical bonds of the atoms in the substrate that they impinge upon lowering the activation energy for the chemical etching reactions and thus increasing the reaction rates with the reactive neutrals that are also incident on the substrate surface. In certain etching chemistries, there may be reaction byproducts that are formed on the surface and act as inhibitors for the chemical etch processes. The energetic

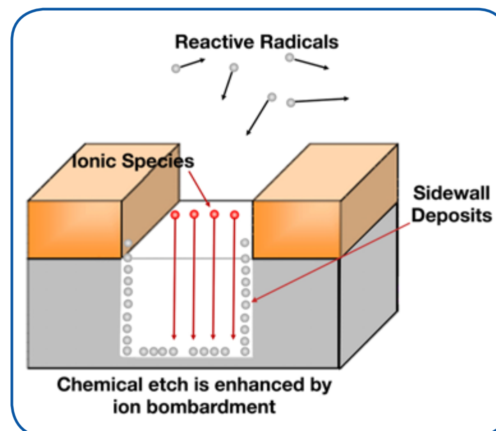


Figure 190. Vertical etch profile in a trench obtained through RIE.

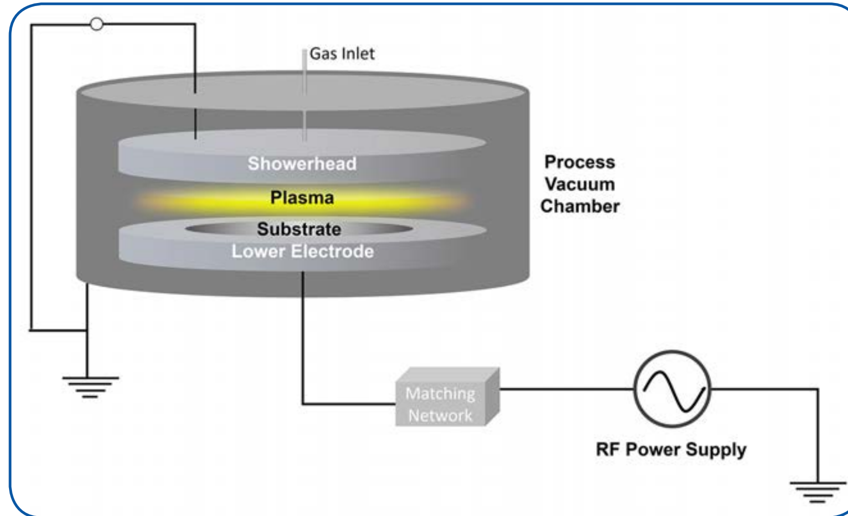


Figure 191. Typical RIE chamber configuration. [230].

ions also act to remove such reaction byproducts from the surface through physical sputtering exposing the underlying material to be removed by the chemically reactive species. Thus RIE is sometimes also referred as Ion-Enhanced Etching or Reactive and Ion Etching. The highly directional nature of ions bombarding the surfaces leads to the anisotropy in the etch process as illustrated in the Figure 190. As the ions are incident perpendicular to the surface, the etch rate is significantly enhanced in the vertical direction as the ions largely interact with the bottom surface of the trench being etched activating the surface and removing any reaction inhibiting byproducts preferentially from the bottom surface while the deposits/ byproducts on the sidewalls act to prevent/minimize the chemical etch from the sidewalls leading to near vertical etch profiles. This high degree of anisotropy is critical in the definition of closely spaced (fine pitch) features in semiconductor manufacturing and is critical in scaling the device features to ever-shrinking dimensions in accordance with Moore's law. The degree of anisotropy in RIE is directly related to the balance between the energy and angular distribution of the incident ion flux and the passivation properties of the reaction byproducts. The incident ion energy and angular distribution is directly related to the electric field drawing the ions towards the substrate and the number of collisions these ions may encounter as they impinge towards the substrate through the accelerating field (function of pressure). Lower pressure and higher substrate bias are key factors that control the ion energy (increase) and angular distribution (reduce) while the polymer formation is largely a function of the polymer forming (CxFy) precursor partial pressure used as one of the etchant gases.

Figure 191 shows a schematic for a typical RIE plasma chamber. RF power is used to excite the plasma and the substrate being processed sits on a powered electrode. The powered electrode enhances the negative bias on the substrate, promoting ion acceleration and surface bombardment. The accelerating voltages experienced by ions in an RIE process are significantly lower than those employed for physical etching processes, producing a much lower potential for substrate damage. Ion energies in RIE are typically between 1 and 30 eV as compared to energies well in excess of 300 eV commonly employed for Physical Vapor Deposition (PVD) approaches. One of the key challenges with such an RIE system is that the same power supply generates the plasma as well as pulls the ions towards the substrate being processed. This configuration is thus

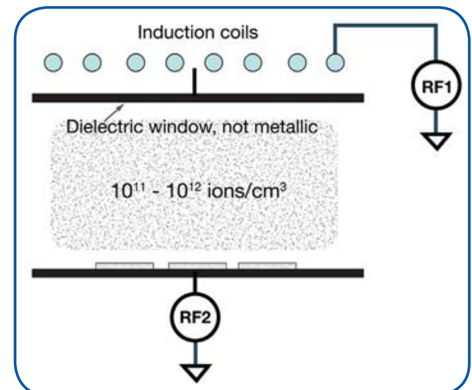


Figure 192. High density plasma etching configuration with dual RF power supplies

limited to capacitively coupled plasmas that typically operate between 10-100 mTorr pressure. For higher etch rates as well as additional process flexibility, inductively coupled plasma sources are used wherein the plasma generation is controlled through a separate RF power supply while the substrate is biased through a second RF supply. Such a system is shown in Figure 192 where the inductively coupled high density plasma is controlled through the first RF power supply (RF1) while the substrate bias is controlled through an independent power supply (RF2). This configuration can enable additional functionality such as modulated pulsed bias as well as operation at lower pressures <10mT for improved etch uniformity across the semiconductor substrates. The pulsing function is especially useful in addressing the issue of Aspect Ratio Dependent Etch (ARDE), wherein the smaller feature structures (higher aspect ratio) etch at a lower rate than the wider features on the same substrate under the same process conditions due to a slower rate of removal/diffusion of the reaction products from high aspect ratio/narrow features. MKS has developed RF generators with unique pulsing capabilities to address these process challenges for semiconductors OEMs manufacturing such HDP etch systems.

D. Physical Sputtering and Ion Milling

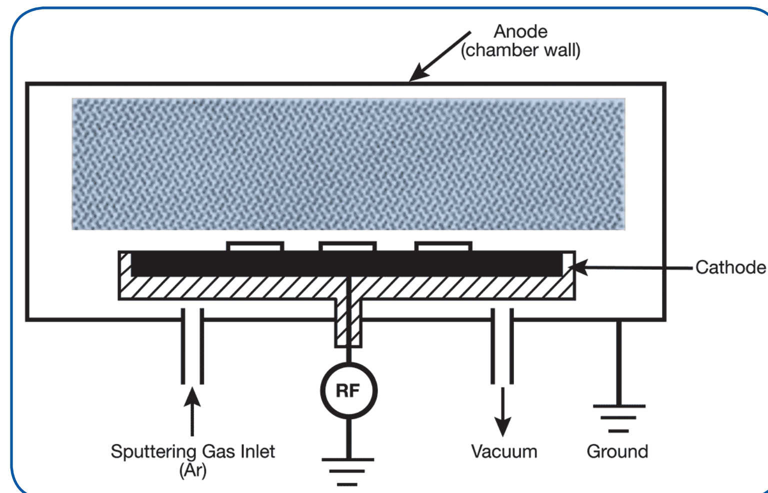


Figure 193. Physical sputter etch tool configuration.

Sputtering (Figure 193) and ion milling (Figure 194) are purely physical processes in which heavy atomic positive ions, created in a plasma, are accelerated to impact a substrate surface to remove material. The ion energies in these processes are high, often greater than 500 eV. Their primary advantage is the fact that they are highly anisotropic processes, producing features with near perfect vertical character. On the negative side, these processes are not particularly selective, having a tendency to etch all materials on the surface. Surface damage is also a problem in these processes. The etch rates that are achievable using physical sputtering are quite low when compared with chemical methods such as plasma etch and RIE. Ion milling is a method that has been developed to increase the etch rate in these processes. In ion milling, heavy ions (Ar^+) are generated in a plasma confined to a discharge chamber separate from the process chamber. A negatively biased grid accelerates the ions out of the discharge chamber to form an ion beam in the process chamber. The process chamber is maintained under high vacuum (10^{-6} Torr) to maintain the coherence of the ion beam until it strikes the mounted substrate. In general, these processes are not widely used in advanced device fabrication. Figure 194 shows a schematic of an ion milling etch system.

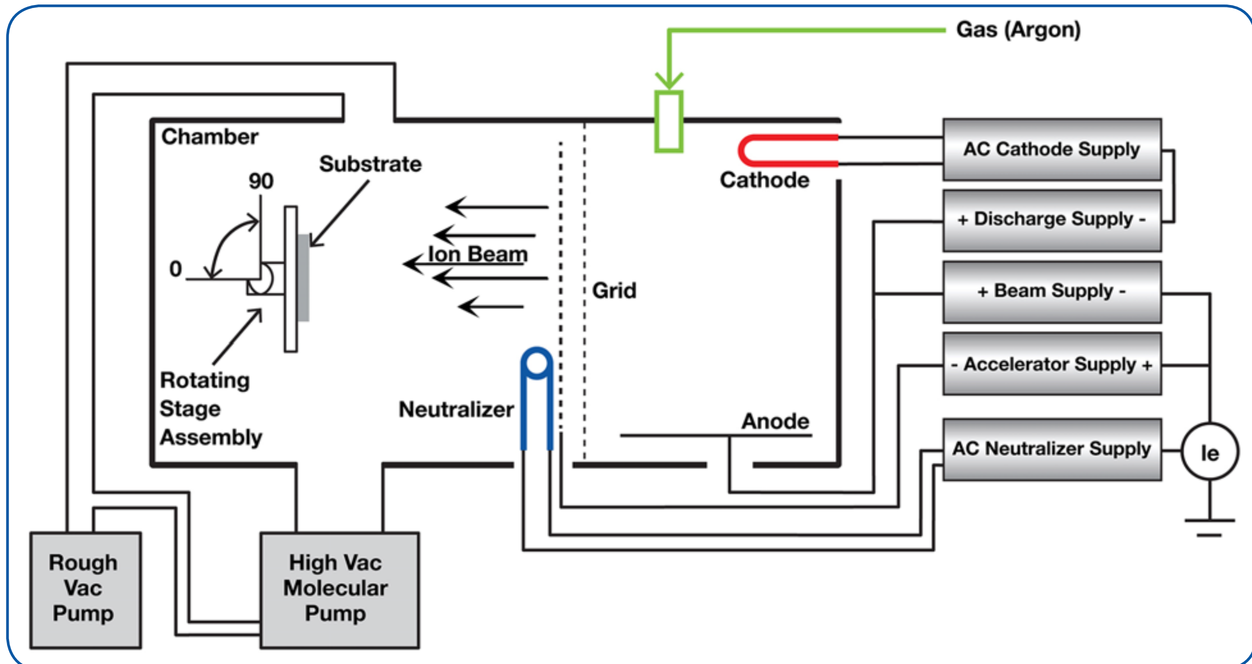


Figure 194. Ion beam etch system. [231] (Used with permission from Microfab Inc., microfabnh.com.)

MKS Product Applications in Etch Technology

Products for Wet Etching

LIQUOZON® Dissolved Ozone Delivery Systems

- See Section B.II

Products for Dry Etching

Medium, High and Ultra-High Vacuum Production, Measurement and Control Solutions

- See Section B.I.
  - MKS Product Applications in Vacuum Production
  - MKS Product Applications for Vacuum Pressure Measurement
  - MKS Product Applications for Vacuum/Pressure Control
  - MKS Product Applications for Residual Gas Analysis and QMS Process Monitoring
- MKS Baratron® Capacitance Manometers are widely used in the measurement and control of process pressures in semiconductor dry etch processes. Up to just a few years ago these processes were typically run at around a few 10's of mTorr. More recently, however, typical processes have extended down to 3 to 4mTorr. Achieving these low pressure set points in etch is necessary for the manufacture of advanced memory and logic devices. However, the etch process is very challenging for capacitance manometers since the gases used in dry etch processes are typically halogen radicals (i.e. neutral, single atomic species of Fluorine, Chlorine and Bromine) which are extremely reactive and corrosive. Thus the capacitance manometer must be able to accurately measure pressure and provide a stable reading without itself succumbing to corrosion which could affect the stability of the pressure reading. MKS Baratron capacitance manometers

are extremely well suited to this environment. The sensor element, which is in intimate contact with the process gas, is composed of high Ni alloy (Inconel®). This alloy provides very good corrosion resistance even in a halogen radical environment. Additionally MKS released a "fluorine-friendly" sensor that provides improved burn-in stability and extended lifetimes in halogen-based processes.

RF Generators for Plasma Etch Applications

- MKS RF Generators deliver reliable solid state power for today's critical etch applications. MKS RF generators insure consistent, repeatable, closed-loop power control.
- SurePower® 13.56 MHz RF Plasma Generators offer state-of-the-art RF topologies, patented intrinsic power amplifier protection, improved design margins, and embedded V-I sensor to achieve superior reliability, reproducibility, and accuracy (Figure 195). Dynamic frequency tuning reduces plasma stabilization time and maintains minimum reflected power.
- The KEINOS™ line of plasma generators is designed for pulsed duty applications in environments that experience fast impedance changes such as PECVD and plasma etch (Figure 196). The KEINOS™ line of generators delivers up to 13 kW of power with pulsing to 50 KHz, multiple set point pulsing, pulse shaping and frequency tuning.
- The GHW Series 13.56 MHz High Reliability RF Plasma Generators offer up to 5.0 kW of power for etching applications. These are field-proven, exceptionally reliable, stable and repeatable generators designed to assure high uptime and process yield (Figure 197).
- MKS also offers pulsed DC power generators for use in PECVD applications.

Reactive Gas Generators

- MKS offers a variety of plasma-based reactive gas generators designed for both general and specific use in etch processing including the ASTRON® Paragon® and R*evolution® ultra-clean remote plasma sources. These generators provide reliable sources of active NF_3 and other fluorine-based radical species as well as reactive species derived from O_2, N_2, H_2 and H_2O. These generators find application where remote sources of plasma species are needed and in maintenance functions such as the generation of reactive species for etch and CVD chamber cleaning (Figure 198).

Figure 195.
SurePower® RF Generator.

Figure 196.
KEINOS™ RF Generators.

Figure 197.
GHW Series RF Generator.

Figure 198.
ASTRON® Paragon®
Remote Plasma Source.

- The AX7610 is a general duty microwave plasma source for use in remote plasma applications and reactive gas generation (Figure 199). It is designed to be integrated into a microwave plasma subsystem comprised of power supplies, microwave magnetron heads, matching systems, etc. MKS offers the SmartPower® intelligent microwave power generator for applications ranging from 180 W to 6 kW. MKS also offers a variety of low power microwave power generators, high power microwave generators (2450 MHz units up to 15 kW and 914 MHz units up to 75 kW), magnetron heads and system accessories.

Figure 199.
AX7610 General Duty Microwave Source.

Plasma Sources for Sputter Etch

- MKS offers a line of RF and DE power generators for applications that include sputter etch applications. Product frequencies range from DC to high RF frequencies with power levels up to 100 kW.

This page intentionally left blank.

VIII. Chemical Mechanical Planarization

A. Chemical Mechanical Polishing/Planarization

Chemical mechanical polishing (CMP) is a planarization technique that was developed for semiconductor applications in the late 1980s and early 1990s. During this period, the number of metal layers increased dramatically and device topographies began to exhibit features that inhibited conformal deposition and gap fill by photoresist, metal, and insulator films. The definition of conformal deposition is demonstrated in Figure 200: a fully conformal film exhibits a ratio of 1:1 for the film thickness on planar versus vertical surfaces; i.e. $T_1=T_2=T_3$. Figure 201(a) shows an idealized image of how conformal films can fill the spaces between features such as metal lines; Figure 201(b) shows how non-conformal films can result in voids in the insulating layers. These voids are electrical weak spots that can result in device failure. As the severity of device topographies increased during the late '80s and '90s many variations on CVD process technology were developed that improved the conformal characteristics of insulating films and allowed their continued use. Eventually, however, the increase in the number of metal layers and more severe topographies forced device manufacturers to move to CMP as a planarization method and, within the context of semiconductor device fabrication, the acronym CMP refers to chemical mechanical planarization. Figure 202 demonstrates the difficulties encountered in fabricating a complex multi-level metal device structure using (a) non-planarized deposition technology versus (b) CMP-based planarization [232].

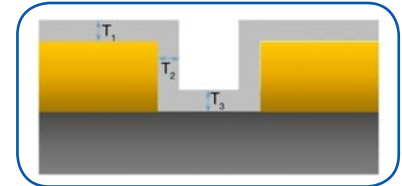


Figure 200. Parameters that define conformal film coverage.

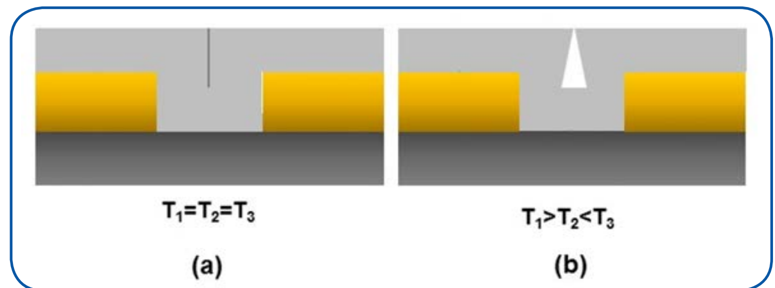


Figure 201. (a) Conformal gap fill; (b) Non-conformal gap fill.

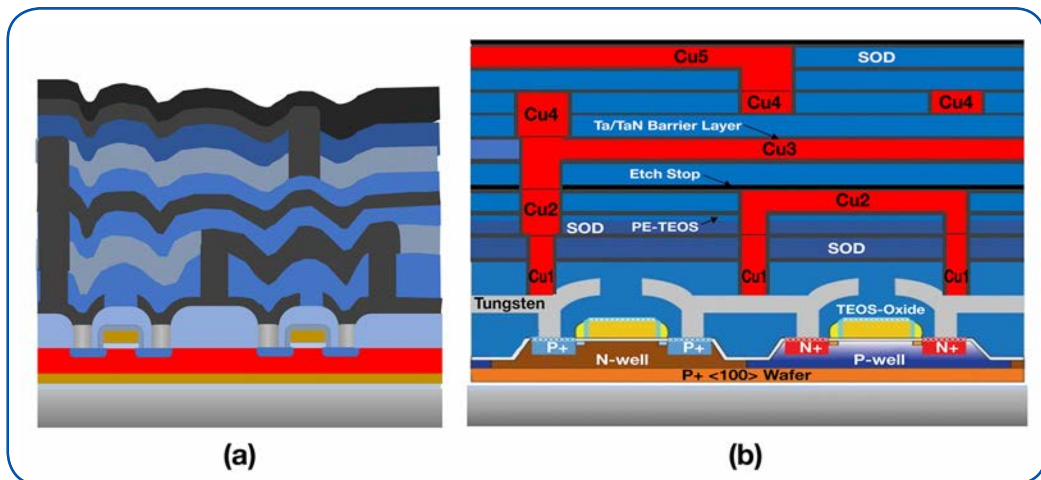


Figure 202. Schematic showing (a) non-planarized and (b) planarized multi-level metal structures [232].

mks

Chemical mechanical planarization, as its name implies, is a physical polishing process in which the surface of a substrate is smoothed and planarized through the combined action of chemical and physical abrasive forces on the surface. CMP combines the best of both techniques while avoiding the pitfalls. Whereas purely abrasive grinding of the surface would cause too much physical damage and purely chemical etching would not achieve planarization, the combined action of the two produces a well-planarized surface with minimal damage. Figure 203 shows a schematic of a chemical mechanical planarization tool (the image has been taken from reference [233]). The tool consists of a rotating platen covered by a polishing pad. The wafer is mounted face down in a carrier that is pressed against the pad with a specified force. This force can be provided using either a defined and regulated gas pressure or a mechanical backing pressure system. The wafer also rotates during the polishing process. The polishing pad is saturated with a slurry of physical abrasive and chemical etchant that is pumped on to the pad. Polishing of the wafer surface occurs as the wafer is rotated on its own axis and moved about the polishing pad while being forced against the pad. During the polishing process, high points on the wafer surface are naturally subjected to more pressure and therefore more abrasive force. This combines with the action of the chemical etchant to produce an enhanced removal rate for material at the high points relative to material at low points in the surface topography. This produces the planarization effect in the process. Since heat is generated in the polishing process and changes in temperature effect etch rates, it is important to maintain a constant temperature at the pad/wafer interface during the CMP process. This is accomplished through active temperature control of the platen.

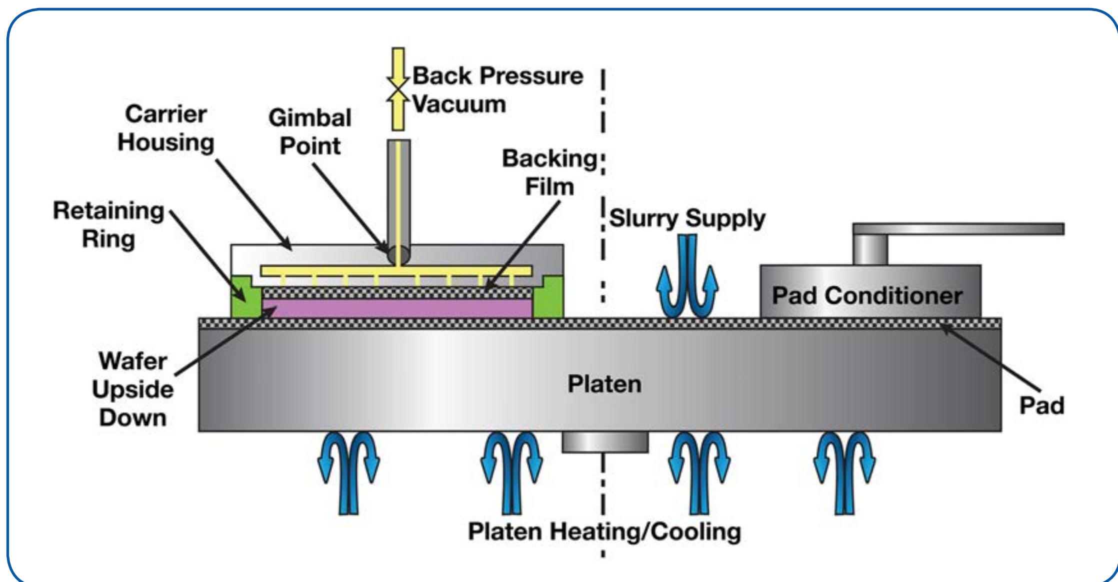


Figure 203. Schematic of a CMP tool (from [233]).

