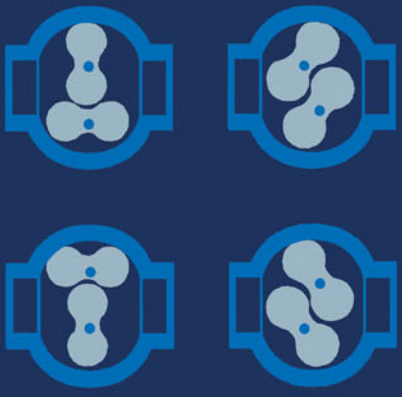
HANDBOOK OF VACUUM SCIENCE AND TECHNOLOGY



Dorothy M. Hoffman
Bawa Singh
John H. Thomas, III

HANDBOOK OF VACUUM SCIENCE AND TECHNOLOGY

This Page Intentionally Left Blank

HANDBOOK OF VACUUM SCIENCE AND TECHNOLOGY

Edited by

Dorothy M. Hoffman
(*deceased*)

Bawa Singh

David M. Sarnoff Research Center
Princeton, New Jersey

John H. Thomas, III

3M Research Laboratories
St. Paul, Minnesota

ACADEMIC PRESS
An Imprint of Elsevier
San Diego London Boston
New York Sydney Tokyo Toronto

ACADEMIC PRESS
An Imprint of Elsevier
525 B Street, Suite 1900, San Diego, CA 92101-4495, USA
1300 Boylston Street, Chestnut Hill, MA 02167, USA
http://www.apnet.com

Academic Press Limited
24–28 Oval Road, London NW1 7DX, UK
http://www.hbuk.co.uk/ap/

Library of Congress Cataloging-in-Publication Data

Hoffman, Dorothy M.
Handbook of vacuum science and technology / Dorothy M. Hoffman, Bawa Singh, John H. Thomas, III.
p. cm.
Includes bibliographical references and index.
ISBN-13: 978-0-12-352065-4 ISBN-10: 0-12-352065-7 (alk. paper).
1. Vacuum technology—Handbooks, manuals, etc. I. Singh, Bawa.
II. Thomas, John H.
TJ940.H59 1997 97-26269
621.5'5—dc21 CIP