Chemical mechanical planarization has a number of advantages for semiconductor device manufacturing beyond the fact that it reduces rough topography to a planarized state. CMP allows the device manufacturer to achieve global planarization of the entire wafer surface in a single step. The approach can be used to planarize a wide range of materials, from different metals to different oxide films. It can planarize different materials in the same step (i.e., metal and insulating films). CMP is essential to the only effective method of copper patterning (damascene processing, see below). The subtractive nature of CMP helps in reducing surface defects. Finally, the process is relatively environmentally benign with no need of the hazardous gases common in dry etch processing.

Chemical mechanical planarization is employed in different planarization applications during device fabrication. Most important are oxide planarization for shallow trench isolation (STI) and Damascene processing. These applications will be described in some detail to provide practical examples of CMP.

B. CMP in Shallow Trench Isolation Processing

Shallow trench isolation (STI) uses trenches etched into the substrate and filled with undoped CVD polysilicon or CVD silicon dioxide as electrical isolation for the active regions in a device. It has replaced the use of thermal oxide for local oxidation of silicon (LOCOS) due to reasons related to thermal budget and the localized physical impacts of thermal oxidation of silicon. Figure 204 shows the STI process flow. Chemical mechanical polishing is used in the next to last and last steps of the STI process sequence. Following the deposition of the insulating oxide that fills the isolating trenches, CMP is used to planarize the oxide level with the nitride layer. In the last steps of the process, the nitride layer is removed followed by planarization of the CVD oxide in the trenches. STI process technology has been one of the enablers of nanometer-scale device fabrication owing to the fact that earlier LOCOS isolation schemes could not be successfully scaled down to the nanometer regime.

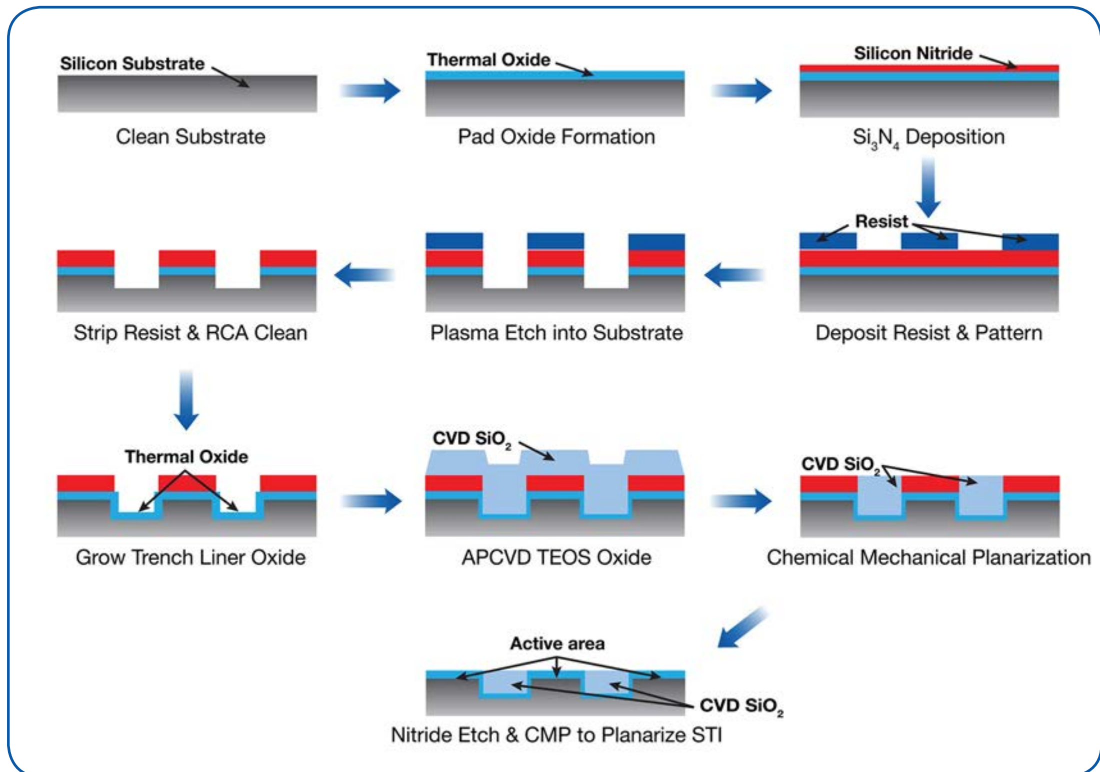


Figure 204. Shallow trench isolation process flow.

C. CMP in Damascene Processing

In the late ‘80s, device speed was becoming limited by a phenomenon known as RC delays (where R is the resistance of the wiring and C is the capacitance of parasitic capacitors in the circuit structure). Signal transmission on a device is delayed in direct proportion to the resistance of the wire on which the transmission occurs and to the capacitance of any unintended capacitors formed when two metal wires are separated by an insulating material. One obvious way to reduce such delays is to reduce the R of the wiring. This prompted the replacement of aluminum (resistivity = 2.8×10^{-8} Ω-cm) wiring by copper (resistivity = 1.7×10^{-8} Ω-cm). It is noteworthy that copper also has other properties that make it preferred over aluminum as the dimensions of the wires on a device shrink (i.e., greater resistance to electromigration).

mks

The use of copper presented serious challenges to device manufacturers. Copper cannot be etched using conventional, halide-based dry etching processes since the copper halide products are not volatile and hence cannot be pumped away by a vacuum system. This means that the conventional technology for patterning metal lines could not be used with copper. Additionally, copper does not adhere well to dielectric material and copper atoms are very mobile in SiO_2. Direct deposition of copper on insulating oxide layers thus presented problems in achieving stable wiring structures and in terms of contamination of the insulating oxide (producing increased leakage).

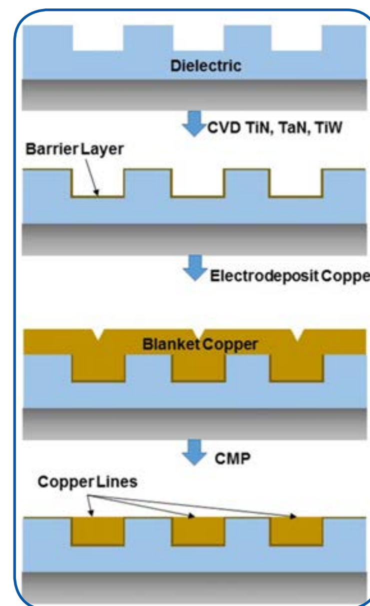


Figure 205. The basic Damascene process.

The Damascene (and Dual Damascene) process, first introduced by IBM in the early 1990s, was a unique, additive processing technique developed to address the challenges presented by the shift to copper wiring in microelectronic circuits. It gets its name from metal inlay techniques developed in Damascus in the Middle East. The process eliminates the need to dry etch copper by using chemical mechanical polishing instead and it employs special barrier layers to prevent diffusion of the copper into the oxide insulating layers. The first step in the process is the formation of the wiring pattern as etched lines in a dielectric layer (the top structure in Figure 205). A barrier layer of, for example, TiN, TaN, or TiW is then deposited on the dielectric layer to act as a barrier between the copper and the insulating dielectric. A thin seed layer of copper is deposited on the barrier, typically using PVD methods, followed by the electrodeposition of a thick copper layer. Chemical mechanical planarization is then used to remove the excess copper leaving only the metal in the etched lines behind. Several sophisticated variations on the Damascene process have been developed, but these simple steps suffice to describe the basics of how the process works.

MKS Product Applications in Chemical Mechanical Planarization

MKS's Pressure Control product line can be applied in the systems that control the backing pressure applied to the wafer holder. Controllers such as the 244E Pressure/Flow Controller can be used with different MKS gas flow control valves to control the backing gas pressure on wafer carriers in CMP tools with pressures ranging from vacuum up to 100 psig (Figure 206).

Figure 206. 244E Pressure/Flow Controller.

IX. Wafer and Reticle Inspection

A. Basic Principles

The progression to ever smaller feature sizes on both substrates and mask/reticles produces lower and lower tolerance to defects in both starting materials and finished devices. Not only are we approaching near zero tolerances to known defect types (particles, crystal defects, etc.), but manufacturers continue to discover device sensitivities to entirely new types of defects as manufacturing progresses into the deep nano-scale. In addition, the available metrologies for defect detection are being required to sense and quantify defects at very near the noise level of their operating principles and new approaches to defect detection are continually being developed.

In this chapter, we provide a brief description of the basics of wafer and reticle inspection techniques and discuss the characteristics of current inspection tools.

1. Unpatterned Wafer Inspection Systems

Historically, device manufacturers have used optical detection systems to inspect wafers and masks for the presence of particles and other types of defects. Modern wafer inspection tools detect particles and pattern defects and determine the position of these defects in an X-Y grid on the wafer which can be output to the user for diagnostic purposes.

Bare substrate wafers are typically inspected when shipped from the wafer manufacturer and on receipt by device manufacturers. Bare or unpatterned wafers may also be monitored before and after being subjected to a passive or active process environment in order to determine a baseline for particle contributions from a given process.

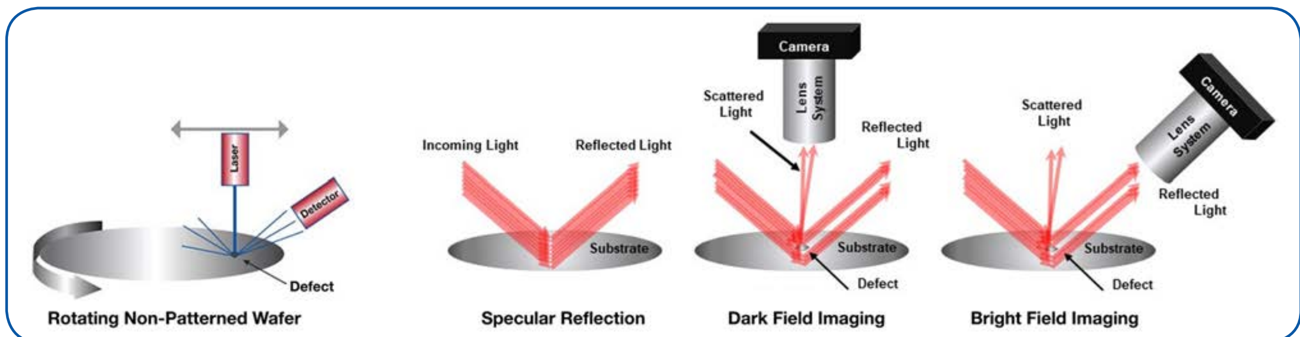


Figure 207. Defect detection on non-patterned wafers showing the difference between dark-field and bright-field image illumination [234].

The basic principle used for defect detection on unpatterned wafers is relatively simple. A laser beam is radially scanned over the surface of a rotating wafer to ensure that the beam is projected onto all parts of the wafer surface. The laser light is reflected from the surface as it would be from a mirror, as is shown in Figure 207 (the term you will hear for this type of reflection is “specular reflection”). When the laser beam encounters a particle or other defect on the surface of the wafer, the defect scatters a portion of the laser light. Depending on the illumination arrangement, the scattered light can be detected either directly (dark-field illumination) or as a loss in intensity in the reflected light beam (bright-field illumination). The rotational position of the wafer and radial position of the beam define the position of the defect. In dark-field imaging, the scattered light is collected to image the defect while bright-field imaging uses the light reflected from the wafer surface. In wafer inspection tools, the light intensity is electronically collected using a photomultiplier tube or similar detector and a map of the scattered or reflected intensity over the wafer

surface is generated, as shown in Figure 208. This map provides information on defect size and location and on wafer surface conditions such as "haze" (microroughness of the surface). As one might imagine, this method requires highly accurate and repeatable rotary and linear motion control of the wafer stage and optical components. In general, dark-field inspection is preferred for non-patterned wafer inspection since high rastering speeds are possible in scanning the wafer surface which permits high wafer throughput. Patterned wafer inspection, on the other hand, may use either bright-field and/or dark-field imaging to inspect patterned wafers, depending on the application, and the inspection takes considerably more time.

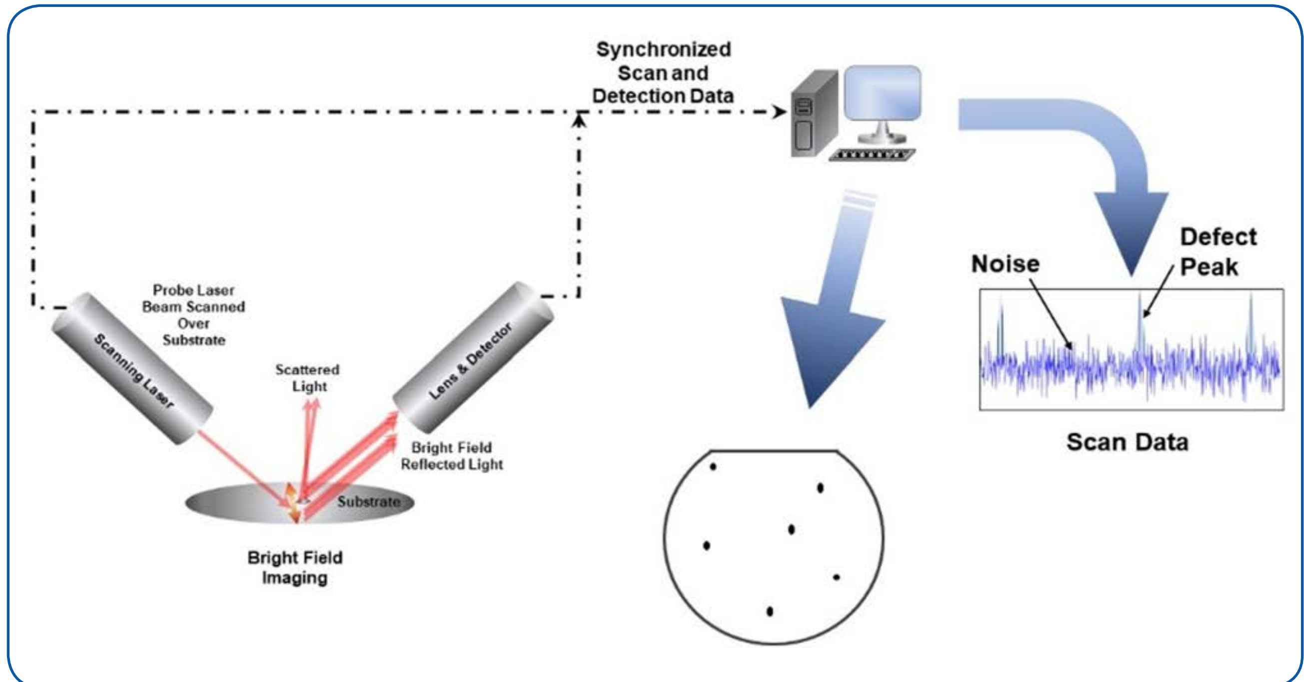


Figure 208. Light collection, processing and wafer mapping in an inspection tool [235].

2. Patterned Wafers

As noted above, optical inspection of patterned wafers can employ bright-field, dark-field illumination or a combination of both for defect detection. Additionally, electron beam (EB) imaging is also used for defect inspection, especially at smaller geometries where optical imaging is less effective. However, it is very slow and only used at the R&D stage. Patterned wafer inspection systems employ a comparative method in which they image a given die on the wafer and compare that image with the image of an adjacent die or that of a "golden" die known to be defect free. Inspection tools use image processing software to subtract one image from the other and if there is a random defect in one of the die, it will not zero out in the subtraction process, showing up clearly in the subtracted image (See Figure 209). The position of the defect is defined in terms of the position within the die and the die position on the wafer allowing a defect map to be generated over the wafer, similar to that generated with unpatterned wafers. As with the inspection of unpatterned wafers, patterned wafer inspection requires precise and repeatable motion control of both the wafer stage and the system optical components.

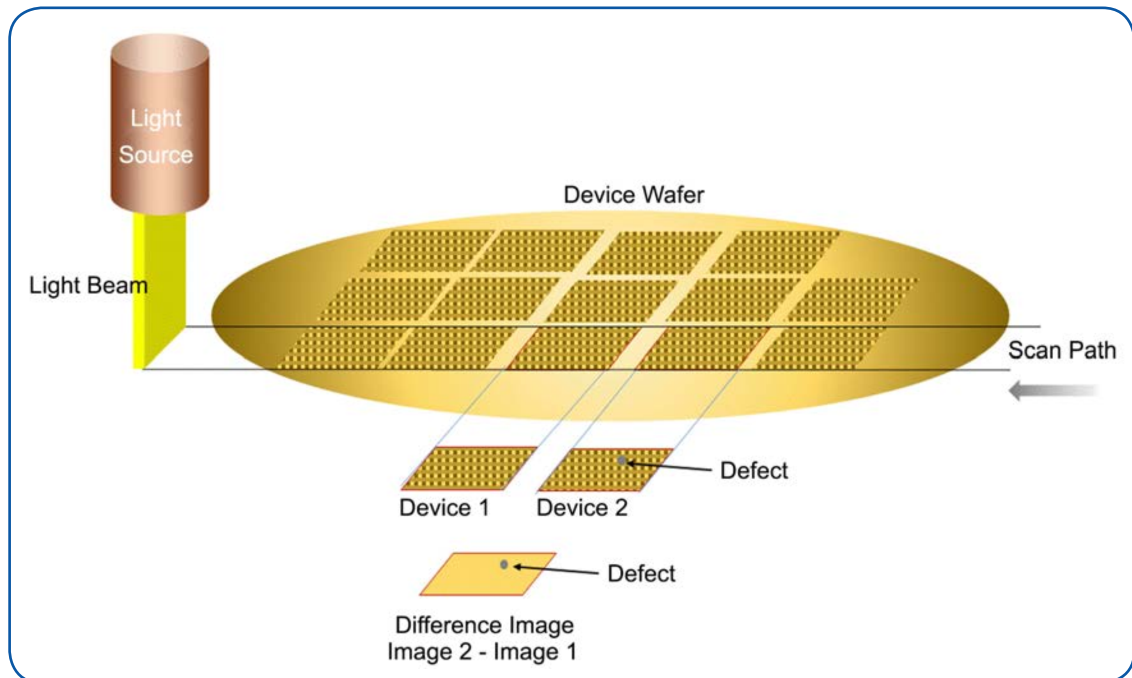


Figure 209. Patterned wafer inspection procedure. [236]).

3. Reticle Inspection

Arguably, reticle inspection is far more important than either non-patterned or patterned wafer inspection. This is due the fact that, while a single defect on a bare or patterned wafer has the potential to "kill" one device, a single defect on a reticle can destroy thousands of devices (or more!) since that defect is replicated on every wafer processed using that particular reticle.

Reticle inspection systems work on the same basic principles and have similar physical requirements to wafer inspection tools, with the exception that reticles are typically inspected using transmitted, rather than, reflected light. Reticle inspection tools employ high resolution imaging optics and either conventional visual or UV wavelength illumination, depending on defect tolerances and/or feature sizes, to find defects on a reticle blank or a patterned reticle. Inspections are routinely performed during both the reticle manufacturing process and throughout the reticle use phase. Reticle inspection tools employ image analysis software algorithms and motion control systems similar to wafer inspection tools. Using UV illumination, the application of conventional optical reticle inspection systems has been extended down to the 90-nm feature sizes; for smaller features, electron beam, EB, reticle inspection is used.

B. Advanced Wafer and Reticle Inspection Tools

1. Sub-100 nm Unpatterned Wafer Inspection

Sub-100 nm inspection tools are currently used in manufacturing environments to provide quality assurance on incoming wafers and for process tool monitoring applications such as system qualification for high volume manufacturing. These tools employ the same basic operating principles as tools designed for larger scale defect detection but use deep UV (DUV) illumination, enhanced optical systems, and sophisticated image analysis algorithms to achieve sub-20 nm sensitivity. A high degree of precision and accuracy is required in the motion control of the wafer stage and the system's optical components at this scale.

mks

Sub-100 nm inspection for unpatterned wafers is complicated by issues of scale. Signal to noise (S/N) ratios are a critical parameter in the detection of particles and other defects on wafer surfaces. As inspection tools are required to sense and quantify ever smaller particles, the impact of factors such as surface microroughness (haze) begins to influence the detectability of small particles via reduction in the S/N ratio of the scattered light signal. Surface chemical contamination from sources such as ambient humidity can also contribute to reductions in S/N. To help counter such effects, inspection tools for sub-100 nm defect detection employ more sophisticated optical spatial filtering and polarization of the scattered signal and specialized signal processing algorithms to optimize their ability to detect small defects in the presence of surface haze.

2. Sub-100 nm Patterned Wafer Inspection

Defect inspection for sub-100 nm patterned wafers presents much greater challenges than does unpatterned wafer inspection. Deep ultraviolet-based (DUV-based) optical inspection for patterned wafer applications operate using the same image comparison principle as older visible light and UV inspection system, albeit requiring a greater degree of sophistication in the optics, motion control and image analysis algorithm.

DUV tools have become the industry standard for patterned wafer inspection tools in applications down to the 65 nm feature sizes (and below). Patterned wafer inspection rates of up to several wafers per hour make these systems suitable for production applications. DUV inspection tools have demonstrated high sensitivities for the detection of defects such as shallow trench isolation, STI, voids, contact etch defects and photoresist micro bridging at sub-100 nm geometries [237]. Using broad-band DUV/UV/Visible illumination, KLA-Tencor's 2810 Bright-Field Patterned Wafer Inspector currently claims the required sensitivity for defect inspection of all layers on DRAM and flash devices down to 55 nm feature sizes and below.

While their well-understood characteristics and relatively low cost and high throughput make the continued use of DUV optical inspection systems desirable, some manufacturers have reported that DUV inspection systems begin to run out of steam at geometries below 65 nm [238]. Tran [238] claimed that the limiting defect sensitivity for DUV dark-field optical pattern inspection systems is about 75 nm in SRAM technology and much larger on logic areas. DUV bright-field systems have somewhat better limiting sensitivities, ~50 nm in SRAM and, as with dark-field, larger in logic. In addition, the use of DUV lasers to illuminate very small and consequently fragile structures on patterned wafers has encountered some unusual problems such as laser ablation of surface material. The solution to these problems may lie in the use of either broadband illumination for optical inspection systems (existing DUV systems employ 266 nm wavelength and are moving to 193-nm illumination) or in the use of production-capable electron beam inspection tools. Recently introduced production-worthy inspection tools based on plasma-generated broadband illumination (3900-series) claim sub-10 nm resolution. New EB tools are available for defect inspection applications at the 10 nm and lower nodes.

3. Sub-100 nm Reticle Inspection

As with the wafer inspection tools, reticle inspection tools for sub-100 nm applications (both blank and patterned reticle inspection) employ DUV illumination, typically single wavelength 257 nm or 193 nm sources. Figure 210 shows a block diagram of a reticle inspection platform. Reticle inspection systems can be configured to employ either transmitted light through the reticle or reflected light from the reticle surface in the inspection process. As with other inspection systems, this reticle inspector requires highly accurate and precise motion control for the optical components and the air bearing reticle stage.