ISBN-13: 978-0-12-352065-4
ISBN-10: 0-12-352065-7

Printed in the United States of America
05 06 07 08 09 IC 9 8 7 6 5 4 3 2

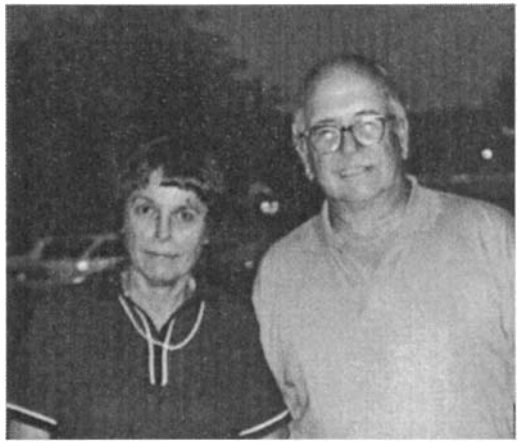
A Short Foreword

H. F. Dylla

When Dorothy Hoffman asked a contingent of her many colleagues in the vacuum science and technology community to contribute chapters to a new *Handbook of Vacuum Science and Technology*, we could not have known that this handbook would become Dorothy's last contribution to the field. Dorothy's untimely death in December of 1996 deprived our world — the international vacuum science and technology community — and the world of music — Dorothy was a concert pianist with the Philadelphia Academy of Music — with a passionate voice for both the science and the arts. Dorothy was among the founding members, and the first woman president, of the American Vacuum Society, incorporated in 1953 to strengthen the coupling of advances in vacuum technology to the burgeoning fields of science (particle and plasma physics, materials, science, space science) and engineering (solid state electronics, thin film optics, materials processing) that require carefully controlled vacuum environments. When Dorothy retired in 1990 from her position as head of the Thin Film Laboratory at the David Sarnoff Research Center in Princeton, New Jersey, she had helped ensure the growth and maturity of thin film technology — an endeavor essential to the manufacture of microelectronics and high-performance optics. Dorothy continued to contribute after she retired by volunteering her talents and the benefits of her experience to the American Vacuum Society. As part of the Society's 40th anniversary in 1993, Dorothy conceived the project that resulted in this handbook. Thanks to the efforts of two of Dorothy's colleagues from Sarnoff Laboratory, Bawa Singh and John Thomas, who guided the final editing, our community now has a new compendium of the underlying science, the technology for producing and controlling vacuum, and the numerous applications that require vacuum. Those fortunate to know Dorothy and those who will share her work in reading this handbook have much to thank her for.

H. F. Dylla
Past President, American Vacuum Society
Jefferson Laboratory
Newport News, VA

This Page Intentionally Left Blank



Dorothy M. Hoffman
A History and Dedication

We wish to dedicate this book to our friend and co-worker Dorothy Hoffman. Dorothy passed away December 13, 1996. This book represents the culmination of a lifetime of work and dedication to the area of thin-film development.

Dorothy graduated from Rensselaer Polytechnic Institute and in 1949 received a M.S. degree from Bucknell University. She then joined the International Resistor Corp. in Philadelphia and worked on thin films. Her first paper in thin-film technology was presented at the AVS National Symposium in 1954 and was one of the first papers on this topic. That same year she joined the AVS and continued as a very active member until her death. Over the years she served the Society as Secretary, as a member of the Long Range Planning Committee, a Trustee and as Thin Film Committee Chair (before divisions existed). In 1974, she became the first woman to serve as President of the AVS. In later years, she served as an advisor to the Thin Film Division and the Board of Directors. She was made an honorary member in 1981. Dorothy was one of the founders of the Thin Film Division and was a founding member of the Delaware Valley Chapter, serving as secretary, vice-chair and chair. All this . . . and more during her 42 years in the area of thin films.

After working at IRC for 10 years, Dorothy joined RCA Laboratories in 1962. When RCA was sold to GE, she stayed on at the labs, then a subsidiary of SRI International, through her retirement in 1990. During her RCA years, she was head of the Thin Film Laboratory. Her technical activities included design and application of thin films to passive optical devices, integrated circuits, thin-film circuits, video disks, and many other common and uncommon applications. When she retired, she couldn't stay inactive. So she started her own consulting firm where she continued her work in the area of thin films. During this time, she was heavily involved in editing and writing the *Handbook of Vacuum Technology*. She wanted to finish this project before she was 70—this was sadly not to be.

In addition to her professional society activities, Dorothy was an active member of the Engineers' Club of Philadelphia, Director of and Representative to Engineering and Technical Societies. She was a long-time member and past president of the Society of Women Engineers and an active member of the local chapter.

In addition to her professional activities, she was a gifted concert pianist and performed at the Academy of Music in Philadelphia. She was a gardener, wildlife lover—an all-around special person. She was married to Earl Hoffman, who died October 1996, and lived in Titusville, New Jersey.

Dorothy brought a passionate devotion to every one of the many things that she did, and she will be deeply missed by all her friends and colleagues.

John H. Thomas, III

Contents

Preface xvii

List of Contributors xxi

Part 1 Fundamentals of Vacuum Technology and Surface Physics 1

1.1 Vacuum Nomenclature and Definitions 3
1.1.1 Basic Definition 3
1.1.2 Pressure Regions of Vacuum 3
1.2 Gas Properties 8
1.2.1 Description of Vacuum as a Low-Pressure Gas 8
1.2.2 Characteristics of a Gas — Basic Definitions 8
1.2.3 Gas Laws 9
1.3 Molecular Processes and Kinetic Theory 11
1.3.1 General Description 11
1.3.2 Molecular Motion 12
1.3.3 Kinetic Theory Derivation of the Gas Laws 14
1.3.4 Pressure 15
1.3.5 Molecular Mean Free Path 17
1.3.6 Number of Impacts with the Chamber Wall 19
1.3.7 Time to Form a Monolayer 20
1.3.8 Thermal Transpiration 20
1.3.9 Coefficient of Thermal Conductivity 21
1.3.10 Coefficient of Diffusion 21
1.4 Throughput, Pumping Speed, Evacuation Rate, Outgassing Rate, and Leak Rate 22
1.5 Gas Flow 25
1.5.1 Nature of Gas Flow 25
1.5.2 Turbulent Flow 27