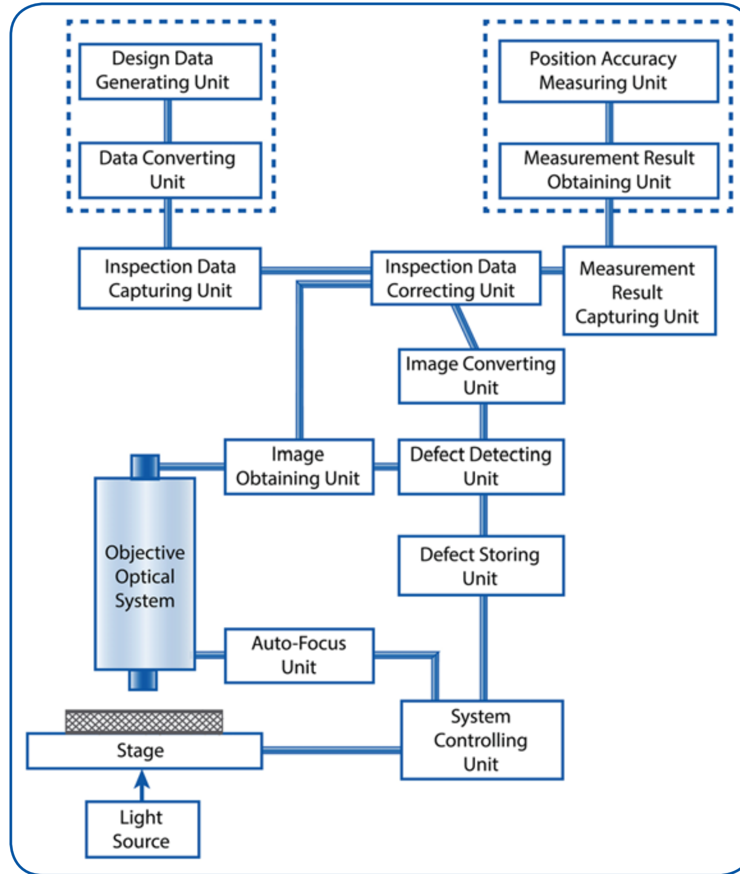


Figure 210. Block diagram of the components in a reticle inspection system (Fujitsu Ltd. [238]).

MKS Product Applications in Wafer and Reticle Inspection

Motion Control for Wafer Inspection

MKS Light and Motion Division offers a variety of high performance air bearing stages suitable for use in wafer inspection tools and other motion control applications. The HybrYX™ XY Hybrid Air Bearing XY Stage (Figure 211) is a relatively low cost, single plane air bearing stage that is well suited for semiconductor wafer inspection systems and many other scanning applications requiring ultra-low velocity ripple & dynamic following error. The HybrYX stage has a true single place XY architecture with available theta and Z-Tip-Tilt solutions. It features scanning velocities of up to 600mm/sec and 0.6G acceleration and large (>1 meter) XY Travel range The HybrYX system has a demonstrated and reliable, long-life operation ideal for high duty cycle environments such as wafer inspection applications.

Figure 211. HybrYX™ XY Hybrid Air Bearing XY Stage.

The DynamYX® family of stages (Figure 212) was specifically designed for semiconductor wafer processing and inspection applications. They offer the highest level of commercially available positioning performance. The extensive use of ceramic materials in these stages provides exceptionally rigid structural stability. They are designed with a low profile that aids in OEM applications.

Figure 212. DynamYX® Datum Ultra-High Performance Stage.

Motion Control for Reticle Inspection

MKS Light and Motion Division offers the DynamYX Reticle Positioning Air Bearing Stage (Figure 213) for use in reticle inspection and repair applications. The DynamYX stage has a much smaller footprint than traditional open-frame solutions and a full-open-aperture that accommodates flexible optical component integration and ease of service access.

Figure 213. DynamYX® RS.

Custom Optical Solutions

MKS Light and Motion Division's Integrated Solutions Business (ISB) group specializes in optical sub-system design and manufacturing and has designed and manufactured assemblies for wafer and reticle inspection tool manufacturers. ISB has designed and manufactured optical sub-systems for lithography, wafer inspection, excimer and EUV light source, metrology and mask writing applications, among others. Table 19 shows the breadth of this experience. ISB has produced custom optical assemblies for applications ranging from illuminators for inspection tools to wafer inspection, as shown in Figures 214, 215 and 216.

Figure 214. 157 nm Objective Lens for VUV 157 nm Photolithography.

Figure 215. High power LED illuminator for use in inspection tools.

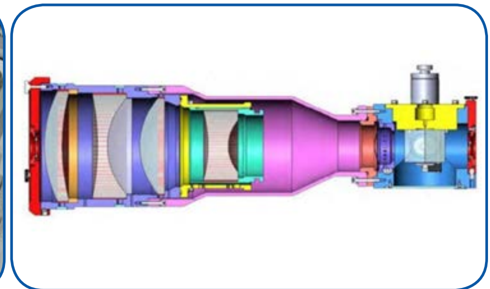


Figure 216. Custom optics.

Opto-mechanical Subsystems

- Wafer alignment
- Wafer reflectometry
- Wafer illumination
- Wafer interferometric inspection
- Brightfield & darkfield optical inspection

Subsystems or Systems

- Photovoltaic scribing & laser edge deletion
- Thermal annealing for LCD flat panel lines

Laser Technology

- Laser beams from DUV to FIR
- Laser beam stabilization
- Laser beam metrology

Precision Optics & Hardware

- Custom optics with wavefront errors $< \lambda/20$
- Custom optical coatings
- Extensive catalog items & custom hardware

Table 19. MKS Light and Motion Division’s design and manufacturing solutions.

This page intentionally left blank.

X. Automation of Process Equipment

The production of microelectronic integrated circuits is, arguably, one of the most complex manufacturing environments in the industrial world. Historically, unit operations such as the fabrication processes (clean, lithography, deposition, etch, etc.), metrology, material handling and others were performed manually. Over the course of decades, however, these unit operations have been converted to automated control owing, primarily, to the growing need to remove human beings from the fab environment. As the feature sizes shrank, device tolerance to particulate and other kinds of contamination became ever lower, until today there is essentially zero tolerance to contamination on wafers, masks and devices in progress. Process technology has evolved to the point where low to zero levels of contamination have become the norm within the process and equipment environments within different unit operations Today, the primary source of particles and other contaminants in the fab ambient is the human staff still needed within the cleanroom.

Over the years, much of the improvement in contamination control has come about through the increased use of automation within the fab. In older 6- and 8-inch fabs, individual process tools were typically automated and computer controlled but there was no automation and coordination of functions such as material handling, product/process characterization, and data collection and analysis between different unit operations; in contrast, modern 12-inch fabs have fully automated wafer processing, metrology, inspection, material handling, and device packaging with little or no human intervention between the introduction of bare substrates to the output of completed devices. Advanced device manufacturers require complex automation solutions for individual unit operations. Subsystems within a unit operation must be seamlessly integrated with intelligent, automated material handling; process control; data collection and analysis; and safety.

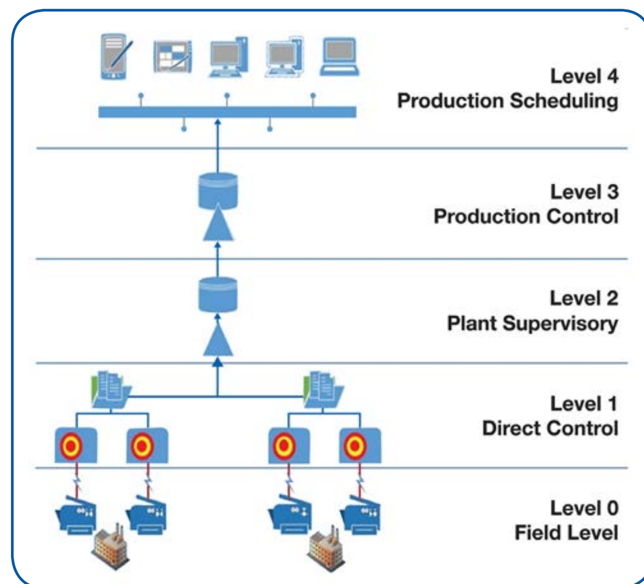


Figure 217. Functional levels of a distributed control system [240].

Automation takes the form of a hierarchical structure as shown in Figure 217, commonly referred to as a distributed control system. Each level of control is connected to the higher levels (and to other computers within the same level) by a high-speed communications network. In Level 0 (Field Level), microprocessor-based controllers execute control loops, perform logic functions and collect and analyze process and machine state data for specific applications within the global manufacturing process. Operational and product data collected at the Field Level are transferred to the unit supervisory systems in Level 2, Plant

mks

Supervisory, where they are used to adjust control loop parameters, analyze process health and track product quality. In Level 3, Production Control, process and product quality data for unit operations are passed to production control systems that coordinate the plant's overall operational activities (i.e. inventory, material handling between unit operations, product quality data, etc) and track historical operational and product characteristics. These latter data are passed along to Level 4, Production Scheduling, where the information is used to optimize asset utilization within the plant. The basic functions that are automated in a semiconductor fab can be considered as sequences or collections of the following activities (taken from [241]):

- Lot selection (or dispatching), to determine which lot to process next
- Transport, to locate and move the lot
- Setting of process condition and recipe, to setup processing conditions
- Process start, to initiate processing
- Process data collection, to record and report measurement data during processing
- Go/no-go quality gating, to determine the acceptance of the processing results
- Exception handling, to handle and solve production exceptions
- Alarm handling, to handle and react predefined alarms

Typical parameters that must be considered in any semiconductor manufacturing automation scheme include:

- Message sequencing standards between different tools and the host computer
- Load/unload port design
- Materials handling
- Wafer cassette/pod identification
- Recipe ID and recipe body check
- Process control
- Engineering review and control
- Manual override

Automation within a fab can be categorized as lying within one of three areas:

1. *Tool automation*, which deals with the automation of the different subsystems within a single unit operations such as a deposition or etch process;

2. *Cell automation*, which deals with managing the connections between tools and tool stations, material movement within and between tools, and advanced process control;

3. *Global fab automation*, which deals with system integration, the scheduling and processing of product lots, business management and fab maintenance.

It is well beyond the scope of this overview to provide an in-depth discussion of global fab automation. Instead, we will provide a brief description of the use of intelligent automation systems for control of the sub-processes in tool automation for a unit operation such as a thin film deposition (Level 1, Direct Control) and on data capture and analysis for real-time monitoring of faults within a manufacturing cell (Level 2, Plant Supervisory), both areas in which MKS has a strong product presence. This discussion will closely follow the available Application Notes and technical magazine publications that focus on these areas, available from the MKS Automation and Control Product Group [242] [243]. Individuals interested in a greater understanding of global automation issues within modern semiconductor fabs (such as Automated Materials Handling Systems (AMHS) and other Computer-Integrated Manufacturing (CIM) systems) can find

useful discussions in publicly available videos that show the fully automated fab functions [244] [245]. There are books and in-depth reviews of automation in semiconductor manufacturing available [246] [247] [248].

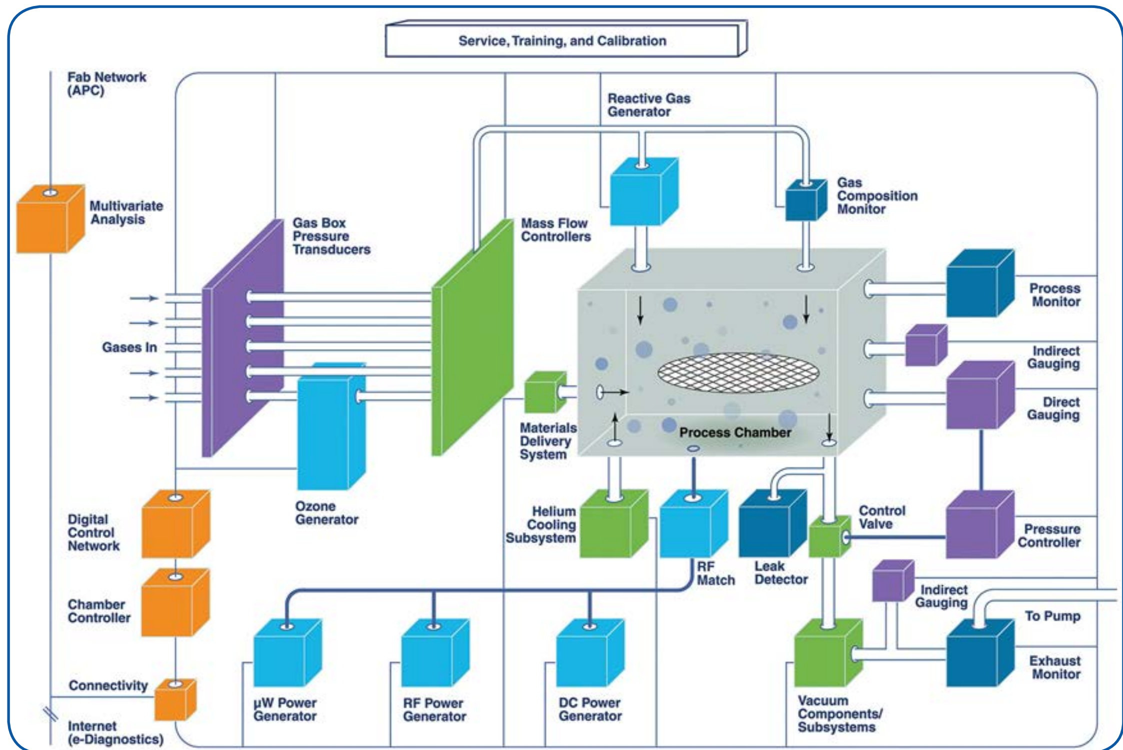


Figure 218. A generic semiconductor wafer processing tool showing different subsystems required for operation.

To understand the issues associated with Level 1 intelligent automation for unit operations, consider the MKS generic semiconductor wafer processing chamber shown in Figure 218. This tool incorporates independent subsystems for, among other things, vacuum and pressure control, gas flow control, process chamber and vacuum line temperature control, plasma power supply, automated wafer handling, process data collection, communication and analysis, and safety. A close examination of this single tool shows that it incorporates well over a dozen independent subsystems. Safe and effective control of units such as this, which contain networks with multiple semi-autonomous subsystems, can be extremely challenging. Care must be taken to avoid control problems due to latency (the delay between the issuance of a command by the control system and the machine response), synchronization, and operational timing challenges at both the local subsystem/device level and at the centralized chamber control level. Adding to the complexity of controlling the many subsystems in a single process chamber is the fact that contemporary semiconductor unit operations are almost invariably configured as cluster tools. Cluster tools incorporate a number of process chambers, usually performing different wafer fabrication processes with concomitant increases in complexity due to the presence of non-redundant systems on different chambers. Furthermore, wafer handling, occasional in situ metrology and the scheduling of internal wafer I/O and processing movements adds even more complexity to the control of these tools. The command information and sensor data needed for synchronized control of the many subsystems in the unit operation must be transferred between master and subsystem controllers in real-time or near real-time. In addition, real-time sensor data is needed by safety interlock systems, should a system failure occur. Finally, with the modular architecture of most advanced process tools, all of this control and communication must be duplicated for multiple process modules.

Process and equipment control in older, conventional semiconductor unit operations typically uses reactive paradigms in which the controller responds to detected changes in process variables with changes to the process control parameters that are intended to bring the process variables back within a specified range. The more advanced process/equipment control employed in modern 8- and 12-inch process tools uses intelligent control paradigms for proactive control of the multiple local nodes based on continually updated, predictive models of the equipment/process functions. Proactive control philosophies combine early fault detection and classification with improved process understanding, and integrated control and automation to achieve higher levels of efficiency and safety (and to lower the cost of a tool's operation). Such proactive control paradigms facilitate reductions in the environmental impact of a process, simplify the efforts needed to comply with government regulations, and help to reduce or eliminate the negative economic and branding costs associated with the release of faulty products to the marketplace.

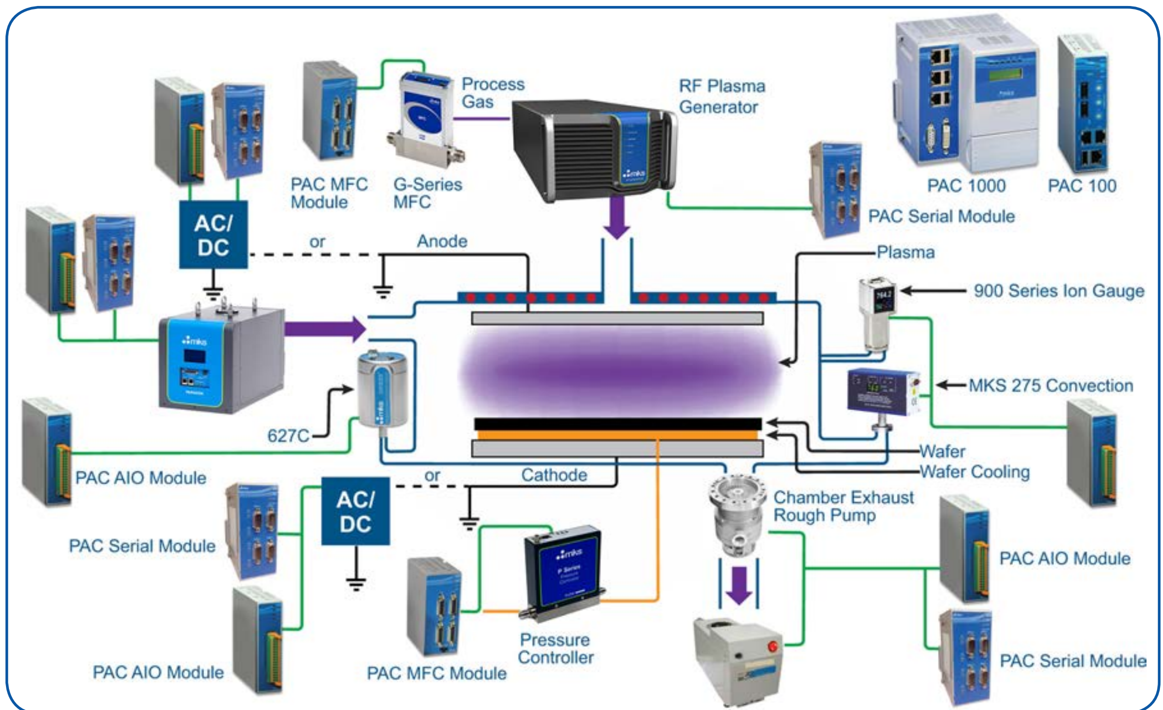


Figure 219. A generic semiconductor wafer processing tool showing different subsystems and the application of MKS's Automation Platform for control of these subsystems.

Network/node-based control can be managed by using a single, centralized controller platform capable of both complete tool and process control. This approach reduces tool complexity through reductions in the number of microprocessor cores and Operating Systems (OS's), employed for control of the various tool functions. Centralized controllers for process tools can also significantly lower the cost of the control architecture as compared with existing approaches.

MKS's Automation Platform addresses the issues discussed above for Level 1 equipment/process control that enables localized and high speed I/O control for MKS and other manufacturers' control elements in a modular fashion. It integrates and automates the sensors and devices resident in a given process tool and communicates equipment and process state data to Level 2 controllers. Figure 219 shows an image of a generic semiconductor wafer processing tool and the different subsystems that can be controlled using the MKS Automation Platform.

In distributed control schemes, the operational and product data collected at Level 1 are transferred to the unit supervisory systems of Level 2. These data carry information that is useful in direct real-time

monitoring and adjusting of the control and health of a process. The data can also be used for updating global empirical control and fault analysis models of overall plant operations. A change in the value of a single independent parameter within a production process has the potential to significantly impact the global process model, forcing an update to that model. Given the very large number of changing independent parameters in most production processes, the need for model updates can quickly become an overwhelming model maintenance issue. Manufacturers thus face a significant challenge in reducing the torrent of machine and sensor data to manageable metrics that summarize the process state and which can be used to maintain global process models. MKS has developed a solution for this problem in their SenseLink™ QM system which enables the collection, storage, retrieval, and analysis of system data, providing data monitoring, visualization, multivariate analysis, modeling and fault detection in a single system for process and product control.

MKS Product Applications in Automation

Automation Platform

Figure 220. Programmable Automation Platform.

The MKS Automation Platform is a modular, scalable and configurable automation and control solution (Figure 220). It integrates seamlessly with other MKS products, improving the utilization of existing tools. The platform consists of two programmable automation control options (PAC 100 & PAC 1000); Communication & Coupler Modules (CMs), a variety of I/O modules interfacing to any type of sensor, actuator, valve, etc.; the Controls Workbench software (CWB) for configuration, process monitoring, tuning, and data storage which supports a standard IEC 61131-3 programming interface. Its modular and open architecture make it scalable and flexible. It supports many fieldbuses and control networks.

SenseLink™ QM Real-Time Quality Monitor

SenseLink™ QM provides process monitoring and part quality prediction through the application of multivariate analysis (MVA) technology from MKS Analytical & Control Solutions (Figure 221). All necessary data acquisition, multivariate processing, and control functions come from one self-contained, compact unit based on the MKS PAC. Adding a SenseLink™ QM monitor to a semiconductor process tool provides in-process fault detection along with contribution details which provide an understanding of your process, not attainable from traditional Supervisory Control and Data Acquisition (SCADA) and Statistical Process Control (SPC) approaches.

Figure 221. SenseLink™ User Interface.

This page intentionally left blank.

XI. Semiconductor Fab Utilities

For many of us, our direct experience with the reality of a semiconductor fabrication facility is limited to meeting rooms, cleanrooms and analytical laboratories. These facilities are, however, part of much more complex structures that contain an intricate infrastructure of ancillary environments housing regulatory, production and delivery systems for the different services and utilities needed for day-to-day fab operation. Figure 222 is a layout diagram for a typical semiconductor fab that shows the ancillary areas needed to provide utility services to the cleanroom production zone. These services include ultrapure water, bulk high purity gases such as nitrogen and argon, exhaust gas handling and disposal, and cleanroom air systems. This section discusses these fab utility systems and supplies context for how different products in MKS's portfolio are employed in these areas. Also, while many of the references in this section are to 200 mm wafer processing, the subject matter is generally applicable to 300 mm wafer processing as well.

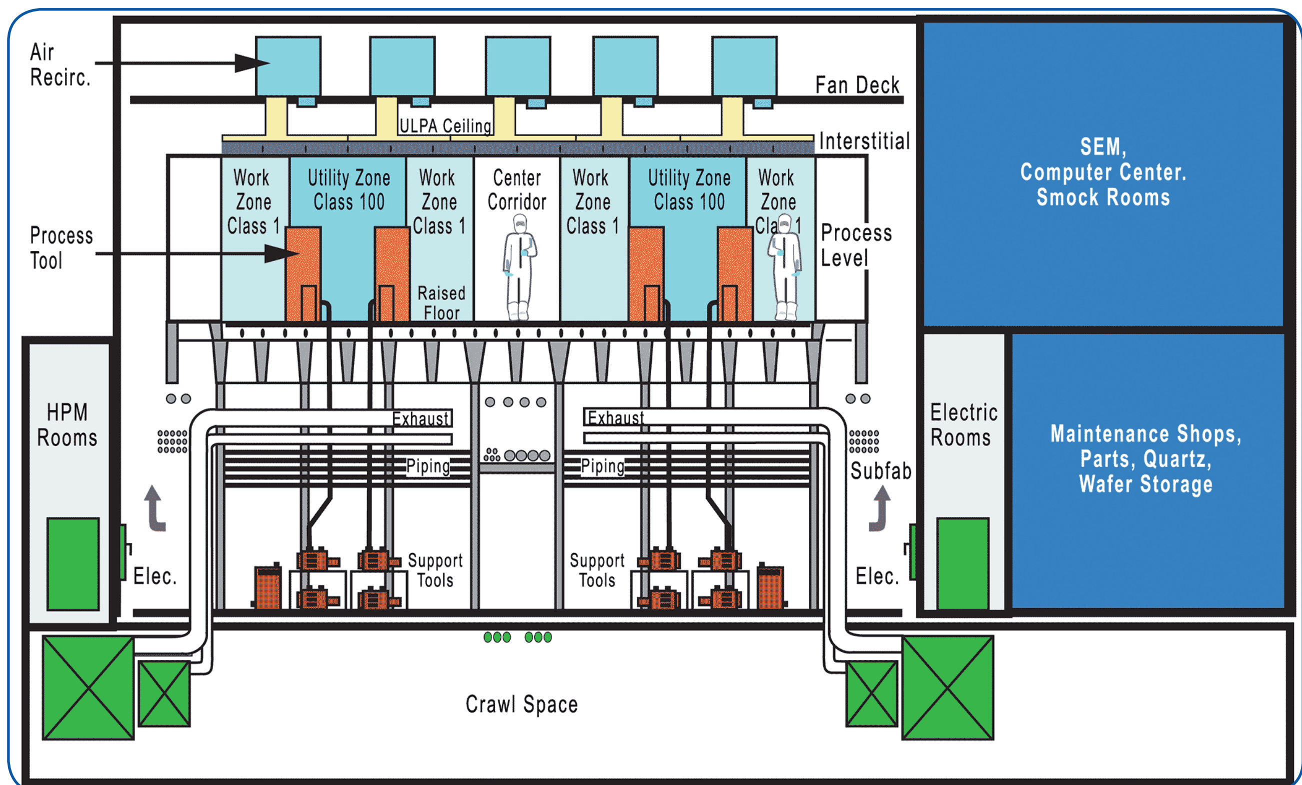


Figure 222. Freescale ATMC factory configuration (from reference [249]).

A. Ultrapure Water

Semiconductor device manufacturing consumes large quantities of water for a variety of purposes ranging from equipment cooling to wafer surface cleaning. Low purity water requirements and conditioning will not be discussed in this overview. Rather, we will focus on the production of ultrapure water (UPW) for use in device fabrication processes. Ultrapure water is required for many process steps. Early stages of device fabrication require repeated steps for wafer cleaning, rinsing and surface conditioning. At many different stages in device manufacturing, it is used for surface cleaning, wet etch, solvent processing, and chemical mechanical polishing/planarization. Indeed, the latter unit process has become one of the largest consumers of UPW within the fab, requiring high volumes for slurry production and rinsing.

Large amounts of UPW are consumed in all fabs – according to the International Technology Roadmap for Semiconductors (ITRS) (2011), device fabs utilized 7 liters/cm^2 of UPW per wafer out. This means that a typical 200 mm wafer fab that processes 20,000 wafers per month can use up to 3,000 m^3 of UPW per day. That is the equivalent of the daily water requirements of a community of 20,000 people. The conversion of raw water to water of ultrahigh purity is thus a significant and costly activity for all semiconductor fabs. Because of the high cost of production and the high-volume needs, there are constant and significant efforts within the industry to reduce the usage of UPW. The ITRS usage target for 2020 was cited as 4.5 liters/cm^2 in the 2015 ITRS Roadmap.

UPW is normally produced using reverse osmosis / deionised resin bed technologies; however, as device linewidths continue to shrink, the requirement for ever higher water purities in semiconductor applications is expected to increase beyond the capabilities of current production technologies. Indeed, modern semiconductor standards for ionic contaminants in UPW are so stringent that some analyses are beyond the detection limits of available analytical tools. This section, will provide the reader a basic familiarity with the design elements and functionalities for UPW systems. We will discuss the main UPW parameters, the treatment sequence for UPW and provide some details of the main treatment steps.

UPW Parameters

Several parameters are monitored for quality control of UPW. These parameters, their points of measurement and measurement method are identified in Table 20. A brief discussion of the main kinds of contaminants, methods of control of their level in UPW, and their typical specified limits is provided below.

Resistivity: This is measured in mega-ohm centimeters or Mohm-cm. Low ion contaminant concentrations in the UPW produce high resistivity values. The theoretical upper limit for UPW with zero ionic contamination is 18.25 Mohm-cm. The 2015 International Technology Roadmap for Semiconductor (ITRS) guideline for UPW resistivity at 25°C is >18.0 Mohm-cm.

Parameter	Measured (POD/POC)	Test Method
TOC	Online	Conductivity/CO_2
Organic Ions	Lab	Ion Chromatography
Other Organics	Lab	LC-MS, GC-MS, LC-OCD
Total Silica	Lab	ICP-MS or GFAAS
Particle Monitoring	Online	Light Scatter
Particle Count	Lab	SEM — capture filter at various pore sizes
Cations, Anions, Metals	Lab	Ion Chromatography, ICP-MS
Dissolved O_2	Online	Electric Cell
Dissolved N_2	Online	Electric Cell

ICP-MS — inductively coupled plasma - mass spectrometry
LC — liquid chromatography; GC — gas chromatography; MS — mass spectrometry
SEM — scanning electron microscopy
GFAAS — graphite furnace atomic absorption spectroscopy
LC-OCD — liquid chromatography - organic carbon detection

Table 20. UPW Parameters, measurement points and methods.

Total Oxidizable Carbon (TOC): The TOC of UPW is measured in parts per billion (ppb). Oxidizable carbon in UPW originates from both inorganic (i.e., mineral carbonates) and organic (including biological and man-made contaminants) carbon contamination in the raw feed water. Typically, reverse osmosis (RO), ion exchange, UV irradiation and degasification are employed to reduce the TOC to acceptable levels in UPW. Tolerable TOC levels in UPW can vary, depending upon the application; however, most applications require very low carbon levels. As an example, the TOC levels needed to avoid lens hazing in immersion

lithography have been a recent driver for this specification, with point of use levels of <1.0 ppb being specified for acceptable performance. The ITRS guideline for TOC is <1.0 ppb.

Dissolved Oxygen (DO): DO is measured using an electrochemical cell. Typically, DO levels in modern fabs are less than 5 ppb. Dissolved oxygen is removed from UPW using vacuum degasification in membrane contactor systems.

Particulate Matter: Raw water sources have high levels of particulate matter. Particles above the micron scale are removed using pre-filters and microfilters, after which the water is polished using increasingly fine filters to remove particles with diameters down to about 0.2 microns. Ultra-filtration at 10,000 molecular weight is used to remove residual particulates beyond this point. Particle specifications for UPW vary, depending on the fab application; in general, particles greater than 0.2 microns cannot be tolerated in any device fabrication, with well-defined limits on particle counts/liter for particles of smaller diameters down to around 0.05 microns. The current ITRS guideline for UPW is <0.3 particles/ml @ 0.05 micron particle diameter. Industry targets are ambitious; suggestions have been made for a specification of the order of <10 particles/ml having diameters greater than 10 nm, a specification that the industry is currently unable to measure, let alone control. In addition to particulate removal in bulk UPW, point of use (POU) ultrafiltration is often employed in the fab environment. Particle counts are normally measured using laser light scattering.

Bacteria: Some bacteria can survive the UPW treatment process and these pose both a biological and particulate threat to integrated devices. Bacterial adhesion occurs naturally in water as pipe walls attract minute quantities of organic nutrients, which attach to the wall and initiate the biofilm process. While regular sanitization programs are employed by some facilities and provide safeguards against microbial activity, biofilms can prove resistant and may permanently coat the inaccessible surfaces of valves and dead-legs. Proper system design and adequate flow velocity are more important than periodic sanitizations to maintaining cleanliness of the system.

Currently, the tests for bacteria and other organisms in UPW employ culture methods that test for viable bacteria and determine the level as "colony forming units/litre" or cfu/liter. These methods lack sensitivity in that only viable bacteria are recovered and large volumes of water may need to be sampled to provide adequate reliability (e.g. <1 cfu/liter cannot be measured with a 100 ml sample). A new technology called Scan RDI may offer a solution for testing total viable organisms. The method is able to detect a single cell based on direct measurements of cell activity and includes bacteria and other live organisms that may be present in biofilm. Another method for bacteria detection is epifluorescence, in which a technician uses a microscope to visually identify both viable and non-viable bacteria that have been stained with dyes that cause biological materials to fluoresce under ultraviolet light. A skilled microscopist can determine much qualitative and quantitative information about the bacteria in the UPW using this method. The ITRS recommended specification for bacterial contamination in UPW is <1 per 1000 ml (by culture).

Silica: Silica, normally measured in ppb, is present in the feed water to UPW systems as silicates and polymeric (or colloidal) silica. Gross removal of silica normally occurs in the RO step of water purification with final removal of residual silica accomplished using anion exchange resin beds followed by ultrafiltration. Typically, the limiting specification for total silica in UPW is 0.2 – 1.0 ppb for dissolved silicates and 0.3 – 2.0 ppb for colloidal silica.

Ions and Metals: Dissolved solids in the feed water to UPW systems consist of a charge-balanced mixture of cations (mostly metals) and anions. These impurities are removed in ion exchange resin beds. Acceptable concentrations of ions and metals in UPW range between 0.02 and 1 ppb, depending on the species and the application.

UPW Unit Operations: Figure 223 provides a schematic of the unit operations in a typical UPW system.

$Al_2(SO_4)_3$ or $Fe_2(SO_4)_2$
Cationic polymer
Lime
NH_3
Cl_2
City
Plant
Acid
Feed water
Raw water
Clarifier
Sand
Filtration
Multi media
Filtration
Scale inhibitor
RO storage
Storage
RO
Reverse osmosis
RO
Reverse osmosis
≤ 5μ
Filtration
Heat
Heat exchanger
Degas
Vacuum degasification
UV
TOC reduction UV irradiation
MB
Primary ion exchange mixed beds
UV
Bacterial inactivation UV irradiation
≤ 1μ
Filtration
N_2
UV
Bacterial inactivation UV irradiation
≤ 1μ
Filtration
MB
Polishing ion exchange mixed beds
UV
TOC reduction UV irradiation
DI storage
≤ 0.02μ
Final filtration
Loop
Ozone
Manufacturing

Figure 223. UPW system process flow abstracted from Sematech guidelines (reproduced from "The Use and Management of Chemical Fluids in Microelectronics", Gottschalk et al, in "Chemistry in Microelectronics", John Wiley & Sons, 2013) [250].

MKS Product Applications in UPW Facilities

MKS currently has no well-established product applications in UPW generating facilities within semiconductor device fabrication facilities. The LIQUOZON® Ultra ozonated DI water system (Figure 224), however, has proven very successful in the sanitization of pharmaceutical UPW systems.

Ozone injection has been shown to be very effective in controlling TOC in these pharmaceutical applications. It may be that, where localized UPW loops of appropriate scale are in place within a semiconductor fab environment, ozone injection using LIQUOZON Ultra or other MKS ozone generation and delivery products [251] can find future applications.

Figure 224. LIQUOZON® Ultra ozonated DI water system.

B. Gas Storage and Delivery

A variety of gases are employed for different purposes within a semiconductor device fabrication plant. These gases range from the pyrophoric and/or toxic specialty gases required for thin film deposition and doping processes (ammonia, methane, silane, germane, dichlorosilane, silicon tetrachloride, phosphine, diborane, arsine and others) through the reactive and corrosive gases needed in different etch processes (chlorine, fluorine, halocarbons, nitrogen trifluoride, etc.). The so-called “atmospherics” (oxygen, hydrogen, nitrogen, argon and helium) are required for purging process systems to reduce contamination and for a number of ancillary needs in process equipment and wafer handling areas. All of the gases used in a semiconductor device manufacturing plant must be of extremely high purity at their source and this purity must be maintained throughout the gas distribution systems that exist within the plant. Some gases may also require further point of use (POU) purification to ensure ultra-low contaminant levels in critical processes.

Nitrogen and argon gases are probably the gases with the highest volumetric use in a semiconductor fab. They are normally supplied to the fab at "5 9s" purity or 99.999% pure. This purity level is necessary since high volumes of these gases are used for purging process equipment and even very low levels of contamination can produce damaging levels of cumulative contamination during long purge cycles. Other atmospherics (i.e., oxygen, hydrogen, helium) are used at somewhat lower volumes and slightly lower purity levels, which are acceptable (but all such gases must be greater than 99% pure at delivery and on-site storage facilities).

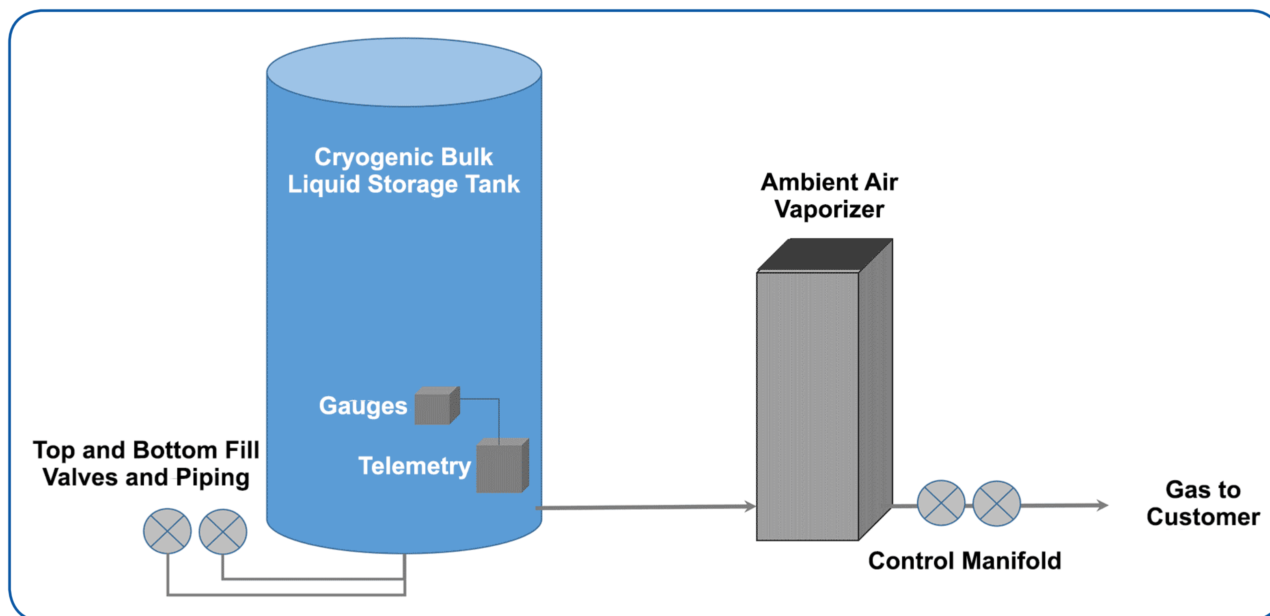


Figure 225. Typical bulk cryogenic gas storage and dispensing system [252].

Within the fab, gases can be stored for distribution in different ways, depending on the purpose and nature of the gas. High volume atmospherics are most often stored in large, cryogenic liquid storage tanks such as those shown in schematic form in Figure 225. These tanks are essentially just large "Dewar" vessels [253] having an inner vessel that contains the cryogenic liquid encased in an outer shell with a high vacuum between the inner vessel and the outer shell. Vacuum is the best possible insulation to maintain the liquid at cryogenic temperatures. In addition to cryogenic storage, tube trailers are often employed for on-site storage of bulk atmospherics.

Hazardous high pressure gases include those with pyrophoric, flammable, corrosive, and toxic characteristics (e.g., silane, most hydrocarbons, fluorides and chlorides, phosphine, arsine, etc). Such gases are typically stored in high pressure cylinders that must be maintained in well-vented gas cabinets that are specially designed to mitigate the hazard associated with a specific gas. Most semiconductor fabrication plants have large numbers of such cabinets on site and there are strict regulations governing the placement and number of hazardous gases in a given facility. (See, for instance, the standards published by the National Fire Prevention Association, specifically NFPA 45, the Standard on Fire Protection for Laboratories using Chemicals). Reference [254] provides some useful information on safe handling and storage of high pressure hazardous gas cylinders. Figure 226 shows a schematic of a typical hazardous gas cabinet designed to contain pyrophoric gases such as silane. These cabinets are equipped with appropriate fittings for connection to the high-pressure cylinders and valve configurations that are designed for cross-purging piping connections to remove residual pockets of gas prior to disconnecting an empty cylinder. Modern gas cabinets are normally equipped with analytical sensors that can detect any leakage from the cylinders or piping.

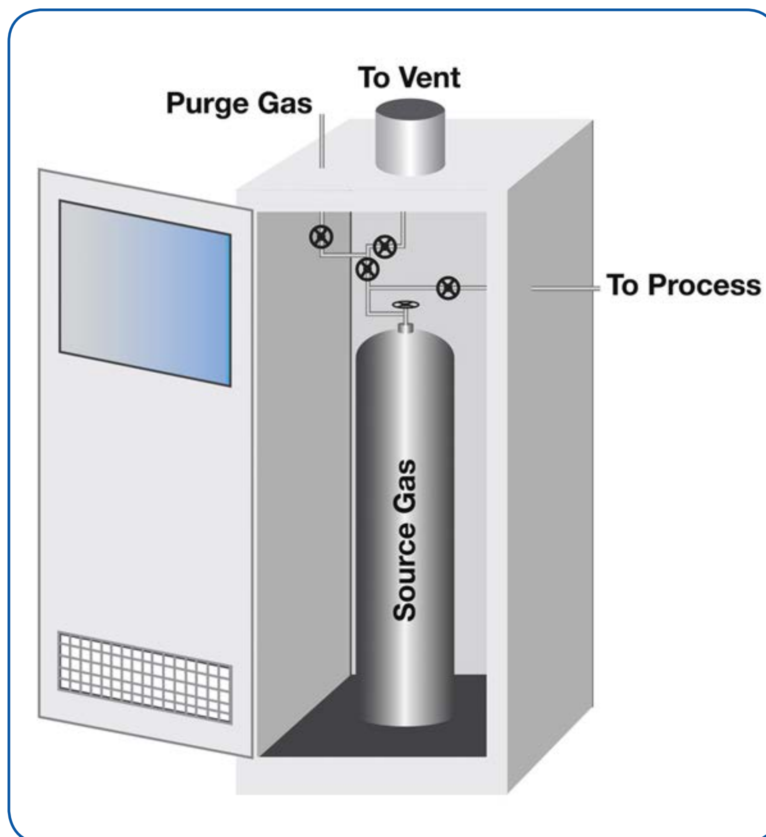


Figure 226. Gas cabinets for pyrophoric gases [254]).

In addition to bulk storage, some high usage gases may be generated on site using gas generators that are supplied, maintained and operated by the gas supplier to a semiconductor plant. On site production of nitrogen, oxygen and hydrogen is becoming relatively common within the industry.

MKS Product Applications in Gas Storage and Delivery

MKS's product line has a number of products that can be used to ensure the quality of bulk gas supplies within a semiconductor fab environment.

MultiGas™ 2032 Purity FTIR Gas Analyzer

The MultiGas Purity Analyzer is an FTIR (Fourier Transform Infrared) spectrometer based gas analyzer that can continuously monitor contaminant levels in bulk process gases at the ppb level (Figure 227). The unit readily detects water, carbon monoxide, carbon dioxide, nitrous oxide, methane, hydrogen fluoride, hydrogen chloride, ammonia, methane, dichloromethane, trichloromethane and many other chemical contaminants in bulk process gases. Bulk gases which can be monitored using the MultiGas 2032 include ammonia, nitrogen trifluoride, silicon tetrafluoride, hydrogen, and others. MultiGas analyzers have a proven utility in bulk gas purity monitoring, including monitoring of gas quality from air separation units, ASUs.

Figure 227. MultiGas™ 2032 FTIR Gas Analyzer.

RGA in situ Process Monitors

The Cirrus 2 benchtop atmospheric pressure gas analysis system can be employed for on-line monitoring and analysis of gases and gas mixtures including trace contaminants in process gases (Figure 228).

Figure 228.
Cirrus™ 2 Atmospheric Gas Analysis System.

Precisive® 5 Application-specific Analyzers

The Precisive 5 application-specific analyzer has been specifically developed for monitoring the quality of hydrocarbon gases in the oil and gas industry, though the operating principles of this technology allow it to be adapted as a gas quality monitor in a variety of industrial settings. This includes those semiconductor process applications in which it is desirable to monitor such bulk gas characteristics as the relative proportions in gas mixtures (Figure 229). The Precisive 5 Analyzer is not suited to the detection of trace levels of contaminants in bulk gases.

Figure 229.
Precisive® 5 Gas Analyzer.

Thermal and Pressure-based Mass Flow Controllers and Meters

MKS mass flow controllers (MFCs) and meters (MFMs) are designed to meet a broad variety of processing applications. Most MKS mass flow control and metering products offer both analog (0-5 VDC; 4-20 mA) and digital (Devicenet, Profibus, EtherCAT, RS485) I/O with embedded Modbus and an Ethernet user interface.

G-Series MFCs and MFMs are broadly applicable instruments that provide cost-effective, high performance measurement and control of gas flows between 5 cc/min (sccm) and 250 L/min (slm) with an accuracy of 1% of setpoint. They have multi-range, multi-gas capability and are available in either elastomer or metal sealed configurations.

ALTA all-digital, metal-sealed MFCs are especially suitable for mass flow control of high flows of hydrogen in pyrogenic oxidation and annealing applications. Hydrogen normally presents a notoriously difficult flow control problem.

I-Series MFCs are also suitable for hydrogen service in pyrogenic oxidation and annealing applications. The I-Series MFCs are IP66-rated for use in harsh industrial environments.

C. Process Exhaust Gas Traps and Scrubbers

The process exhaust downstream from the cleaning, deposition and metal etch processes employed in a semiconductor fab often experiences problems due to fouling. This is most often caused by the condensation and build-up of solid or liquid process by-products on the internal surfaces of vacuum components and exhaust lines. Such build-up can be controlled by maintaining the exposed surfaces at elevated temperatures; therefore fabs require technologies that accurately and precisely control the surface temperatures of the different components in the vacuum train and exhaust gas handling systems. Exhaust streams in a fab frequently contain very corrosive and/or toxic gases that must be removed by chemical scrubbing prior to release to the atmosphere. This entails the use of equipment in most fabs that is dedicated to managing and treating the process exhaust gas stream. Maintenance of this equipment constitutes an important cost factor in all fabs and there is a keen interest in equipment that can reduce or prevent build-up and/or corrosion and prolong maintenance cycles.