1.5.3 Viscous, Streamline, or Laminar Flow 28
1.5.4 Molecular Flow 29
1.5.5 Flow Relationships 29
1.6 Conductance 32
1.6.1 Conductance 32
1.6.2 Conductances in Parallel 33
1.6.3 Conductances in Series 33
1.7 Flow Calculations 35
1.7.1 Equations for Viscous Flow 35
1.7.2 Equations for Molecular Flow 37
1.7.3 Knudsen's Formulation 37
1.7.4 Clausing Factors 38
1.8 Surface Physics and Its Relation to Vacuum Science 40
1.8.1 Physical Adsorption or "Adsorption" 40
1.8.2 Chemisorption 42
1.8.3 Sticking Coefficient 43
1.8.4 Surface Area 44
1.8.5 Surface Adsorption Isotherms 45
1.8.6 Capillary Action 47
1.8.7 Condensation 48
1.8.8 Desorption Phenomena 49
1.8.9 Thermal Desorption 50
1.8.10 Photoactivation 52
1.8.11 Ultrasonic Desorption 53
1.8.12 Electron- and Ion-Stimulated Desorption 53
1.8.13 Gas Release from Surfaces 54
References 55

Part 2 Creation of Vacuum 57

2.1 Technology of Vacuum Pumps — An Overview 59
2.1.1 Vacuum Pump Function Basics 59
2.1.2 Gas Transport: Throughput 61
2.1.3 Performance Parameters 62
2.1.4 Pumping Speed 64
2.1.5 Pumpdown Time 65
2.1.6 Ultimate Pressure 69
2.1.7 Forevacuum and High-Vacuum Pumping 71
2.1.8 Pump System Relationships 73
2.1.9 Crossover from Rough to High-Vacuum Pumps 78
2.1.10 Pumping System Design 79
References 83

2.2 Diaphragm Pumps 84
2.2.1 Introduction: Basics and Operating Principle 84
2.2.2 State-of-the-Art Design and Manufacturing 87
2.2.3 Performance and Technical Data 91
2.2.4 Modular Concept for Specific Application Setups: Standalone Operation 92
2.2.5 Diaphragm Pumps as Backing and Auxiliary Pumps in Vacuum Systems 93
References 96
2.3 Vacuum Blowers 97
2.3.1 Introduction 97
2.3.2 Equipment Description 97
2.3.3 Blower Operating Principle 100
2.3.4 Blower Pumping Efficiency 101
2.3.5 Blower Pumping Speed Calculations 103
2.3.6 Power Requirements 104
2.3.7 Temperature Considerations 106
2.3.8 Flow and Compression Ratio Control Mechanisms 108
2.3.9 Liquid-Sealed Blowers 112
2.3.10 Selected System Arrangements 112
2.4 Vacuum Jet Pumps (Diffusion Pumps) 116
2.4.1 Basic Pumping Mechanism 117
2.4.2 Pumping Speed 122
2.4.3 Throughput 127
2.4.4 Tolerable Forepressure 128
2.4.5 Ultimate Pressure 132
2.4.6 Backstreaming 137
2.4.7 Other Performance Aspects 144
References 148
2.5 Cryogenic Pumps 149
2.5.1 Introduction 149
2.5.2 Cryopump Basics 156
2.5.3 Advanced Control Systems 167
2.5.4 Cryopump Process Applications 173
2.5.5 Cryogenic Pumps Specifically for Water Vapor 177
2.5.6 Comparison of Cryopumps to Other Types of Pumps 179
2.5.7 Future Developments 181
References 181
2.6 Turbomolecular Pumps 183
2.6.1 Turbomolecular Pumps (TMP) 183
2.6.2 Molecular Drag Pumps (MDP) 195
2.6.3 Combination of Pumps (TMP + MDP) 197