CVD processes, for example, convert significantly less than 100% of the process precursor within the processes chamber, and unreacted precursor gases are sent through the vacuum conductance lines to the vacuum pumping and exhaust systems. In many CVD processes, precursor compounds undergo gas phase reactions that produce highly reactive gas-phase intermediate compounds that live long enough to react and produce solids and liquids as they pass through the downstream conductance and exhaust

system. The best understood example of this phenomenon is the pyrolysis of TEOS to produce silicon dioxide. When TEOS is introduced into the process chamber it initially undergoes the following gas-phase decomposition reaction [255]:

$$Si(OC_2H_5)_4 \rightarrow Si(OC_2H_5)_3OH + C_2H_4$$

Since this reaction is endothermic by only 10 kcal/mol, almost all of the TEOS going into the deposition chamber (at ~720°C in low pressure CVD TEOS) is rapidly converted to $Si(OC_2H_5)_3OH$ in the gas phase. The OH group in this molecule makes the intermediate much more reactive than pure TEOS. Residual gases containing this intermediate are pumped out of the deposition chamber and into the exhaust lines of the system where they continue to react even up to and through the vacuum pumps and into the exhaust gas collection system. These reactions produce loosely adhering solid deposits on the walls and other surfaces in the vacuum/exhaust train. These deposits build up and eventually create flow constrictions that impact the conductance characteristics of the system and, ultimately, the yield of the process. The deposits are also a particle source within the system. The stainless steel or aluminum surfaces within the process system and other system components have thermal expansion coefficients that are very different from those of any solid that coats the surface, resulting in high levels of stress between the surface and the coating. This stress causes the coating to flake off as the surface expands and contracts from temperature and pressure changes, which generates particles that remain loosely adhered to the surface. Additionally, particulates produced by gas phase nucleation can also lodge in the uneven surfaces of the stressed film. These loosely adhering particles can be transported back into the process chamber during gas cycling, causing unwanted and unpredictable particle excursions in the process. Effective traps for particles and condensation products, when coupled with heating of the exhaust lines and other downstream components, have been found to resolve this issue since they keep much of the precursors or byproducts in the vapor phase until they reach the waste treatment unit, and the constant temperature of the components reduces the flaking of any deposited material due to film stress and thermal cycling. Figure 230 shows a CVD installation having a particle trap and vacuum conductance lines and vacuum control components that are encased in conformal heaters to control solid deposition and condensation on the internal surfaces.

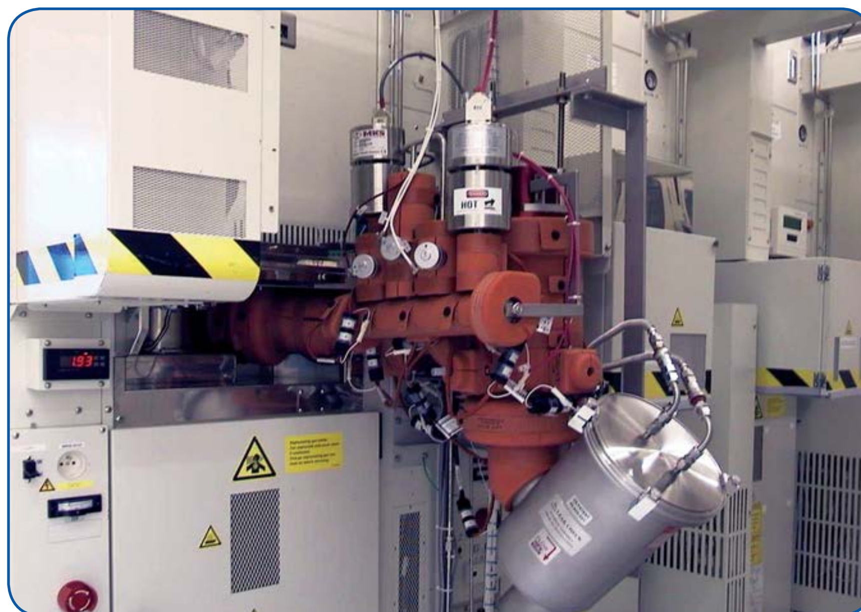
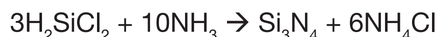


Figure 230. CVD system vacuum lines equipped with MKS S49 thermal management system process line heaters that reduce solid build-up and particle generation.

mks

Silicon nitride CVD provides another example in which reaction by-products must be maintained in the vapor state until they can be effectively trapped or chemically scrubbed from the exhaust stream. The process employs a large excess of ammonia to ensure that the corrosive HCl reaction product is rendered less corrosive through reaction to produce ammonium chloride:

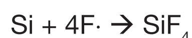


$$3H_2SiCl_2 + 10NH_3 \rightarrow Si_3N_4 + 6NH_4Cl$$

The ammonium chloride, NH_4Cl, is pumped through the downstream vacuum lines and these lines must be heated to a temperature that keeps this abrasive solid in the vapor phase until it can be condensed in an effective particle trap. If allowed to condense in unheated downstream vacuum lines, the ammonium chloride produces a loose, powdery deposit that will backstream into the process chamber during pressure cycling. Heating the vacuum lines to 130 – 150°C in an arrangement such as that shown in Figure 230 keeps the ammonium chloride in the gas phase until it reaches the unheated particle trap.

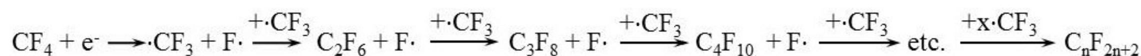
Etch processes have similar issues with the accumulation of deleterious build-up in the conductance and exhaust lines. Source gases for etching are usually fluorine- or chlorine-containing molecules such as CF_4, NF_3, SF_6, CCl_4, Cl_2, BCl_3 or CCl_2F_2. The products of etch reactions can form deposits on downstream surfaces. For example, the plasma etching of silicon using CF_4 proceeds according to:

$$CF_4 + e^- \rightarrow \cdot CF_3 + F\cdot$$



$$Si + 4F\cdot \rightarrow SiF_4$$

Here the dot by a chemical formula denotes a radical fragment. In any etch process that uses CF_4 or other fluoro- or chlorofluoro-carbons, $\cdot CF_3$ and other radical fragments can sequentially react to produce polymers for a CF_4-based etch as illustrated below:



$$CF_4 + e^- \longrightarrow \cdot CF_3 + F\cdot \xrightarrow{+\cdot CF_3} C_2F_6 + F\cdot \xrightarrow{+\cdot CF_3} C_3F_8 + F\cdot \xrightarrow{+\cdot CF_3} C_4F_{10} + F\cdot \xrightarrow{+\cdot CF_3} \text{etc.} \xrightarrow{+x\cdot CF_3} C_nF_{2n+2}$$

For example, Teflon-like polymer build-up, due to this or a similar reaction sequence, is observed in MEMS manufacturing using the Bosch etch process.

Chlorine-based metal etching processes such as aluminum etch can also produce downstream deposits, in this case crystalline aluminum trichloride, $AlCl_3$ [256]. The vapor pressure curve in Figure 231 shows that $AlCl_3$ is a solid at room temperature (25°C) at pressures as low as sub-millitorr. Thus, in an aluminum etch process, with unheated vacuum lines, most of the downstream surface will be coated with aluminum trichloride when the process pressure is in the millitorr range unless the byproduct is trapped close to the source. Heating the vacuum lines to temperatures that maintain $AlCl_3$ in the vapor state at the process pressure resolves this problem. Many other etch processes that use fluorocarbon and inorganic halide gasses produce either polymeric fluorocarbon or solid halide deposits on downstream surfaces.

Downstream of the vacuum pumping system, the process exhaust is fed to a centralized exhaust treatment facility in most semiconductor fabs. These facilities are generally described as exhaust "scrubbers." Typically, fabs employ more than one such unit, differentiated by the type of exhaust stream that the unit must treat. "Acid" gases such as $AlCl_3$ emanate from etch processes and include hydrogen fluoride, hydrogen chloride, chlorine, fluorine, silicon tetrafluoride, Perfluorocarbons (PFCs), nitric and sulphuric acids, as well as with other acidic compounds. Wet scrubbers are used to react acid gas residues with aqueous neutralizing solutions prior to discharge to atmosphere. "Basic" effluents such as ammonia are normally treated separately from acid gases, both because they react with the acid gases and because they require different scrubbing chemistries. Wet scrubbers are normally used for neutralization of basic gases as well. Furthermore, many CVD processes employ flammable, pyrophoric, and corrosive gases that are best removed or rendered harmless by combustion in a dedicated chamber. Finally, dry, chemisorption scrubbers are commonly employed for the removal of toxic gas residues from the exhaust stream.

In addition to acid and ammonia process effluent abatement, the U.S. Environmental Protection Agency has amended its Green House Gas (GHG) Mandatory Reporting Rule 40 Part 98 to include semiconductor fabs in the requirement to report GHG emissions from their facilities. GHG's are emitted from many semiconductor processes including CVD chamber cleans, etch, and wafer cleaning.

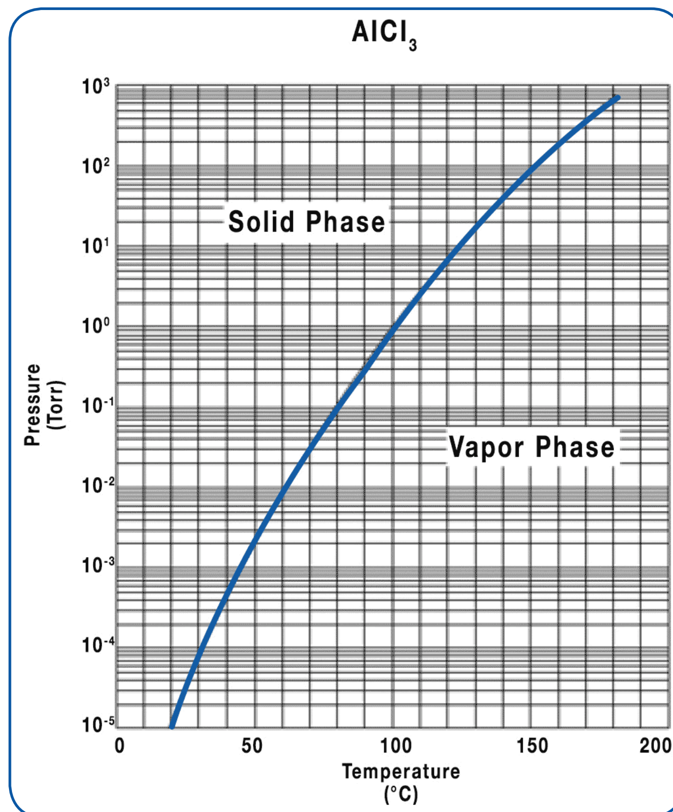


Figure 231. AlCl3 vapor pressure curve.

MKS Product Applications in Exhaust Gas Treatment

Series 49 Thermal Management System

MKS offers a series of advanced heaters and heater control systems. The MKS Series 49 thermal management system (Figure 232) is well-suited for vacuum line heating applications in ALD, CVD, and etch processes, as well as in many other deposition applications. These heaters are conformal, use less energy than conventional heat tracing/insulation installations and have superior control of temperature uniformity. The Series 49 simplifies maintaining accurate and precise temperature control while simultaneously providing improved safety functions and more comprehensive communications options. Applications of the Series 49 are detailed in the MKS white paper titled: "Controlling Vapor Condensation and Downstream Deposits in Deposition and Etch Tools".

Figure 232. Series 49 Thermal Management System.

Heated Vacuum Valves

MKS's *Jalapeño* Series vacuum valves are internally heated valves that prevent condensation of reactive vapors and volatile solids. These include regular and low-profile (Figure 233) heated bellows isolation valves. MKS also offers manual and pneumatic heated vacuum ball valves.

Figure 233.
Jalapeño LoPro™
Isolation Valve.

Process Traps

MKS offers a variety of traps designed for specific process applications. These range from particle and condensation traps for conventional deposition and etch processes to heated traps appropriate for ALD applications.

The Metal Etch Trap (Figure 234) is an efficient, high capacity trap specifically designed to capture the condensable chloride vapors generated in aluminum etch processes. This trap is especially useful in preventing condensable, corrosive vapors such as $AlCl_3$ from entering dry chemisorption scrubbers where they can rapidly clog the unit and destroy the scrubbing capacity. It is specifically designed to capture large quantities of condensable chlorides without reducing the vacuum pumping speed in the etch system.

MKS also supplies the TEOS Trap (Figure 235) as one element of a TEOS Effluent Management Subsystem™. The TEOS trap collects TEOS byproducts, preventing them from backstreaming into the process chamber and/or causing pump oil contamination and abrasive damage of internal pump components. The use of TEOS traps have been demonstrated to produce a reduction in particulates of greater than 20% in process. System maintenance is simplified with a TEOS trap in place, replacing the cleaning of many feet of piping with cleaning a single component. The large trapping capacity leads to longer preventative maintenance cycles. High trapping efficiency provides better protection to the pump, valves, and other downstream instrumentation.

CVD silicon nitride processes probably have the longest history of maintenance issues arising from pump wear due to abrasion by ammonium chloride particles in the exhaust stream. Newer material deposition processes such as CVD titanium nitride exhibit similar issues that require preventive action. MKS offers the Vapor Sublimation Trap (Figure 236) for use in silicon nitride and titanium nitride processes. In combination with heated lines, valves, and other vacuum components, the vapor sublimation trap effectively eliminates ammonium chloride build up in vacuum forelines of these CVD systems.

Recently, MKS introduced a Heated Trap (Figure 237) that prevents coating and clogging of chamber exhaust lines. This trap is specifically designed for ALD systems employing condensable organometallic precursors. These precursors, while not reactive during their additive cycle in the ALD process, can condense on exhaust lines and react when co-reactant is exhausted from the process chamber, producing nitride or oxide films that are very difficult to remove. The heated trap, in combination with heated vacuum forelines, effectively prevents this problem by thermally reacting excess precursor to produce the metal and organic by-products that can be safely pumped to the exhaust system.

Figure 234.
Metal Etch Trap.

Figure 235.
TEOS Trap.

Figure 236.
Vapor Sublimation Trap.

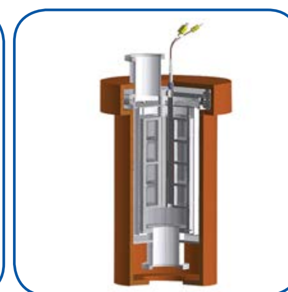


Figure 237.
Heated Trap.

Nitrogen Barriers and Scrubber Inlet Transition Devices

MKS offers a unique Virtual Wall™ nitrogen barrier device (Figure 238) that controls particulate and deposits using a nitrogen gas boundary layer between the inner surface of vacuum forelines and the exhaust gas stream. The Virtual Wall, originally developed as part of the TEOS Effluent Management Subsystem, is finding application in a number of semiconductor applications, including tungsten deposition and metal etch. The MKS Application Note titled: "The Versatile MKS Virtual Wall™ Gas Barrier Device" includes a discussion of the different applications of the Virtual Wall.

MKS also offers custom scrubber inlet transition kits for wet scrubbers that prevent water back-migration into exhaust lines. Kits are customized to process chemistries to keep precursors and byproducts moving into the scrubber system.

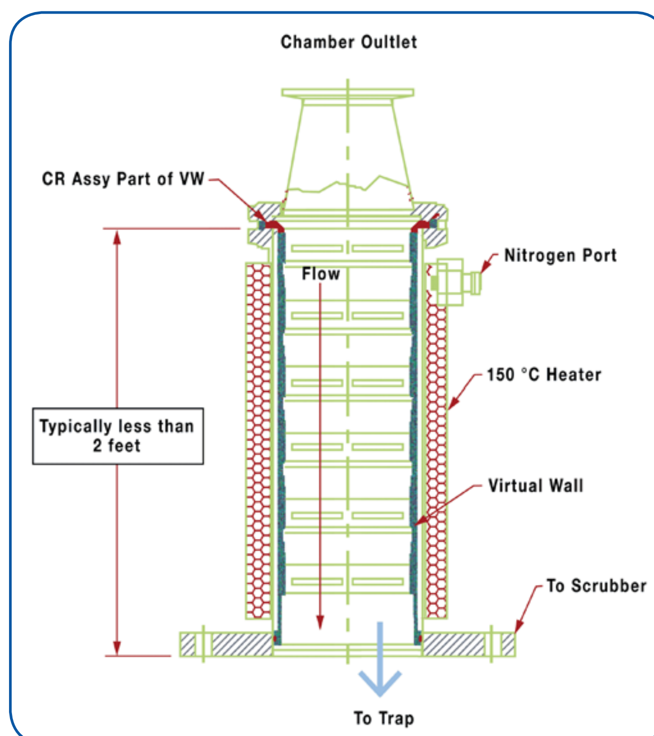


Figure 238. Virtual Wall® gas barrier device.

Exhaust Monitoring

MKS provides a variety of analytical solutions for monitoring semiconductor fab exhaust streams. Data from these tools can be used for both process optimization and regulatory compliance. MKS's analytical tools are based on either optical spectrometry or mass spectrometry.

MKS's MultiGas 2030 continuous gas analyzer (Figure 239) is based on an Fourier Transform Infrared (FTIR) Technology. It provides accurate and precise analyses of gas streams that can be configured for applications in both vacuum forelines and process exhausts. The 2030 has been certified EPA-compliant for measuring the destruction efficiencies of abatement systems for F-GHGs. MKS has publications available that describe the use of the MultiGas™ 2030 for both process optimization (e.g., 1, 2, and 3) and exhaust monitoring for regulatory compliance.

Figure 239. MultiGas™ 2030 Continuous Gas Analyzer.

MKS's MultiGas 2030 continuous gas analyzer and Precisive 5 Trace Hydrocarbon/VOC analyzer (Figure 240) have been proven as effective monitors in processes where monitoring the efficiency of volatile organic compound (VOC) abatement systems is necessary.

The Cirrus 2 benchtop atmospheric pressure gas analysis system has also been proven as an analytical system for exhaust monitoring, especially for greenhouse gas (GHG) emissions.

Figure 240. Precisive® 5 Trace Hydrocarbon/VOC Analyzer.

D. Cleanroom Air Supply System

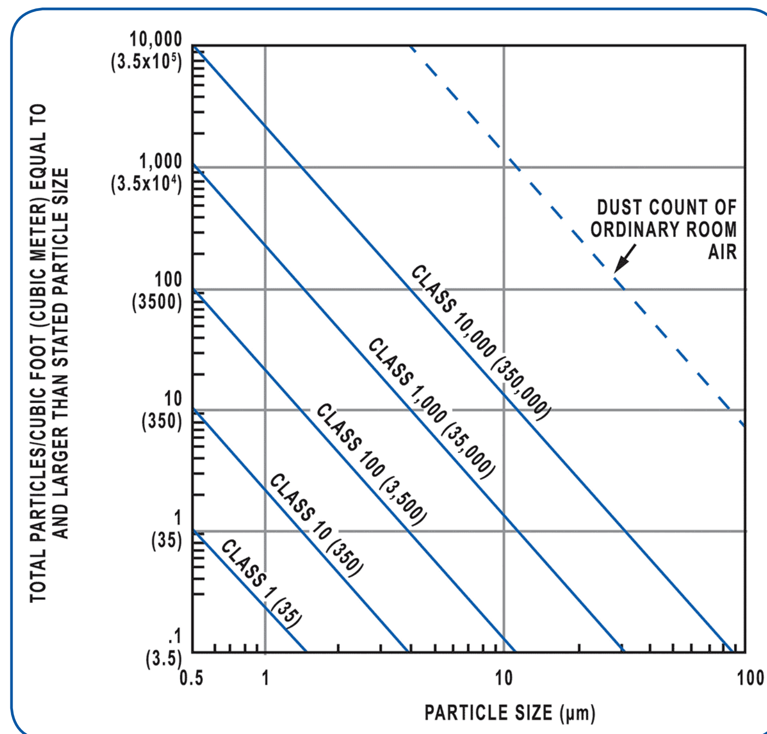


Figure 241. Cleanroom classification [257].

Semiconductor fab environments require ultra-clean conditions to ensure low numbers of product defects due to particulate contamination. A critical aspect in creating an ultra-clean environment is proper design of an air purification and handling system to deliver ultra-low particulate air to the work environment.

Cleanrooms are classified according to the size and number of particles per cubic foot in the ambient cleanroom air. Figure 241 shows a graph detailing the different cleanroom classifications. This graphic is based on an older standard, General Services Administration (GSA) standard FED-STD-209E, which has been recently replaced by a more rigorous ISO standard — ISO 14644, Clean Rooms and Controlled Environments (14644-1: Classification of Air Cleanliness shown in Table 21).

Class	Maximum Particles/m^3						Fed Std 209E Equivalent
	>=0.1 µm	>=0.2 µm	>=0.3 µm	>=0.5 µm	>=1 µm	>=5 µm	
ISO 1	10	2					
ISO 2	100	24	10	4			
ISO 3	1,000	237	102	35	8		Class 1
ISO 4	10,000	2,370	1,020	352	83		Class 10
ISO 5	100,000	23,700	10,200	3,520	832	29	Class 100
ISO 6	1,000,000	237,000	102,000	35,200	8,320	293	Class 1,000
ISO 7				352,000	83,200	2,930	Class 10,000
ISO 8				3,520,000	832,000	29,300	Class 100,000
ISO 9				35,200,000	8,320,000	293,000	Room Air

Table 21. ISO 14644-1 cleanroom standards.

Ultra-clean air is produced using advanced air filtration technology. The filters employed are either highly efficient particulate air (HEPA) or ultra-low particulate air (ULPA) filters that remove roughly 99.9% of the microparticulates in room air. The flow pattern employed for the clean air within a fab is commonly referred to as “laminar flow”; it directs the flow of ultra-low particulate air so as to avoid moving any existing particles to work environments into the work zone. Depending on the fab environment, different modes of laminar air flow may be used. In the larger, general work area, laminar air flow is most often vertical with clean, filtered air introduced at the ceiling level and withdrawn through vents in the sub-floor. Clean, filtered air may also be introduced in a horizontal laminar flow in the general cleanroom environment, although this is not encountered as often as vertical laminar flow. Rather, individual work stations will frequently employ dedicated filtration and horizontal laminar air flow to ensure the lowest levels of particulate contamination in those areas.

Activity	Particles/Minute (0.3 microns and larger)
Motionless (Standing or Seated)	100,000
Walking ~2 mph	5,000,000
Walking ~3.5 mph	7,000,000
Walking ~5 mph	10,000,000
Horseplay	100,000,000

Table 22. Particulates generated by human activity.