2.6.4 Evaluation of Combinations of Backing Pumps and TMPs, Etc. ... 200
2.6.5 The Use of TMP in Applications: Specific Effects and Demands ... 208
2.6.6 Avoiding Operational Mistakes ... 211
References ... 212
2.7 Pumps for Ultra-High Vacuum Applications ... 214
2.7.1 System Design for Ultra-High Vacuum ... 215
2.7.2 The Selection of Pumps for Ultra-High Vacuum Applications ... 216
2.7.3 Sputter-Ion Pumps ... 220
2.7.4 Getter Pumps ... 242
References ... 252

Part 3 Vacuum Measurements ... 255

3.1 The Measurement of Low Pressures ... 257
3.1.1 Overview ... 258
3.1.2 Direct Reading Gauges ... 260
3.1.3 Indirect Reading Gauges ... 265
3.1.4 Calibration of Vacuum Gauges ... 286
References ... 288
3.2 Mass Analysis and Partial Pressure Measurements ... 290
3.2.1 Overview and Applications ... 290
3.2.2 Inlet Systems ... 300
3.2.3 Ion Generation and Ion Sources ... 303
3.2.4 Ion Separation Analyzers ... 308
3.2.5 Detection of Ions ... 323
References ... 326
3.3 Practical Aspects of Vacuum System Mass Spectrometers ... 335
3.3.1 Historical Insight ... 335
3.3.2 Expected Gases in a Vacuum System ... 336
3.3.3 The Ion Generation Process ... 340
3.3.4 Techniques for Analysis ... 351
3.3.5 Calibration of Vacuum System Mass Spectrometers ... 364
3.3.6 Some Applications ... 370
References ... 374
3.4 Mass Flow Measurement and Control ... 376
3.4.1 General Principles of Mass Flow Measurement ... 376
3.4.2 Overview of Thermal Mass Flow Controller Technology 378
3.4.3 Performance Characteristics ... 382
3.4.4 Troubleshooting ... 386
References ... 387

Part 4 Systems Design and Components 389

4.1 Selection Considerations for Vacuum Valves 391
4.1.1 Introduction 391
4.1.2 Valves for Shutoff 391
4.1.3 Valves for Control 397
4.1.4 Valve Construction 398
4.1.5 Specialty Valves 404
4.1.6 Installation Considerations for Vacuum Valves 407
References 408
4.2 Flange and Component Systems 409
4.2.1 Introduction 409
4.2.2 Selecting a Flange System 410
4.2.3 Common Flange Systems 410
4.2.4 Components with Flanges Attached 425
Trademarks 430
References 432
4.3 Magnetic-Fluid-Sealed Rotary Motion Feedthroughs 433
4.3.1 Basic Sealing Principle 433
4.3.2 Application Factors 434
4.3.3 Impact of Feedthrough on Process 436
4.3.4 Impact of Process on Feedthrough 437
4.3.5 Materials Considerations 438
4.3.6 Application Examples 440
4.3.7 Comparison to Other Types of Feedthroughs 442
4.4 Viewports 444
4.4.1 Materials 444
4.4.2 Mounting Systems and Precautions 445
4.5 Construction Materials 446
4.5.1 Properties Defining Material Performance 446
4.5.2 Vacuum Chamber Materials 451
4.5.3 Special-Purpose Materials 455
References 462
4.6 Demountable Seals for Flanges and Valves 463
4.6.1 Sealing Overview: Polymer and Metal Seals 463
4.6.2 The Elastomeric and Nonelastomeric Polymers Used in Vacuum Sealing 464
4.6.3 Metal Seals 474
References 482
4.7 Outgassing of Materials 484
4.7.1 Relationships Among System Pressure, Pumping Speed, and Outgassing 484

4.7.2 Initial Pumpdown from Atmospheric Pressure 494
4.7.3 Pressure Vs. Time During Outgassing..... 495
4.7.4 The Outgassing Rate of Elastomers and Plastics..... 497
4.7.5 The Outgassing Rate of Metals and Ceramics..... 501
4.7.6 The Outgassing Rate of Preconditioned Vacuum Systems After Short Exposure to the Atmosphere 504
4.7.7 Methods of Decreasing the Outgassing Rate..... 506
4.7.8 Measurement of the Outgassing Rate of Materials 507
References..... 508
4.8 Aluminum-Based Vacuum Systems 509
4.8.1 Outgassing..... 509
4.8.2 Demountable Seals..... 512
4.8.3 Cleaning and Surface Finishing..... 518
4.8.4 Mechanical Considerations 520
4.8.5 Thermal Conductivity and Emissivity..... 536
4.8.6 Corrosion..... 538
4.8.7 Welding Aluminum for Vacuum Applications 541
References..... 548
4.9 Preparation and Cleaning of Vacuum Surfaces..... 553
4.9.1 Surface Modification..... 554
4.9.2 External Cleaning 567
4.9.3 Assembly, Handling, and Storage 587
4.9.4 *In Situ* Cleaning..... 591
4.9.5 Documentation..... 599
4.9.6 Conclusion..... 601
Trade Names..... 601
References..... 601

Part 5 Vacuum Applications 607

5.1 High-Vacuum-Based Processes: Sputtering 609
5.1.1 Sputtering and Deposition..... 611
5.1.2 Sputter Deposition Technologies..... 612
5.1.3 Magnetron Applications 624
5.1.4 Future Directions in Sputtering..... 626
References..... 627
5.2 Plasma Etching 628
5.2.1 Introduction..... 628
5.2.2 Review of Plasma Concepts Applicable to Etching Reactors..... 628
5.2.3 Basic Plasma Etching Requirements..... 633
5.2.4 Plasma Diagnostics 641
5.2.5 Basic Plasma Etch Reactors..... 643

5.2.6 Advanced Plasma Etch Reactors 649
5.2.7 New Trends 665
References 667
5.3 Ion Beam Technology 672
5.3.1 Introduction 672
5.3.2 Ion Beam Etching 678
5.3.3 Ion Beam Sputter Deposition 683
5.3.4 Ion-Beam-Assisted Deposition 687
5.3.5 Ion Beam Direct Deposition 689
5.3.6 Conclusion 690
References 691
5.4 Pulsed Laser Deposition 694
5.4.1 Introduction 694
5.4.2 Pulsed Laser Deposition System 695
5.4.3 The Ablation Mechanism 698
5.4.4 Advantages and Limitations 700
5.4.5 Materials Survey 705
5.4.6 Future Outlook 708
References 708
5.5 Plasma-Enhanced Chemical Vapor Deposition 711
5.5.1 Introduction 711
5.5.2 Equipment and Other Practical Considerations 717
5.5.3 Process Scaleup 723
5.5.4 Conclusion 727
References 728
5.6 Common Analytical Methods for Surface and Thin Film 731
5.6.1 Introduction 731
5.6.2 The Electron Spectroscopies 732
5.6.3 Methods Based on Ion Bombardment 745
5.6.4 UHV Generation and System Considerations for Surface Analysis 755
References 757

Part 6 Large-Scale Vacuum-Based Processes 759

6.1 Roll-to-Roll Vacuum Coating 761
6.1.1 Overview of Roll-to-Roll Vacuum Coating 761
6.1.2 Typical Products 764
6.1.3 Materials and Deposition Processes Commonly Used in Roll-to-Roll Coating 765
6.1.4 Vacuum Systems for Roll-to-Roll Coating Applications. 775
6.1.5 Substrates (Webs) 779
6.1.6 Process Control 783