When proper cleaning methods and maintenance are coupled with effective air filtration and laminar flow design, it is possible to achieve cleanroom classifications as low as Class 1 or ISO 3 (sometimes better). However, it is important to understand that a cleanroom’s rating refers to the room without equipment or people. Equipment, when properly installed and maintained, does not cause significant deterioration in cleanroom performance. The primary source of degradation in cleanroom performance is, typically, the people that work in that environment. Through the simple act of being present in the cleanroom they contribute to particle contamination by shedding particles from their skin and clothing. Table 22 shows the rate of particle generation that accompanies a variety of simple, unprotected human activities. This is why it is necessary to undergo special gowning and entry procedures when entering a semiconductor work environment. Even so protected, humans in the cleanroom tend to cause a degradation in particle performance and this is the reason why fabs are moving, as much as possible, to fully automated work environments.

MKS Product Applications in Cleanroom Air Supply Systems

MKS offers the AIRGARD® ambient air analyzer to monitor the quality of filtered, recirculated air in cleanroom environments (Figure 242). The extensive use of toxic gases and solvents in semiconductor processing requires such monitoring for the safety of workers. The AIRGARD ambient air analyzer is a self-contained, high sensitivity FTIR-based gas analyzer. FTIR is a proven method for the detection and quantification of a wide range of volatile chemical species in ambient air, with detection limits in the low parts per billion (ppb) range, depending on the species and matrix. The AIRGARD analyzer employs FTIR technology for the rapid detection of a range of toxic chemicals at ppb levels and response times of less than 20 seconds. It has been successfully employed for the detection and monitoring of numerous toxic industrial chemicals as well as for continuous monitoring for chemical warfare agents. The AIRGARD analyzer was SAFETY Act designated by the Department of Homeland Security (DHS) in June, 2011.The method is proven to deliver no false positives (and hence no costly false alarms in the fab environment) and 97% detection efficiency.

Figure 242.
AIRGARD®
Ambient Air Analyzer.

Bibliography

[1] E. M. Purcell and D. J. Morin, Electricity and Magnetism 3rd Edition, Cambridge University Press, 2013.

[2] D. Owens, “Physics 12.4.1a - Electric Potential and Potential Difference,” [Online]. Available: https://www.youtube.com/watch?v=wT9AsY79f1k.

[3] R. Chang, Essential Chemistry: a core text for general chemistry - 2nd Edition, McGraw-Hill Higher Education, 2000.

[4] “General Chemistry,” [Online]. Available: https://www.youtube.com/watch?v=FSyAehMdpyI&list=PL8dPuuaLjXtPHzzYuWy6fYEaX9mQQ8oGr.

[5] “Energy and Chemistry - Change, Chapter 6,” [Online]. Available: https://www.youtube.com/watch?v=GqtUWyDR1fg .

[6] “Electronic structure of atoms,” [Online]. Available: https://www.youtube.com/watch?v=rcKilE9CdaA.

[7] “Chemical Bonding,” [Online]. Available: https://www.youtube.com/watch?v=a8LF7JEb0IA.

[8] H. T. Grahn, Introduction to Semiconductor Physics, Singapore: World Scientific Co. Pte. Ltd., 1999.

[9] “Intrinsic And Extrinsic Semi Conductors,” [Online]. Available: https://www.youtube.com/watch?v=o-PPbmMm0eA.

[10] J. Walker, “Part 3,” in *Halliday & Resnick Fundamentals of Physics Extended 9th Edition*, John Wiley & Sons, Inc..

[11] “The Electric Field Around a Charge,” [Online]. Available: https://www.youtube.com/watch?v=vaDT4GwAZ2I&list=PL6io3nDcMmqAwIL5TeBx_cvBGpiH1XRjj.

[12] “Intrinsic Silicon and Extrinsic Silicon,” [Online]. Available: https://www.youtube.com/watch?v=uFaZiVy3bZo&index=59&list=PL8TDSRZRF0t9wYCFIEToRXkBqbIXqzC3r.

[13] B. Anderson and R. L. Anderson, Fundamentals of Semiconductor Devices, McGraw-Hill Education, 2005.

[14] “PN Junction,” [Online]. Available: https://www.youtube.com/watch?v=ar7xDMR4P_U.

[15] “Passing Gases: Effusion, Diffusion and the Velocity of a Gas,” [Online]. Available: https://www.youtube.com/watch?v=TLRZAFU_9Kg.

[16] “PN Junction Theory,” [Online]. Available: http://www.electronics-tutorials.ws/diode/diode_2.html.

[17] “How a transistor works,” [Online]. Available: https://www.youtube.com/watch?v=DXvAlwMAxiA.

[18] “Bipolar Transistor,” [Online]. Available: http://www.electronics-tutorials.ws/transistor/tran_1.html.

[19] “How MOSFETs and Field-Effect Transistors Work,” [Online]. Available: https://www.youtube.com/watch?v=tz62t-q_KEc&index=2&list=PLUmzZYg1_YlAGG2tw28fW7wrPFYDlnn6n.

[20] “Self-aligned gate,” [Online]. Available: https://en.wikipedia.org/wiki/Self-aligned_gate.

[21] Wikipedia, “Current-Voltage Characteristics,” Wikipedia, [Online]. Available: https://en.wikipedia.org/wiki/Current%E2%80%93voltage_characteristic.

mks

[22] "Why MOSFET Pinchoff occurs," [Online]. Available: http://electronics.stackexchange.com/questions/77198/why-mosfet-pinchoff-occurs.

[23] D. Bhattacharya and N. K. Jha, "FinFETs: From Devices to Architecture," *Advances in Electronics,* 2014.

[24] P. Pavan, R. Bez, P. Olivio and E. Zanoni, "Flash memory cells-an overview," *Proceedings of the IEEE,* vol. 85, no. 8, pp. 1248-1271.

[25] "CMOS," Wikipedia, [Online]. Available: https://en.wikipedia.org/wiki/CMOS.

[26] https://www.youtube.com/watch?v=1rZyGL1K5QI&t=31s

[27] C. Smith and A. R. Barron, Synthesis and Purification of Bulk Semiconductors, Openstax CNX, 20112.

[28] H. Rossotti, Diverse Atoms: Profiles of the Chemical Elements, Oxford, UK: Oxford University Press, 1999.

[29] L. D. Cunningham, "Silicon," USGS Mineral Resources Program, [Online]. Available: http://minerals.usgs.gov/minerals/pubs/commodity/silicon/760494.pdf.

[30] F. A. Cotton, G. Wilkinson, C. Murillo and M. Bochmann, Advanced Inorganic Chemistry, John Wiley & Son, 1999.

[31] "Quartzite," Encyclopaedia Britannica, [Online]. Available: http://www.britannica.com/science/quartzite.

[32] S. Wolf and R. N. Tauber, Silicon Processing For the VLSI Era, Vol. 1, Sunset Beach, California: Lattice Press, 1986.

[33] A. R. Barron, C. R. Hamilton and C. Smith, "Chemistry of the Main Group Elements," Rice University, 20 August 2010. [Online]. Available: http://cnx.org/contents/9G6Gee4A@25.9:bGvrPT9L@3/Semiconductor-Grade-Silicon.

[34] A. Ciftja, T. A. Engh and M. Tangstad, "Refining and Recycling of Silicon: A Review," 1990.

[35] Wacker Chemistry AG, "Silicon Building Blocks for Microchips and Solar Cells," [Online]. Available: https://www.youtube.com/watch?v=eypAfmrRpB0.

[36] W. C. O'Mara, R. B. Herry and L. P. Hunt, Handbook of Semiconductor Silicon Technology, Park Ridge, NJ: Noyes Publications, 1990.

[37] "EKZ 2700 Crystal Puller," PVA TePla AG 2017, [Online]. Available: http://www.pvatepla.com/en/products/crystal-growing-systems/pva/czochralski-process/typ-ekz-2700.

[38] "Miller Index Notation," Southwest Center for Microsystems Education, 2015. [Online]. Available: https://www.youtube.com/watch?v=iFqRUqwAWvo&index=10&list=PLjRlTrCFouyPIZ32vqp_U93dnV7WSPG7V.

[39] "Zone Refining," TutorVista.com, [Online]. Available: https://www.youtube.com/watch?v=pgX4Xlw0NbM.

[40] C. Honsberg and S. Bowden, "Float Zone Silicon," PVEducation.org, [Online]. Available: http://www.pveducation.org/pvcdrom/manufacturing/float-zone-silicon.

[41] "Lapping and Polishing Basics," South Bay Technology, [Online]. Available: http://www.southbaytech.com/appnotes/54%20Lapping%20&%20Polishing%20Basics.pdf.

[42] "Chemical Mechanical Planarization, CMP Process Fundamentals: Sec 2 - CMP Tools and Process," Cabot Microelectronics, [Online]. Available: https://www.youtube.com/watch?v=2z4lq-Ms_OU.

[43] H. Hayashi, "Development of Compound Semiconductor Devices - In Search of Immense Possibilities," *SEI Technical Review,* no. 72, p. 5, 2011.

[44] "Direct and indirect band gaps," [Online]. Available: https://en.wikipedia.org/wiki/Direct_and_indirect_band_gaps.

[45] W. Koechner and M. Bass, Solid-State Lasers - A Graduate Text, New York: Springer-Verlag, 2003.

[46] R. C. Powell, Physics of Solid-State Laser Materials, New York: AIP Press - Springer-Verlag, 1998.

[47] W. Kern, "The Evolution of Silicon Wafer Cleaning Technology," *Journal of the Electrochemical Society,* vol. 137, p. 1887, 1990.

[48] "Crystal Growth," Alineason Materials/Technology, [Online]. Available: http://www.alineason.com/index.php/en/knowhow/crystal-growth.

[49] X. Nie, "Patents of Methods to Prepare Intermetallic Matrix Composites: A Review," *Recent Patents on Materials Science,* vol. 1, no. 3, pp. 232-240, 2008.

[50] T. Kamins, Polycrystalline Silicon for Integrated Circuits and Displays, New York: Springe-Verlag, 1998.

[51] M. Z. Lee, Y. A. Chang, C. L. Lee and T. F. Lei, "Characteristics of Vertical Thermal/PECVD Polysilicon Oxides formed on the Sidewalls of Polysilicon Films," *Journal of the Electrochemical Society,* vol. 150, no. 1, pp. G28-G32, 2003.

[52] M. M. Mandurah, K. C. Saraswat and T. I. Kamins, "A Model for Conduction in Polycrystalline Silicon Part 1: Theory," *IEEE Transactions on Electron Devices,* vol. 28, no. 10, pp. 1163-1171, 1981.

[53] "Chemical Vapor Deposition," [Online]. Available: https://www.youtube.com/watch?v=k27ukETADD8&list=PLv1hJnaxTVPAw81r07W8bbpB0cbzcOctv.

[54] A. C. Jones and M. L. Hitchman, "Chemical Vapour Deposition - Precursors, Processes and Applications," in *An Overview of Chemical Vapour Deposition*, Cambridge, Royal Society of Chemistry, 2009, p. Chapter 1.

[55] "Physico-chemical properties of substance," [Online]. Available: http://chemister.ru/Database/search-en.php.

[56] "Tetrasilane MSDS," Gelest, [Online]. Available: http://www.gelest.com/wp-content/uploads/product_msds/SIT7880.0-msds.pdf.

[57] "The Basics of Cylinder Handling," [Online]. Available: https://www.youtube.com/watch?v=u5UUAlESvHs.

[58] "Compressed Gas Safety Video," [Online]. Available: https://www.youtube.com/watch?v=9-f5zfMH7QI.

[59] "SiH4 Silane exposed to ai," [Online]. Available: https://www.youtube.com/watch?v=ROtw2PjESRU.

[60] "Silane Safety/Lessons Learned and Accident Prevention," Air Products, [Online]. Available: https://www.bnl.gov/pv/files/pdf/IEEE_May2008_Silane_Tutorial.pdf.

[61] K. Huang, "Silane Systems and Operation," Dow Corning, [Online]. Available: http://www.dowcorning.com/content/publishedlit/Silane_Systems_and_Operation_SC_CorpT.pdf.

[62] "Silane Safety Notes," [Online]. Available: http://www-ehs.ucsd.edu/training/CG_Handouts/SilaneSafetyNotes.pdf.

[63] A. T. Voutsas and M. K. Hatalis, "Structure of As-Deposited LPCVD Silicon Films at Low Deposition Temperatures and Pressures," *Journal of the Electrochemical Society,* vol. 139, no. 9, pp. 2659-2665, 1992.

[64] J. G. M. Mulder, P. Eppenga, M. Hendriks and J. E. Tong, "An Industrial LPCVD Process for In Situ Phosphorus-Doped Polysilicon," *Journal of the Electrochemical Society,* vol. 137, no. 1, pp. 273-279, 1990.

[65] D. Houghton, “SiGe CVD, fundamentals and device applications,” ICPS, 2004. [Online]. Available: https://www.aps.org/units/fiap/meetings/presentations/upload/houghton.pdf.

[66] M. Bohr, “The Invention of Uniaxial Strained Silicon Transistors at Intel,” [Online]. Available: http://download.intel.com/pressroom/kits/advancedtech/pdfs/Mark_Bohr_story_on_strained_silicon.pdf.

[67] M. L. Hammond, Handbook of Semiconductor Thin Film Processes and Technology Second Edition, K. Seshan, Ed., New York: Noyes Publications, 2002.

[68] M. Bohr, “Silicon Technology Leadership for the Mobility Era,” 2008. [Online]. Available: http://www.intel.com/content/dam/www/public/us/en/documents/presentation/silicon-technology-leadership-presentation.pdf.

[69] J. D. Cressler, Ed., SiGe and Si Strained-Layer Epitaxy for Silicon Heterostructure Devices, Boca Raton, Florida: CRC Press, 2010.

[70] G. E. Pikus and G. L. Bir, Symmetry and Strain-Induced Effects in Semiconductors, New York: J. Wiley, 1974.

[71] K. E. A. Ismail, “Electron transport properties in Si/SiGe heterostructures: Measurements and device applications,” *Appl. Phys. Lett.,* vol. 63, p. 660, 1993.

[72] D. K. E. A. Nayak, “Enhancement-mode quantum-well GeSi PMOS,” *IEEE Elect. Dev. Lett.,* vol. 12, p. 154, 1991.

[73] J. D. Cressler, Ed., Silicon Heterostructures Handbook, Boca Raton, Florida: CRC Press, 2005.

[74] M. Gupta, “Stress Generation Techniques in Advance CMOS,” *MR International Journal of Engineering and Technology,* vol. 7, p. 1, 2015.

[75] A. Chaudry, G. Joshi, J. N. Roy and D. N. Singh, “Review of Current Strained Silicon Nanoscale MOSFET Structures,” *Acta Technica Napocensis,* vol. 51, p. 15, 2010.

[76] S. W. Bedell, A. Khakifirooz and D. K. Sadana, “Strain Scaling for CMOS,” *MRS Bulletin,* vol. 39, p. 131, 2014.

[77] C. Schriever, C. Bohley, J. Schilling and R. B. Wehrspohn, “Strained Silicon Photonics,” *Materials,* vol. 5, p. 889, 2012.

[78] M. A. Herman and H. Sitter, Molecular Beam Epitaxy - Fundamentals and Current Status, Berlin Heidelberg: Springer-Verlag, 1996.

[79] A. Mosleh, M. A. Alher, L. C. Cousar, W. Du and S. A. Ghetmiri, “Direct Growth of Ge1-xSnx Films on Si using a cold-wal ultra-high vacuum chemical-vapor-deposition system,” *Frontiers in Materials,* vol. 2, 2015.

[80] D. W. Greve, “UHV/CVD and related growth techniques for Si and other materials,” [Online]. Available: https://users.ece.cmu.edu/~dwg/research/UHVCVD.pdf.

[81] B. S. Meyerson, “Low-temperature silicon epitaxy by ultrahigh vacuum/chemical vapor deposition,” *Applied Physics Letters,* vol. 48, p. 797, 1986.

[82] P. H. Holloway and G. E. McGuire, Eds., Handbook of Compoud Semiconductors - Growth, Processing, Characterization and Devices, New Jersey: Noyes Publications, 1995.

[83] S. Tiwari, in *Compound Semiconductors Device Physics*, San Diego, Academic Press, 1992.

[84] B. Mason, M. Larson, Y. Akulova and S. Kalluri, “InP based photonic integrated circuits for DWDM optical communications,” in *2013 International Conference on Indium Phosphide and Related Materials (IPRM)*, May 2013.

[85] R. Michalzik, "VCSELs: A Research Review," in *VCSELs Fundamentals, Technology and Applications of Vertical-Cavity Surface-Emitting Lasers*, R. Michalzik, Ed., Berlin Heidelberg, Springer-Verlag, 2013, p. 166.

[86] Y. Qian, A. Hohl-AbiChedid, J. Li, F. Ye, F. Yang, A. Rice, X. Chen, S. D. Solimine, R. J. Bhat, M. H. Hu and C.-E. Zah, "High performance AlGaInAs/InP 14xx-nm semiconductor pump lasers for optical amplifications," *SPIE Proceedings, Vol. 4905, Materials and devices for Optical and Wireless Communications,* vol. 37, 2002.

[87] Dow Corning, "Compound semiconductor solutions for advanced power electronics," Dow Corning, [Online]. Available: https://www.dowcorning.com/content/electronics/electronicsmarkets/compound-semiconductor-overview.aspx.

[88] Circuits Today, "Energy Band Structure of Gallium Arsenide (GaAs)," [Online]. Available: http://www.circuitstoday.com/energy-band-structure-of-gallium-arsenide-gaas.

[89] E. O. Odoh and A. S. Njapba, "A Review of Semiconductor Quantum Well Devices," *Advances i Physics Theories and Applications,* vol. 46, 2015.

[90] N. Kondo and Y. Nanishi, "Low-Temperature Surface Cleaning of GaAs by Electron Cyclotron Resonance (ECR) Plasma," *Japanese Journal of Applied Physics (Part 2),* vol. 28, no. No. 1, p. L7, 1989.

[91] M. Rahman and N. Hasan, "Compound Semiconductor Epitaxial Growth Techniques," *International Journal of Thin Films Science and Technology,* vol. 5, p. 45, 2016.

[92] A. C. Jones and M. L. Hitchman, Chemical Vapour Deposition: Precursors, Processes and Applications, Cambridge: Royal Society of Chemistry, 2009.

[93] A. C. Jones and P. O'Brien, CVD of Compound Semiconductors, Weinheim, Germany: VCH, 1997.

[94] k-Space Associates, "MOCVD Higher Qualithy Deposition with Advanced Tools," k-Space Associates, Inc., [Online]. Available: http://www.k-space.com/applications/mocvd/.

[95] Aixtron SE, "How MOCVD works," Aixtron Group, [Online]. Available: http://www.aixtron.com/fileadmin/documents/Brochures/How_MOCVD_works.pdf.

[96] S. M. Sze, VLSI Technology - 2nd Edition, India: McGraw-Hill, 2003.

[97] A. Javey, "Berkeley EECS Microfabrication Technology Lecture: Section 4: Thermal Oxidation," [Online]. Available: http://www-inst.eecs.berkeley.edu/~ee143/fa08/lectures/Section%204%20-%20Thermal%20 Oxidation.pdf.

[98] S. Kuppurao, H. S. Joo and G. Miner, "In situ steam generation: A new rapid thermal oxidation technique," *Solid State Technology,* no. July, pp. 233-239, 2000.

[99] K. O. Ng and D. Vanderbilt, "Structure and Oxidation Kinetics of the Si(100)-SiO2 Interface," *Physical Review B,* vol. 59, p. 10132, 1999.

[100] R. Buczko, S. J. Pennycook and S. T. Pantelides, "Bonding arrangements at the Si-SiO2 and SiC-SiO2 interfaces and a possible origin of their contrasting properties," *Physical Review Letters,* vol. 84, p. 943, 2000.

[101] A. Bongiorno, A. Pasquarello, S. H. M. and C. F. L., "Transition Structure at the Si(100)–SiO2 Interface," *Physics Review Letter,* vol. 90, no. 18, p. 186101, 06 May 2003.

[102] F. Giustino, A. Bongiorno and A. Pasquarello, "Atomistic models of the Si(100) - SiO2 interface: structural, electronic and dielectric properties.," *Journal of Physics: Condensed Matter,* vol. 17, p. S2065, 2005.

mks

[103] F. Giustino, A. Bongiorno and A. Pasquarello, “Atomistic models of the Si(100)–SiO2 interface: structural, electronic and dielectric properties,” *Journal of Physics: Condensed Matter,* vol. 17, no. 21, 2005.

[104] C. Hollauer, “The Deal-Grove Model,” [Online]. Available: http://www.iue.tuwien.ac.at/phd/hollauer/node16.html.

[105] B. E. Deal and A. S. Grove, “General Relationship for the Thermal Oxidation of SIlicon,” *Journal of Applied Physics,* vol. 36, p. 3770, 1965.

[106] L. Lie, R. Razouk and B. Deal, “High Pressure Oxidation of Silicon in Dry Oxygen,” *Journal of The Electrochemical Society,* vol. 129, 1982.

[107] Y. v. d. Meulen, “Kinetics of thermal growth of ultra-thin layers of SiO2 on silicon,” *Journal of the Electrochemical Society,* vol. 119, pp. 530-534, 1972.

[108] E. Irene, “Silicon Oxidation Studies: some aspects of the initial oxidation regime,” *Journal of the Electrochemical Society,* vol. 125, p. 1708, 1978.

[109] T. K. Whidden, P. Thankikasalan, M. J. Rack and D. K. Ferry, “Initial oxidation of silicon (100): A unified chemical model for thin and thick oxide growth rates and interfacial structure,” *Journal of Vacuum Science and Technology B,* vol. 13, p. 1618, 1995.

[110] C. Helms and B. Deal, Eds., The Physics and Chemistry of SiO2 and the Si-SiO2 Interface, New York: Springer Science+Media, 1988.

[111] A. Bongiorno and A. Pasquarello, “Multiscale modeling of oxygen diffusion through the oxide during silicon oxidation,” *Physical Review B,* vol. 70, p. 195312, 2004.

[112] T. Zeng and H. Doumanidis, “A study on growth of ultrathin silicon dioxide films by rapid thermal oxidation,” in *Technical Proceedings of the 2003 Nanotechnology Conference and Trade Show*, 2003.

[113] R. Beresford, “Analysis of the thin-oxide growth kinetic equation,” *Semiconductor Science and Technology,* vol. 18, p. 973, 2003.

[114] R. de Almeida, S. Goncalves, I. Baumvol and F. Stedile, “Dynamics of thermal growth of silicon oxide films on Si,” *Physical Review B,* vol. 61, p. 12992, 2000.

[115] T. Cale, “Conformality and Composition of Films Deposited at Low Pressure,” *MRS Proceedings,* vol. 389, 1995.

[116] K. Jensen and W. Kern, “Thermal Chemical Vapor Deposition,” in *Thin Film Processes II*, J. Vossen and W. Kern, Eds., San Diego, Academic Press, 1991, pp. 283 - 367.

[117] L. Tedder, M. Cappelli and K. Saraswat, “Mechanistic studies of dielectric thin film growth by low pressure chemical vapor deposition: The reaction of tetraethoxysilane with SiO2 surfaces,” *Journal of Vacuum Science and Technology,* vol. B11, p. 720, 1991.