6.1.7 Specific Problems Exhibited by Coatings 784
References 787
6.2 The Development of Ultra-High-Vacuum Technology for Particle Accelerators and Magnetic Fusion Devices 789
6.2.1 Introduction 789
6.2.2 Storage Rings and the Need for UHV 790
6.2.3 UHV for Early Storage Rings 793
6.2.4 Storage Ring Vacuum Vessel and Pumping System Developments 796
6.2.5 Cold-Bore Machines 798
6.2.6 Superconducting RF Accelerators 800
6.2.7 The Next-Generation Big Accelerator? 801
6.2.8 The Magnetic Fusion Road Map 801
6.2.9 The Early History of Magnetic Fusion 803
6.2.10 Model C: The First UHV Fusion Device 804
6.2.11 The Russian Revolution in Fusion: Tokamaks 805
6.2.12 Plasma Impurities and Vacuum Technology 806
6.2.13 Toward the Breakeven Demonstrations 808
6.2.14 The Next Step in Fusion 810
Acknowledgments 810
References 812

Preface

The *Handbook of Vacuum Technology* was conceived in 1994 as an addition to Academic Press's *Methods of Experimental Physics,* Vol. 14: *Vacuum Physics and Technology,* edited by G. L. Weissler and R. W. Carlson (1979). It was decided that rather than add to the series *Methods of Experimental Physics*, a new book should be produced. This handbook covers areas of vacuum technology not covered in detail in Volume 14 or those where the technology has changed significantly. A significant number of world-known authors have written chapters for this handbook, where the chapters do somewhat overlap with the original *Methods of Experimental Physics,* Volume 14, to provide continuity. Readers are referred to *Methods of Experimental Physics,* Volume 14, for chapters covering concepts such as basic vacuum equations, molecular transport, details on welding, soldering, glass systems (generally not employed in modern vacuum technology), protective devices, and many other significant older technology bases. We emphasize the importance of Volume 14 of the *Methods of Experimental Physics* as a basis for this introduction to vacuum technology.

This book has five parts: (1) an introduction; (2) creation of vacuum (pumps and the technology presently used in their operation and design); (3) vacuum measurements (pressure, partial pressure, gas flow, leak detection, calibration and associated technology); (4) system components and design, including construction materials, valves, flanges, operation and maintenance, and system cleaning and cleanliness and finally; (5) applications of high- and ultra-high-vacuum technology. The information included in this handbook is directed toward the practitioner of vacuum technology and includes many references for detailed information in specific areas of interest.

Part 1 is a brief overview of what *Methods of Experimental Physics,* Volume 14, presented in detail. It is intended to give the reader details and references to past literature. Readers should see Volume 14 and other cited references for fundamental information on the basic laws employed in vacuum technology as it is practiced today.

Vacuum technology has progressed in many different ways since its origin. Today, as in many other areas, vacuum technology has been driven toward automation of equipment and pumps. System design is automated through powerful CAD programs available from commercial suppliers. Although very important, software tends to have a short lifetime and changes rapidly. Consequently, the use of these programs and their application is beyond the scope of this text. Readers who need up-to-date information on this topic should contact suppliers of technical programs or use the Internet facilities that many software houses provide free of charge.

Part 2 includes a user-oriented overview and complete discussion of modern pumps used in the practice of vacuum. The discussions are written by experts in the area of their specific pump technology. Included in this chapter are discussions of the following: mechanical pumps (a continuation from *Methods of Experimental Physics,* Volume 14), diaphragm pumps, blowers, diffusion pumps, cryogenic pumps, turbomolecular pumps, UHV pumps (sputter ion, sorption, getter, including nonevaporable getter pumps). In addition to a description of pump operation, we emphasize maintenance and areas of application. Future trends in vacuum pump design are also presented.

Methods of monitoring vacuum processes have evolved over the past decade. Part 3 of this handbook addresses this topic in detail. New concepts have changed vacuum pressure measurement apparatus. In addition to covering the older pressure measurement devices, this handbook covers in detail the use and maintenance of rotating disk and capacitance manometers. The "art" of leak detection and partial pressure analysis has become highly automated. Manufacturers of leak detection equipment have included online data logging, improved sensitivity, and automatic calibration of leak rates into the 10^{-11} torr-liter/sec range. Partial pressure analysis uses the now standard quadrupole-based mass spectrometer. Automated instrumentation has interfaces that generally are acceptable to high-speed desktop computers. We present applications of partial pressure analysis and leak detection to provide the novice with a useful procedure for applying these methods to current vacuum systems. Similar features are discussed for mass flow (gas flow) measurements in Part 3.

With the accessibility of high-speed desktop computers, local control and monitoring of various parameters encountered in vacuum technology have become an integral part of modern vacuum systems. Using various monitors, dynamic real time process control and automatic logging of process performance are provided. These new attributes make vacuum production and processing routine, and provide a paper or digital log of the process parameters for later examination.

Part 4 discusses components used in modern vacuum system construction and system design with the user in mind. With the advent of powerful high-speed desktop computation available at low cost, many new computational programs have been developed based on either the classical analog approach to mathemati-

cal solutions or matrix manipulation approaches. These techniques allow rapid system design, preconstruction understanding of actual operation, and methods of automation in construction. Engineering CAD/CAM methods are regularly used in producing vacuum components and systems.

Many components used in earlier vacuum systems have been refined with better materials. Components tend to be automated through the use of electrical/pneumatic controls. New system design materials, including aluminum, are covered in detail. Vacuum system operation, maintenance and the problems of cleanliness and contamination are discussed for users. Pump selection and integrated system design conclude this section, oriented to a "systems approach" to designing integrated vacuum systems in processing materials.

Part 5 is dedicated to applications of high-vacuum and ultra-high-vacuum in commercial products. The main applications of vacuum have been to deposit and etch thin films of materials used in many areas of modern technology, including the semiconductor industry (silicon wafer fabrication, high-density interconnections, etc.), large-area deposition and web coating, and many other areas. To keep the focus of this volume on vacuum technology, we have limited the number of applications covered. These include sputtering, plasma etching, laser ablation, chemical vapor and plasma enhanced chemical vapor deposition, and surface analytical applications of ultra-high-vacuum apparatus.

Ultra-high-vacuum systems have been employed in both ultra-clean film deposition and surface analysis. The use of UHV in analytical apparatus follows from the requirements of keeping the surface clean during a typical measurement. This is true in most analytical approaches, because they are oriented toward the investigation of the surface or near surface structure, chemistry and composition. High vacuum is required in applications requiring particle (and photon) mean free paths great enough to be detected by *in situ* detectors. These techniques include X-ray- or electron-probe-based equipment (for example, x-ray photoelectron spectroscopy and auger electron spectroscopy). These methods are reviewed in some detail, as they are practiced today and include many references to further reading. Unique applications include a section on large-scale vacuum based processes.

This handbook, along with the earlier text *Methods of Experimental Physics,* Vol. 14, *Vacuum Physics and Technology*, provides a very complete and up-to-date series on the generation of vacuum, vacuum fixturing, vacuum measurements, system maintenance and operation, and vacuum applications as it is applied today. Our orientation has been on the practical use of vacuum technology. We are sure that as a user of vacuum technology, you will find this handbook invaluable.

This Page Intentionally Left Blank

List of Contributors

Numbers in parentheses indicate the pages on which the author's contributions begin.