[118] G. Hill, A. Harrus and M. Thoma, “Plasma TEOS as an Intermetal Dielectric in Two Level Metal Technology,” *Solid State Technology,* p. 127, 1 April 1990.

[119] S. Wilson and B. Liang, “Characterization of TEOS oxides in the isolation module of high-speed bipolar and BiCMOS Circuits,” *Thin Solid Films,* vol. 220, p. 59, 1992.

[120] J. Perchard, H. Smith, R. O’Connor, J. Olsen and K. Law, “Characterization of a multiple-step in situ plasma enhanced chemical vapor deposition (PECVD) tetraethylorthosilicate (TEOS) planarization scheme for submicron manufacturing,” in *SPIE Proceedings vol. 1188 Multichamber and Insitu Processing of Electronic Materials*, 1989.

[121] K. Maeda and S. M. Fisher, "CVD TEOS/O3: development history and applications," *Solid State Technology,* June 1993.

[122] L. Zanotti, S. Rojas, G. Doghieri and F. Santarelli, "Process chracterization for LPCVD TEOS-ozone based SiO2 films," *Journal de Physique IV,* vol. 3, 1993.

[123] K. Fujino, Y. Nishimoto, N. Tokumasu and K. Maeda, "Silicon Dioxide Deposition by Atmospheric Pressure Low Temperature CVD Using TEOS and Ozone," *Journal of the Electrochemical Society,* vol. 137, p. 2993, 1990.

[124] I. Shareef, G. Rubloff, M. Anderle, W. Gill, J. Cotte and D. K. Kim, "Subatmospheric chemical vapor deposition ozone/TEOS process for SiO2 trench filling," *Journal of Vacuum Science and Technology,* vol. B13, p. 1888, 1995.

[125] T. Whidden and S. Lee, "In Situ Fourier Transform Infrared Spectroscopy near the Substrate in Tetraethoxysilane/Ozone Chemical Vapor Deposition," *Electrochemical and Solid State Letters,* vol. 2, p. 527, 1999.

[126] S. Romet, M. Couturier and T. Whidden, "Modeling of Silicon Dioxide Chemichal Vapor Deposition from Tetraethoxysilane and Ozone," *Journal of the Electrochemical Society,* vol. 148, p. G82, 2001.

[127] C. Beaudry and J. Dukovic, "Faraday in the Fab: A Look at Copper Plating Equipment for On-Chip Wiring," *The Electrochemical Society INTERFACE,* vol. Winter, pp. 40-44, 2004.

[128] B. P. Shieh, L. C. Bassman, D. -K. Kim, K. C. Saraswat, M. D. Deal, J. P. McVittie, R. S. List, s. Nag and L. Ting, "Integration and reliability issues for low capacitance air-gap interconnct structures," in *Proceedings of the IEEE* , 1998.

[129] G. Manera, J. Diniz, I. Doi and J. Swart, "Ultra-Thin Silicon Oxynitride Gate-Dielectric Made by ECR Plasmas," in *Advanced Short-time Thermal Processing for Si-based CMOS Devices II, Electrochemical Society Proceedings*, 2004.

[130] R. Beck and A. Jakubowski, "Ultrathin oxynitride films for CMOS technology," *Journal of Telecommunications and Information Technology,* vol. 1, p. 62, 2004.

[131] K. E. A. Wörhoff, "Design and Application of Compact and Highly Tolerant Polarization-Independent Waveguides," *Journal of Lightwave Technology,* vol. 25, p. 1276, 2007.

[132] C. Woo, C. Ahn, Y. Kwon, J.-H. Han and H. Cho, "Transparent and Flexible Oxide Thin-Film-Transistors Using an Aluminum Oxide Gate Insulator Grown at Low Temperature by Atomic Layer Deposition," *Metals and Materials International,* vol. 18, p. 1055, 2012.

[133] P. Nayak, J. Caraveo-Frescas, Z. Wang, M. Hedhili, Q. Wang and H. Ashareef, "Thin Film Complementary Metal Oxide Semiconductor (CMOS) Device Using a Single-Step Deposition of the Channel Layer," *Scientific Reports,* vol. 4, p. 4672, 2014.

[134] M. Eizenberg, Interlayer Dielectrics for Semiconductor Technologies, S. P. Murarka, M. Eizanberg and A. K. Sinha, Eds., London: Elsevier, 2003, pp. 1-6.

[135] Y. Pauleau, "Deposition Fundamentals and Properties of Metallic and Diffusion Barrier Films," in *Multicomponent and Multilayered Thin Films for Advanced Microtechnologies: Techniques, FUndamentals and Devices*, vol. 234, O. Auciello and J. Engemann, Eds., Kluwer Academic Publishers, 1993, pp. 471-521.

[136] "Metallization," [Online]. Available: http://www2.isu.edu.tw/upload/341/7/files/dept_7_lv_2_31755.pdf.

[137] A. Javey, “Berkeley EECS Microfabrication Technology Lecture: Section 4: Thermal Oxidation,” [Online]. Available: http://www-inst.eecs.berkeley.edu/~ee143/sp08/lectures/Sec8-Metallization.pdf.

[138] S. Zaima and Y. Yasuda, “Low resistivity contact materials for ULSI applications and metal/semiconductor interfaces,” in *ULSI Process Integration II, Proceedings of the International Symposium*, 2001.

[139] C. Tan and A. Roy, “Electromigration in ULSI Interconnects,” *Materials Science and Engineering,* vol. 58, p. 1, 2007.

[140] “Metallization: Aluminum Technology,” [Online]. Available: https://www.halbleiter.org/en/metallization/aluminum-technology/.

[141] D. K. Schroder, “Electromigration,” [Online]. Available: http://schroder.personal.asu.edu/Electromigration.pdf.

[142] J. Bakke, “Extending tungsten metallization for next-generation devices,” *Solid State Technology,* 08 2016.

[143] R. Rosenberg, D. Edelstein, c.-K. Hu and K. Rodbell, “Copper Metallization for High Performance SIlicon Technology,” *Annual Review of Materials Science,* vol. 30, p. 229, 2000.

[144] Q. Jiang, Y. Zhu and M. Zhao, “Copper Metallization for Current Very Large Scale Integration,” *Recent Patents in Nanotechnology,* vol. 5, p. 106, 2011.

[145] T. Ohba, “A Study of Current Multilevel Interconnect Technologies for 90 nm Nodes and Beyond,” *Fujitsu Science Technical Journal,* vol. 38, p. 12, 2002.

[146] “Novellus debuts integrated copper system, signs strategic alliances with Lam, IPEC,” Micromagazine, [Online]. Available: http://micromagazine.fabtech.org/archive/98/07/breakout.html.

[147] Y. Shacham-Diamand and V. Dubin, “Copper electroless deposition technology for ultra-large-scale-integration (USLI) metallization,” *Microelectronic Engineering,* vol. 33, p. 47, 1997.

[148] J. Reid, “Demascene Copper Electroplating,” in *Handbook of Semiconductor Manufacturing Technology 2nd Edition*, R. Doering and Y. Nishi, Eds., Boca Raton, CRC Press, 2008, pp. 16-1 - 16-47.

[149] P. Andricacos, C. Uzoh, J. Dukovic, J. Horkans and H. Deligianni, “Demascene copper electroplating for chip interconnects,” *IBM Journal of Research and Development,* vol. 42, p. 567, 1998.

[150] S. E. A. Venkatesan, “A high performance 1.8V, 0.20/spl μ/m CMOS technology with copper metallization,” in *Electron Devices Meeting, 1997 IEDM’97. Technical Digest, International*, 1997.

[151] Y. Shacham-Diamand, “Barrier Layers for Cu USLI Metallization,” *Journal of Electronic Materials,* vol. 30, p. 336, 2001.

[152] “CMOS Manufacturing Process,” [Online]. Available: http://bwrcs.eecs.berkeley.edu/Classes/icdesign/ee141_s02/Lectures/Lecture5-Manufacturing.pdf.

[153] K. A. Reinhardt and W. Kern, Eds., Handbook of Semiconductor Wafer Cleaning Technology (Second Edition), New York City: William Andrew Inc., 2008.

[154] T. Ohmi, Ed., Scientific Wet Process Technology for Innovative LSI/FPD Manufacturing, Boca Raton, Florida: CRC Press, 2006.

[155] S. R. Kasi, M. Liehr, P. A. Thiry, H. Dallaporta and M. Offenberg, “Hydrocarbon reaction with HF-cleaned Si(100) and effects on metal-oxides semiconductor device quality,” *Appl. Phys. Lett.,* vol. 59, p. 108, 1991.

[156] Y. Shiramizu and H. Kitajima, “Effect of organic compounds on gate oxide reliability,” in *Int. Conf. Solid State Devices and Materials - Extended Abstracts 1995*, Osaka, Japan, 1995.

[157] A. Licciardello, O. Puglisi and S. Pignataro, “Effect of organic contaminants on the oxidation kinetics of silicon at room temperature,” *Appl. Phys. Lett.,* vol. 48, p. 41, 1986.

[158] T. Hattori, Ed., Ultraclean Surface Processing of Silicon WAfers - Secrets of VLSI Manufacturing, New York: Springer, 1998.

[159] K. Saga and T. Hattori, “Influence of surface organic contamination on the incubation time in low-pressure chemical vapor deposition,” *J. Electrochem Soc.,* vol. 144, p. L253, 1997.

[160] J. A. Lebens, W. C. McColgin, J. B. Russell, E. J. Mori and L. W. Shive, “Unintentional doping of wafers due to organophosphates in the clean room ambient,” *J. Electrochem. Soc.,* vol. 143, p. 2906, 1996.

[161] M. Morita, T. Ohmi, E. Hasegawa, M. Kawakami and M. Ohwada, “Growth of native oxide on a silicon surface,” *J. Appl. Phys.,* vol. 68, p. 1272, 1990.

[162] A. Chin, B. C. Lin, W. J. Chen, Y. B. Lin and C. Tsai, “The effect of native oxide on thin gate oxide integrity,” *IEEE Electron Device Letters,* vol. 19, p. 426, 1998.

[163] J. M. Lai, W. H. Chieng, B. C. Lin, A. Chin and C. Tsai, “The Leakage Current Effect of Thin Gate Oxides (2.4-2.7 nm) with In Situ Native Oxide Desorption,” *J. Electrochem. Soc.,* vol. 146, p. 2216, 1998.

[164] J. Yugami, T. Itoga and M. Ohkura, “Highly-reliable gate oxide formation for giga-scale LSI’s by using closed wet cleaning system and wet oxidation with ultra-dry unloading,” *IEDM Tech. Dig.,* pp. 855-858, 1995.

[165] M. Miyawaki, S. Yoshitake and T. Ohmi, “Improvement of aluminum-silicon contact performance in native oxide free processing,” *IEEE Electron Device Letters,* vol. 11, p. 448, 1990.

[166] M. Morita and T. Ohmi, “Characterization and Control of Native Oxide on Silicon,” *Jap. J. Appl. Phys.,* vol. 33, pp. Part 1, Number 1B, 1994.

[167] S. Ciampi, J. B. Harper and J. J. Gooding, “Wet chemical routes to the assembly of organic monolayers on silicon surfaces via the formation of Si-C bonds: surface preparation, passivation and functionalization,” *Chem. Soc. Reviews,* vol. 39, p. 2158, 2010.

[168] S. Franssila, Introduction to Microfabrication, Wiley, 2010, p. 158.

[169] A. Bohg and A. K. Gaind, “Influence of film stress and thermal oxidation on the generation of dislocations in silicon,” *Appl. Phys. Lett.,* vol. 33, p. 895, 1978.

[170] A. Roth, Vacuum Technology, Third, Updated and Enlarged ed., Amsterdam: Elsevier Science B. V., 1990.

[171] N. Yoshimura, Vacuum Technology Practice for Scientific Instruments, Heidelberg: Springer-Verlag, 2008.

[172] J. F. O’Hanlon, A User’s Guide to Vacuum Technology, Hoboken, New Jersey: John Wiley & Sons, 2003.

[173] P. Danielson, “Articles on Vacuum Technology,” Normandale Community College, [Online]. Available: http://www.normandale.edu/departments/stem-and-education/vacuum-and-thin-film-technology/vacuum-lab/articles.

[174] N. S. Harris and L. Budgen, “Design and manufacture of modern mechanical vacuum pumps,” *Vacuum,* vol. 26, no. 12, pp. 525-529, 1976.

[175] “Rotary Vane Pump Animation,” Precision Plus Vacuum, [Online]. Available: https://www.youtube.com/watch?v=AFHogF-9eGA.

[176] “Edwards CXS Dry Screw Vacuum Pump,” Acclon Technologies, [Online]. Available: https://www.youtube.com/watch?v=aSLYNAxolV0.

[177] “Oerlkon Leybold Vacuum - DRYVAC Dry Compressing Screw Vacuum Pump,” Oerlikon Group, [Online]. Available: https://www.youtube.com/watch?v=3iiSfo9h_Fc.

[178] “Pfeiffer Vacuum multi-stage Roots Pump ACP,” Pfeiffer Vacuum, [Online]. Available: https://www.youtube.com/watch?v=rtoXctEvJxc.

[179] “busch rotary claw,” Lapierre Equipment, [Online]. Available: https://www.youtube.com/watch?v=VZvYTRJ6vGM.

[180] J. Hennings, “Thirty years of turbomolecular pumps? A review and recent developments,” *Journal of Vacuum Science and Technology A,* vol. A 6, p. 1196, 1988.

[181] W. Umrath, Ed., Fundamentals of Vacuum Technology, Cologne: Oerlikon Leybold Vacuum, 2007.

[182] M. Vitale, “Blower & Vacuum Best Practices”.*Blower & Vacuum Best Practices.*

[183] Wikipedia, “Turbomolecular Pump,” [Online]. Available: https://en.wikipedia.org/wiki/Turbomolecular_pump.

[184] L. N. Pfeiffer and K. W. West, “Apparatus Comprising a High Vacuum Chamber”. Patent 4,873,833, 19 October 1989.

[185] C. R. Tilford, “Pressure and Vacuum Measurements,” in *Physical Methods of Chemistry* , Second Edition ed., vol. Volume Six: Determination of Thermodynamic Properties, B. W. Rossiter and R. C. Baetzold, Eds., John Wiley & Son, 1992, pp. 101-173.

[186] T. I. O. M. A. Control, Guide to the Measurement of Pressure and Vacuum, London: Institute of Measurement and Control, 1998.

[187] R. D’Angelo and D. Thomas, “Pressure Manometer Basics,” 1 October 2001. [Online]. Available: http://www.sensorsmag.com/sensors/pressure/manometer-basics-1073. [Accessed November 2016].

[188] “Ionization Gauge,” SS Scientific Ltd., [Online]. Available: http://www.ss-sci.com/ssig.htm.

[189] L. Network, “A closer look at gauge types,” Laboratory Network, [Online]. Available: http://www.laboratorynetwork.com/doc/a-closer-look-at-gauge-types-0001.

[190] E. D. Hoffman and V. Stroobant, Mass Spectrometry Principles and Applications, Chichester: John Wiley & Sons Ltd., 2007, pp. 88-100.

[191] SCREEN, “Wet Station FC-3100,” Dainippon Screen Mfg. Co, [Online]. Available: http://www.screen.co.jp/spe/products/fc-3100/index.html.

[192] C. Gottschalk, “Ozonated water - Where the green choice is better,” *Cleanrooms,* 2005.

[193] H. Sundstrom and C. Gottschalk, “The Evolution of Ozone Subsystems,” *Semiconductor International Equipment Components and Substystems,* September 2006.

[194] AKrion, “Process and Equipment Solutions for Integrated Circuit Manufacture,” AKrion Systems LLC, [Online]. Available: http://www.akrionsystems.com/applications/semiconductor.

[195] SCREEN, “SCREEN Launches SU-3300 Single Wafer Cleaning System,” SCREEN Semiconductor Solutions Co., [Online]. Available: http://www.screen.co.jp/eng/spe/information/SPE 16091.html.

[196] C. Gottschalk and J. Schweckendiek, “Using dissolved ozone in semiconductor cleaning applications,” *MICRO Magazine,* March 2004.

[197] T. Hattori, “Implementing a Single-Wafer Cleaning Technology Suitable for Minifab Operations,” *MICRO,* vol. 21, no. 1, pp. 49-57, 2003.

[198] C. V. Ciufia, K. G. Knoch and J. Osterkamp, “Metal Removal and Drying by HF/O3,” in *Semiconductor Wet Processing Conference*, 2004.

[199] I. Kashkoush, G. Chen and R. Novak, “Today’s binary and EAPSMs need advanced mask cleaning methods,” *Solid State Technology,* vol. 47, no. 2, February 2004.

[200] A. Hand, “Getting Masks Clean With Minimal Optical Degradation,” *Semiconductor International,* vol. 26, no. 11, October 2003.

[201] “Ion Implant Technology - 92,” Amsterdam, 1993.

[202] E. Rimini, Ion Implantation: Basics to Device Fabrication, New York: Springer Science + Business Media, LLC, 1995.

[203] J. F. Ziegler, Ed., Ion Implantation Science and Technology, San Diego: Academic Press, 1988.

[204] C. Hu, “Device Fabrication Technology,” in *Modern Semiconductor Devices for Integrated Circuits*, Pearson, 2009, pp. 59-88.

[205] A. Srivastava, A. Wilson and I. Koo, “Using a remote plasma source for n-type Plasma Doping chamber cleans,” in *20th International Conference on Ion Implantation Technology (IIT), 2014*, Portland, OR, 2014.

[206] A. M. Haider, “Kinetics of Silicon Oxidation in a Rapid Thermal Processor,” 2010. [Online]. Available: http://www.ncp.edu.pk/docs/iss/talks/group_ii/g2_d1_asad.pdf.

[207] [Online]. Available: http://www.chem.qmul.ac.uk/surfaces/scc/scat2_5.htm.

[208] B. Cui, “NE 343 Microfabrication and Thin Film Technology,” University of Waterloo, [Online]. Available: https://ece.uwaterloo.ca/~bcui/?page_id=20.

[209] E. Granneman, P. Fischer, D. Pierreux and P. Zagwign, “Batch ALD: Characteristics, comparisons with single wafer ALD and examples,” *Surface and Coatings Technology,* vol. 201, no. 22, pp. 8899-8907, 2007.

[210] G. Dilliway, “ALD for Si Nanoelectronics,” in *ALD for semiconductor appliclatons_Workshop_2010*.

[211] FST Inc., [Online]. Available: http://fstinc.en.ec21.com/Pellicle_Division--389708_389711.html.

[212] K. Suzuki and B. W. Smith, Eds., Microlithography Science and Technology, Boca Raton, FL: CRC Press, 2007.

[213] “Photomasks,” Toppan Printing Co. Ltd., [Online]. Available: http://www.toppan.co.jp/electronics/english/semicon/photo_mask/.

[214] C. Mack, Fundamental Principles of Optical Lithography, Chichester: John Wiley & Son, 2007.

[215] Y. Wei and R. L. Brainard, Advanced Processes for 193-nm Photolithography, Bellingham: SPIE, 2009.

[216] M. Abramowitz and M. W. Davidson, “Optical Microscopy Primer,” National High Magnetic Field Laboratory, 2015. [Online]. Available: http://micro.magnet.fsu.edu/primer/anatomy/numaperture.html.

[217] [Online]. Available: http://www.compugraphics-photomasks.com/our-capabilities/opcphase-shift/.

[218] M. K. Kwak and L. J. Guo, “Phase Shift Lithography,” in *Encyclopedia of Microfluidics and Nanofluidics*, Springer-Verlag, 2014, p. 1.

[219] [Online]. Available: http://spie.org/publications/fg06_p78-80_phase-shift_masks/.

[220] [Online]. Available: http://spie.org/newsroom/immersionlitho-intro.

[221] "The Quad Patterning Era Begins...," Monolith IC3D, 24 10 2011. [Online]. Available: http://www.monolithic3d.com/blog/the-quad-patterning-era-begins.

[222] R.-H. M. Schmidt, "Ultraprecision engineering in lithographic exposure equipment for the semiconductor industry," *Philosophical Transactions of the Royal Society A,* 16 July 2012.

[223] G. Stix, "Water immersion Arrangement," *Scientific American,* vol. 293, pp. 64-67, 2005.

[224] V. Bakshi, Ed., EUV Lithography, Bellingham, Washington: SPIE Press, 2009.

[225] S. Hill, "UV Lithography: Taking extreme measures," Phys.Org, [Online]. Available: https://phys.org/news/2011-12-uv-lithography-extreme.html.

[226] P. French, G. Krijnen and F. Roozeboom, "Precision in Harsh Environments," *Microsystems & Nanoengineering,* vol. 2, no. October, 2016.

[227] D. D. Simone, M. Mao and G. Vandenberghe, "Novel metal-oxide photoresist materials for extreme UV lithography," *SPIE Newsroom Micro/Nano Lithography,* 1 July 2016.

[228] "Etching," Penn State Material Research Institute, [Online]. Available: http://www.mri.psu.edu/nanofabrication-lab/capabilities/etching.

[229] Coburn and Winters, Journal of Vacuum Science and Technology 16, 391 (1979)

[230] AZoNano, [Online]. Available: http://www.azonano.com/article.aspx?ArticleID=2740..

[231] "Technology," MicroFab, [Online]. Available: http://www.microfabnh.com/ion_beam_etch_technology.php.

[232] P. B. Zantye, A. Kumar and A. K. Sikder, "Chemical mechanical planarization for microelectronics applications," *Materials Science and Engineering,* vol. R 45, pp. 89-220, 2004.

[233] Alpsitec, "CMP Process," [Online]. Available: https://www.crystec.com/alpovere.htm.

[234] C. Van Dommelen. [Online]. Available: http://www.embedded.com/print/4382211.

[235] KLA-Tencor, "Surfscan SP1 Operations and Overview Course," [Online]. Available: http://slideplayer.com/slide/10915236/.

[236] Hitachi, "Wafer Defect Inspection Systems," [Online]. Available: http://www.hitachi-hightech.com/global/products/device/semiconductor/inspection.html.

[237] N. Raqccah, D. Rogers, E. Tzuri and A. Shavit, "Extending inspection limits using a DUV-laser-based bright-field inspection tool," *MICRO Magazine,* July 2006.

[238] T. K. Tran, "Intel Presentation," [Online]. Available: https://www.nist.gov/sites/default/files/documents/pml/div683/conference/tran.pdf).

[239] Tsutomu Horie and Yuichi Tokumaru, US Patent # 2006/0147104, July 6, 2006.

[240] B. Lydon, "PLC vs. DCS - Competing Process Control Philosophy," Automation.com, [Online]. Available: http://www.automation.com/automation-news/article/plc-vs-dcs-competing-process-control-philosophy.

[241] D.-Y. Liao, "Automation and Integration in Semiconductor Manufacturing," in *Semiconductor Technologies*, J. Grym, Ed., InTech Open Science, 2010.

[242] “Seamless Integrated Automation Using the MKS Automation Platform,” 2017. [Online]. Available: https://www.mksinst.com/docs/R/IntegratedAutomationPlatform-AppNote.pdf.

[243] D. Hazen, “Modular Platforms for Integrated Automation,” *Processing Solutions for the Process Industries,* 22 August 2016.

[244] “GlobalFoundries ITDC Fab 8,” [Online]. Available: https://www.youtube.com/watch?v=-SCskPV00kU.

[245] “The Process of Manufacturing the Chip,” [Online]. Available: https://www.youtube.com/watch?v=zoBw2cSmFC0.

[246] D.-Y. Liao, “Automation and Integration in Semiconductor Manufacturing,” in *Semiconductor Technologies*, J. Grym, Ed., InTech, 2010.

[247] “Semiconductor manufacturing intelligence and automation,” *Computers & Industrial Engineering,* vol. 99, pp. 315-473, September 2016.

[248] S. J. Quin, G. Cherry, R. Good, J. Wang and C. A. Harrison, “Semiconductor manufacturing process control and monitoring: A fab-wide framework,” *Journal of Process Control,* vol. 16, no. 3, pp. 179-191, 2006.

[249] J. Cotner, “Semiconductor 101: Functionality and Manufacturing of Integrated Circuits AMF-ENT-T0194,” NXP Semiconductors, [Online]. Available: http://www.nxp.com/files/training/doc/dwf/DWF13_AMF_ENT_T0194.pdf.

[250] C. Gottschalk, K. McLaughlin, J. Cren, C. Peyne and P. Valenti, “The Use and Management of Chemical Fluids in Microelectronics,” in *Chemistry in Microelectronics*, C. L. T. Yannick, Ed., London, John Wiley & Sons, 2013.

[251] “Ozone Products,” MKS Instruments Inc., [Online]. Available: https://www.mksinst.com/product/Category.aspx?CategoryID=554.

[252] “Typical Bulk Liquid Storage,” Air Products, [Online]. Available: http://www.airproducts.com/products/Gases/supply-options/bulk-deliveries-and-storage-systems/typical-bulk-liquid-storage-systems.aspx.

[253] “A Quick Guide to Cryogenics,” National Optical Astronomy Observatory, [Online]. Available: https://www.noao.edu/kpno/manuals/irim/cryo.html.

[254] “Compressed Gas Cylinder Safety,” University of Illinois, Division of Research Safety, [Online]. Available: https://www.drs.illinois.edu/SafetyLibrary/CompressedGasCylinderSafety.

[255] M. E. Coltrin, P. Ho, H. K. Moffat and R. J. Buss, “Chemical kinetics in chemical vapor deposition: growth of silicon dioxide from tetraethoxysilane (TEOS),” *Thin Solid Films,* vol. 365, p. 251, 2000.

[256] Y. Gu, “Method for removing condensable aluminum chloride vapor from aluminum etch effluent”. USA Patent 6790258 B2, 14 September 2004.

[257] “Cleanroom Classification”.*Brigham Young University.*

[258] [Online]. Available: https://www.youtube.com/watch?v=KrXE_SzRoqw.

This page intentionally left blank.

Appendix A

Electrical Conduction in Solids

The physical mechanism of electrical conduction in solids is best understood using a model that physicists refer to as "band theory", which is based on the idea that the state of an electron in a solid is described by a four-fold scheme of quantum numbers [3] [258]. A useful analogy for electrons in a solid is that of spectators in an amphitheater. The electrons can be considered as spectators sitting in the amphitheater and the quantum numbers describe how many rows and seats are available. The energy state of individual electrons is described by a combination of four quantum numbers; this can be considered as analogous to the position of a spectator in an amphitheater which is described by a particular row and seat. Like spectators in an amphitheater moving between seats and rows, electrons can change their status, provided that there are spaces available for them to move into and provided that they have sufficient energy to make the move. If an electron (spectator) is to move into a higher-energy electron shell (higher row seat), it requires that additional energy to be given to the electron from an external source to move it further out in the electrical potential of the atom's nucleus. Using the amphitheater analogy, it takes an increase in energy (i.e., work) for a person to move into a higher row of seats, because that person must climb to a greater height against the force of gravity. Conversely, an electron "leaping" into a lower shell gives up some of its energy, like a person jumping down into a lower row of seats.

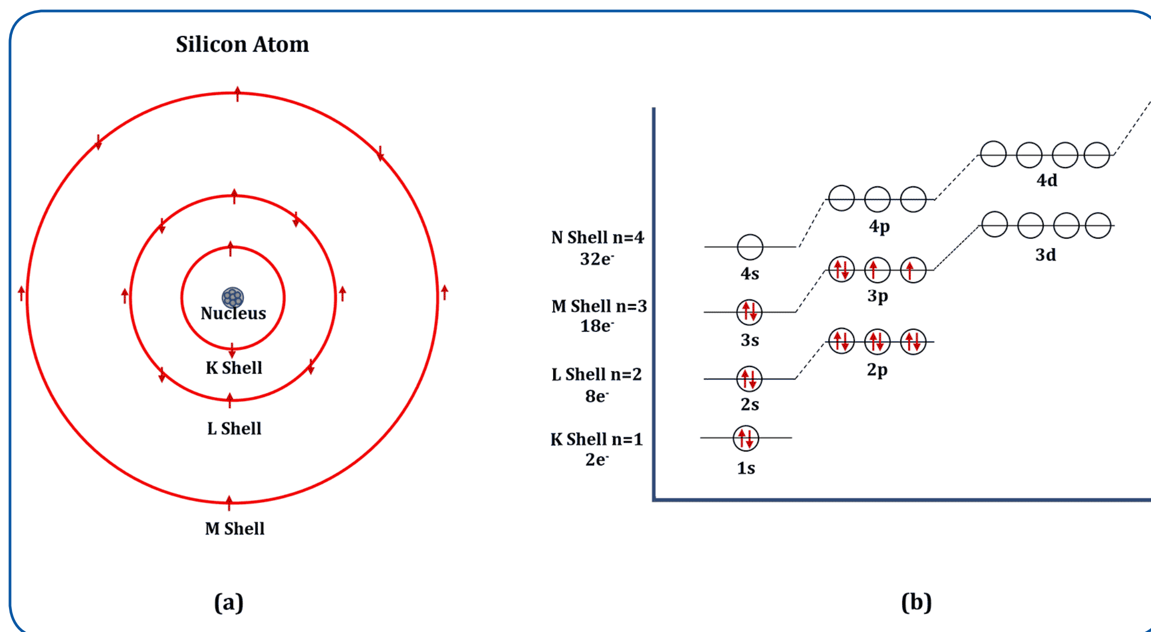


Figure 243. The silicon atom, showing (a) electron orbitals and (b) relative orbital energies.

In band theory, the electrons in a solid exist within "energy bands" (using our amphitheater analogy, energy bands can be viewed as seating sections containing many rows with equal quality seating). The model for the manner by which energy bands arise in a solid is most easily understood by first considering the energy of electrons in a single atom. Figure 243(a), shows a Bohr depiction of a single silicon atom. In this model of the atom, electrons occupy discrete orbitals or "shells" around the central atomic nucleus and each shell has a characteristic energy. For the silicon atom depicted in Figure 243(a), the atom's 14 electrons occupy three shells. The shells in (a) are designated, proceeding from the inside out, as K, L and M (there are higher shells, designated N, O, P, etc. in larger atoms). Different shells hold different numbers

of electrons: the K shell holds two electrons; the L shell holds eight electrons; and the M shell holds 18 electrons (the N shell holds 32). In the very simplistic terms of the Bohr model of the atom, the number of electrons that a shell can hold correlates, very roughly, with the available surface area of the sphere defined by that shell, as one might expect from simple considerations of the repulsive forces atoms. In isolated atoms, the electrons exist in discrete electron orbital shells (labelled *s* and *p* in Figure 243). The energy gap between particles having the same charge (modern models of atomic structure are much more complex than this simple view).

Within each of the electron shells of an atom, electrons can occupy sub-shells of varying energy; these are designated by the letter *s, p, d* or *f*. Figure 243(b) shows these sub-shells represented as circles and the relative energies and ordering of the shells and sub-shells. Each sub-shell (circle) can hold a maximum of two electrons. The smallest K shell is only big enough to accommodate a single *s* sub-shell and therefore contains a maximum of two electrons. The larger L shell can accommodate one *s* sub-shell and three higher energy *p* sub-shells. Each *p* sub-shell is iso-energetic with the other two but is differentiated from its two partners by quantum numbers that we will not consider in this simplified discussion. The availability of *s* and *p* sub-shells allows the L shell to hold up to 8 electrons. The M shell includes one *s* sub-shell, three *p* sub-shells and five *d* sub-shells (as with the *p* sub-shells, the five *d* sub-shells are iso-energetic) and accommodates up to 18 electrons; higher energy shells hold correspondingly greater numbers of electrons. When the available sub-shells within a shell are all filled by electrons, then that shell is said to be "closed" and the electrons within that shell are no longer available for chemical interactions or electrical conduction. Silicon, shown in Figure 243(a), has 14 electrons. These electrons are sequentially accommodated in the available shells, starting at the lowest energy K shell. For silicon, the K and L shells are completely filled (10 electrons) and the M shell is unfilled, with only four electrons occupying the available nine sub-shells (the nine sub-shells can accommodate up to 18 electrons); these electrons are distributed in the lower energy *s* and *p* sub-shells. These four electrons in the outer or highest energy electron shell, are commonly termed the "valence" electrons; they determine silicon's chemical and electrical properties. An alternate way of depicting the energy shells and sub-shells in the silicon atom and the way that the sub-shells are filled is shown in Figure 244. Figure 244 clearly shows the "closed" shell core electrons in the silicon K and L shells and the valence electrons in the unfilled M shell.

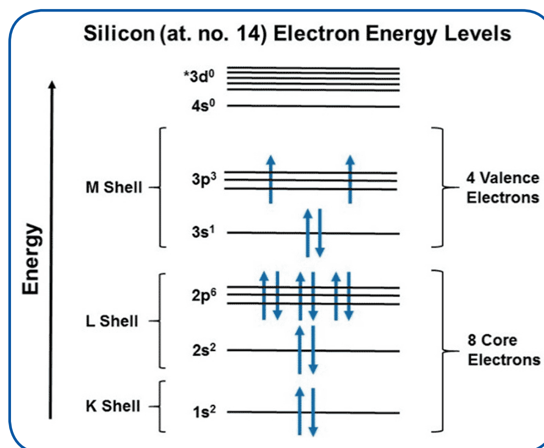


Figure 244. Relative electron energy levels and sub-shell filling pattern for the silicon atom.

Within each of the electron shells of an atom, electrons can occupy sub-shells of varying energy; these are designated by the letter *s, p, d* or *f*. Figure 243(b) shows these sub-shells represented as circles and the relative energies and ordering of the shells and sub-shells. Each sub-shell (circle) can hold a maximum of two electrons. The smallest K shell is only big enough to accommodate a single *s* sub-shell and therefore contains a maximum of two electrons. The larger L shell can accommodate one *s* sub-shell and three higher energy *p* sub-shells. Each *p* sub-shell is iso-energetic with the other two but

is differentiated from its two partners by quantum numbers that we will not consider in this simplified discussion. The availability of *s* and *p* sub-shells allows the L shell to hold up to 8 electrons. The M shell includes one *s* sub-shell, three *p* sub-shells and five *d* sub-shells (as with the *p* sub-shells, the five *d* sub-shells are iso-energetic) and accommodates up to 18 electrons; higher energy shells hold correspondingly greater numbers of electrons. When the available sub-shells within a shell are all filled by electrons, then that shell is said to be "closed" and the electrons within that shell are no longer available for chemical interactions or electrical conduction. Silicon, shown in Figure 243(a), has 14 electrons. These electrons are sequentially accommodated in the available shells, starting at the lowest energy K shell. For silicon, the K and L shells are completely filled (10 electrons) and the M shell is unfilled, with only four electrons occupying the available nine sub-shells (the nine sub-shells can accommodate up to 18 electrons); these electrons are distributed in the lower energy *s* and *p* sub-shells. These four electrons in the outer or highest energy electron shell, are commonly termed the "valence" electrons; they determine silicon's chemical and electrical properties. An alternate way of depicting the energy shells and sub-shells in the silicon atom and the way that the sub-shells are filled is shown in Figure 244. Figure 244 clearly shows the "closed" shell core electrons in the silicon K and L shells and the valence electrons in the unfilled M shell.

Now, when two or more atoms bind together, the electrons in each atom's valence sub-shell combine to create *molecular* sub-shells that accommodate all of the electrons from the atoms' valence sub-shells; it is this combination that constitutes the interatomic bond. Consider two silicon atoms coming together to produce a hypothetical "Si_2" molecule. Each silicon atom contributes four electrons from its valence shell (two electrons in the single 3*s* sub-shell and two electrons in two of the three available 3*p* sub-shells) to form the molecular bonds between the two atoms in the Si_2 molecule with eight electrons occupying eight molecular sub-shells (Remember, the core electrons don't participate in the bonding or electrical conduction). This energy levels of the molecular orbitals produced by the combination of the two Si atoms' atomic orbitals are shown graphically in Figure 245.

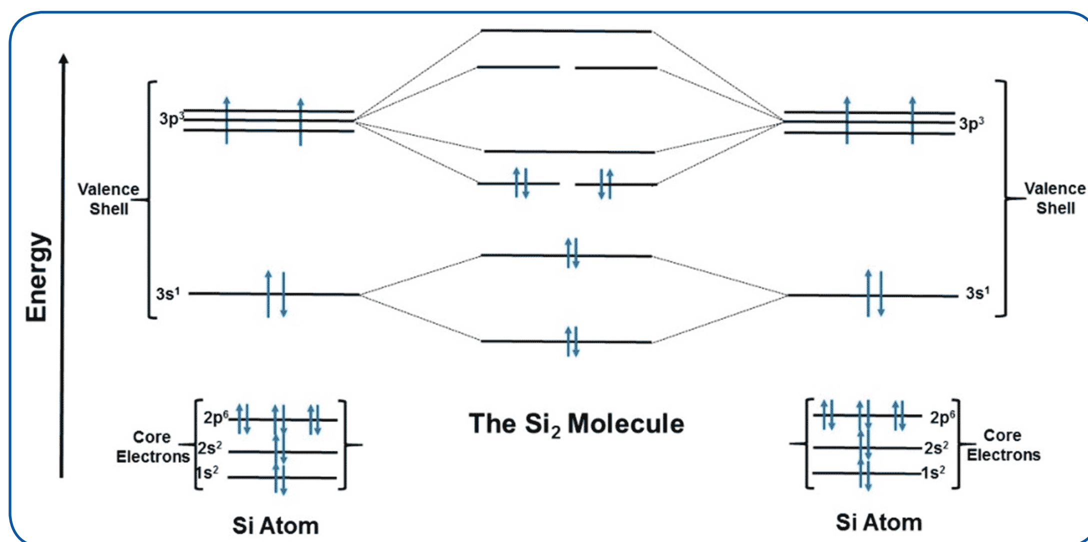


Figure 245. *Relative electronic energy levels and the available molecular sub-shells in the Si_2 molecule.*

Two molecular sub-shells are produced for every sub-shell in the individual silicon atom and the molecular sub-shell energy levels are arrayed above and below the energy of the corresponding atomic sub-shell, as depicted in Figure 245. Not all of the available molecular sub-shells in the Si_2 molecule are needed to accommodate the valence electrons of the two silicon atoms. This is the case for all molecules; indeed, if all of the available molecular sub-shells were filled with electrons, no bonding could occur between the two Si atoms (this arises from the fact that molecular sub-shells are produced as bonding and anti-bonding pairs –a subject that we don't need to get into in this discussion).

So, when two identical atoms combine, the available energy levels for the electrons from those atoms are doubled (the situation is somewhat more complex when two dissimilar atoms combine). Similarly, if three identical atoms combine, then there are three sub-shells generated for each sub-shell in the original atom; with four atoms, there are four sub-shells for each sub-shell in the original atom, etc. For descriptive purposes, let's use the example of a chain of *n* bonded hydrogen atoms each with a single available energy sub-shell to contribute to the molecule (in this case one electron in the 1*s* sub-shell). As atoms are progressively added to our hypothetical linear poly-hydrogen molecule, additional energy sub-shells are created between the upper and lower energy limits of the 1*s* shell. With each additional atom, the energy increment between sub-shells becomes smaller and the energy gap between the shells becomes slightly smaller. By the time enough atoms are added to the assemblage for it to be considered an extended solid, the number of atoms can be considered as effectively infinite. In terms of the sub-shells available for electrons within the solid, the ever-decreasing energy increment between sub-shells produces a band of available energy levels within which an electron can freely move rather than the discrete levels characteristic of isolated atoms or molecules. The continuous band of energy levels created in our hypothetical poly-hydrogen molecule corresponds to the 1*s* sub-shell in the isolated hydrogen atom. This band contains the accumulated electrons contributed by all of the hydrogen atoms in the molecule, creating a half-filled 1*s energy band*, as shown in Figure 246.

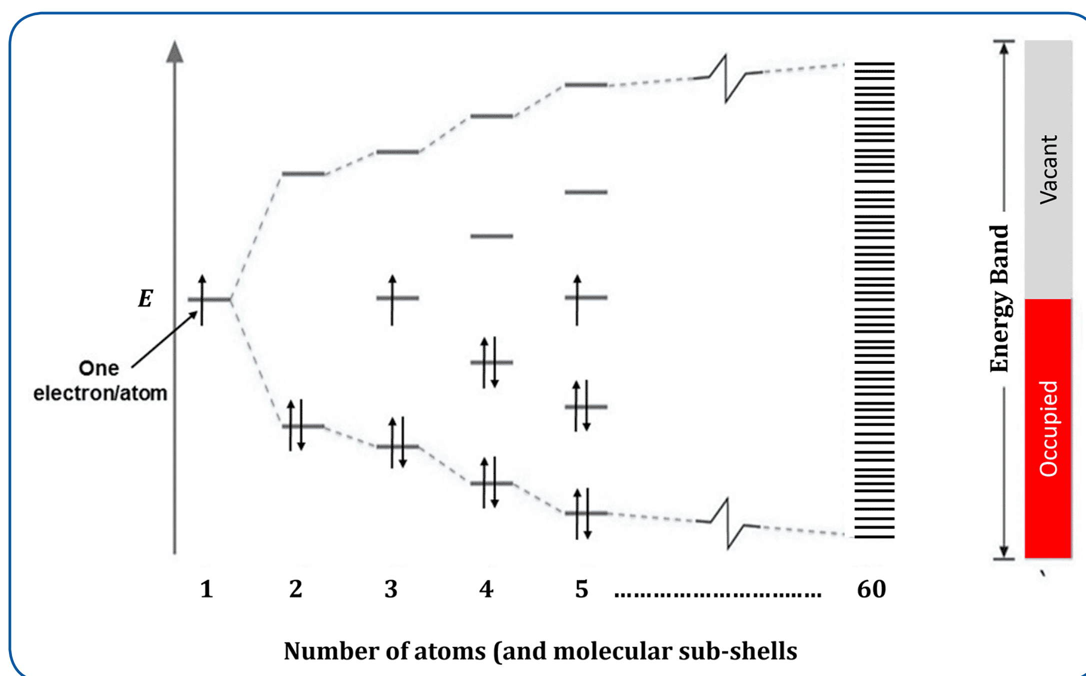


Figure 246. The arrangement of molecular energy sub-shells for a linear chain of n hydrogen atoms, *each of which contains a singly occupied s orbital.*

By analogy, similar molecular arrangements of heavier atoms such as silicon can produce more than one energy band through a combination of the available orbitals in the valence shell (whereas hydrogen only has a single sub-shell in the valence band, silicon has nine, four of which, the 3*s* and 3*p* sub-shells, are relevant to our discussion). When silicon atoms combine to produce the infinite atomic array that is a silicon crystal, energy bands are formed corresponding to the atomic s and p sub-shells. An examination of Figure 245 shows that the *s* and two of the three *p* sub-shells can each be expected to form an energy band that is completely filled by the available electrons. Additionally, an empty energy band, due to the vacant third *p* sub-shell, is also formed at an energy level relatively close to the filled *p* bands. These filled and empty energy bands respectively constitute the *valence band* and *conduction band* in the electronic energy model used to describe the electrical characteristics of solids.

Metal

Valence Shell (s) (Occupied by Electrons)

Conduction Shell (p) (Empty)

Energy

P

S

Single Atom

Single Atom

Significant energy required to move an electron from occupied S level to empty P level

Five Atoms in Close Proximity

10 sub-shells and reduced gap between occupied and empty states

Conduction Band, E_c

Overlap

Valence Band, E_v

Multiple Atoms in Close Proximity

Infinite number of sub-shells form an energy band; In metals, the Valence and Conduction bands overlap, allowing free movement of the electrons throughout the solid

Insulator

Valence Shell (s) (Occupied by Electrons)

Conduction Shell (p) (Empty)

Energy

P

S

Single Atom

Single Atom

Significant energy required to move an electron from occupied S level to empty P level

Five Atoms in Close Proximity

10 sub-shells and reduced gap between occupied and empty states

Conduction Band, E_c

Energy Gap

Valence Band, E_v

Multiple Atoms in Close Proximity

Infinite number of sub-shells form energy bands; In an insulator the Valence and Conduction bands are separated by a large Energy Gap, preventing free movement of electrons in the solid

Semiconductor

Valence Shell (s) (Occupied by Electrons)

Conduction Shell (p) (Empty)

Energy

P

S

Single Atom

Single Atom

Significant energy required to move an electron from occupied S level to empty P level

Five Atoms in Close Proximity

10 sub-shells and reduced gap between occupied and empty states

Conduction Band, E_c

Energy Gap

Valence Band, E_v

Multiple Atoms in Close Proximity

In a semiconductor, the Valence and Conduction bands are separated by a small Energy Gap, allowing movement of electrons in the solid with the addition of a small amount of energy

Figure 247. Electron energy band arrangement in a metal, an insulator and a semiconductor.

Figure 247 shows the relationship between filled (valence) and empty (conduction) bands in the different electrical classifications of solids. The Valence Band, denoted as E_v in Figure 247 has every sub-shell occupied by two electrons. The Conduction Band, denoted as E_c, is completely made up of vacant sub-shells. If an electron can be promoted from the Valence to the Conduction band, electrical current will flow within the solid since there are vacant sub-shells in that band that energy bands is denoted by E_g in Figure 247 and it is this gap that determines a solid material's permit electrons to flow when an electrical potential is applied. The energy gap separating these electrical properties. Figure 247 shows the band gap arrangement in a metal, a semiconductor and an insulator. The band gaps overlap in a metal so that almost no energy is required for an electron to move into the conduction band, becoming available for electrical conduction. The value of the band gap (E_g) in a metal is therefore 0, allowing free flow of electrons. In a semiconductor, the band gap is relatively small (E_g of silicon is approximately 1 electron volt, E_g in other semiconductors can be up to 2 eV) and electrons can be promoted into the conduction band by either thermal or photonic energy (hence the photoconductive behavior of semiconductors). Insulators such as glass (silicon dioxide) have large band gaps (E_g of at least 5 electron volts) and large amounts of energy are required to promote an electron into the conduction band, effectively preventing electrical conduction in these materials.