GARY S. ASH (149), CTI-Cryogenics Division, Helix Technology Corporation, Mansfield, Massachusetts 02048

WILLIAM H. BAYLES, JR. (257), The Televac Division, The Fredericks Company, Inc., Huntingdon Valley, Pennsylvania 19006

HOWARD M. BRADY (257), Electron Technology Division, The Fredericks Company, Inc., Huntingdon Valley, Pennsylvania 19006

JEFFREY T. CHEUNG (694), Rockwell Science Center, Thousand Oaks, California 91360

BENJAMIN B. DAYTON (484), Consultant, East Flat Rock, North Carolina 28726

EMIL DRUBETSKY (257), The Televac Division, The Fredericks Company, Inc., Huntingdon Valley, Pennsylvania 19006

H. F. DYLLA (789), Jefferson Laboratory, Newport News, Virginia 23606 and Department of Physics, College of William and Mary, Williamsburg, Virginia 23185

F. J. ECKLE (84), VACUUBRAND GMBH + CO, D-97877, Wertheim, Germany

JAMES L. GARNER (509), SMC Corporation, Turtin, California 92780

RON GOEHNER (257), Electron Technology Division, The Fredericks Company, Inc., Huntingdon Valley, Pennsylvania 19006

MARSBED HABLANIAN (59, 116), (Retired) Varian Associates, Wellesley, Massachusetts 02181

WALTER HELGELAND (433), Rigaku/USA, Inc., Danvers, Massachusetts 01923

HINRICH HENNING (183), Leybold Vakuum GmbH, D-50968, Cologne, Germany

L. D. HINKLE (376), MKS Instruments, Inc., Andover, Massachusetts 01810

DOROTHY HOFFMAN (deceased)

FRANK JANSEN (711), BOC Coating Technology, 4020 Pike Lane, Concord, California 94524

CHUCK KRAFT (444), Larson Electronic Glass, Redwood City, California 94064
LASZLO V. LIESZKOVSZKY (290), GE Lighting, Cleveland, Ohio 44112
DONALD M. MATTOX (553), Management Plus, Inc., Albuquerque, New Mexico 87122
R. A. OUTLAW (335), Teledyne Brown Engineering: Hastings Instruments, Hampton, Virginia 23669
JANDA K. PANITZ (446), Sandia National Laboratories, Albuquerque, New Mexico 87185
V. PATEL, Ph.D. (628), Sarnoff Corporation, Princeton, New Jersey 08543
NEIL T. PEACOCK (391, 409), HPS Division of MKS Instruments, Inc., Boulder, Colorado 80301
R. N. PEACOCK (463), Vice President for Development (Retired), HPS Division of MKS Instruments, Inc., Boulder, Colorado 80301
MICHAEL J. POWERS (672), Commonwealth Scientific Corporation, Alexandria, Virginia 22314
JAY RICHMAN (97), Consultant, Stokes Vacuum, Inc., King of Prussia, Pennsylvania 19406
WILLIAM B. ROBBINS (761), 3M Corporate Research Laboratories, Thin Film Technology Resources, St. Paul, Minnesota 55144
STEPHEN M. ROSSNAGEL (609), I.B.M., T. J. Watson Research Center, Yorktown Heights, New York 10598
BAWA SINGH (1), David Sarnoff Research Center, Princeton, New Jersey 08543
JACK H. SINGLETON (214), Consultant, Monroeville, Pennsylvania 15146
JOHN H. THOMAS, III (1, 731), 3M Research Laboratories, St. Paul, Minnesota 55144

HANDBOOK OF VACUUM SCIENCE AND TECHNOLOGY

This Page Intentionally Left Blank

PART 1

Fundamentals of Vacuum Technology and Surface Physics

Bawa Singh
David Sarnoff Research Center

John H. Thomas, III
3M Research Laboratories

1.1 Vacuum Nomenclature and Definitions 3
1.1.1 Basic Definition 3
1.1.2 Pressure Regions of Vacuum 3
1.2 Gas Properties 8
1.2.1 Description of Vacuum as the Characteristics of a Low-Pressure Gas 8
1.2.2 Characteristics of a Gas — Basic Definitions 8
1.2.3 Gas Laws 9
1.3 Molecular Processes and Kinetic Theory 11
1.3.1 General Description 11
1.3.2 Molecular Motion 12
1.3.3 Kinetic Theory Derivation of the Gas Laws 14
1.3.4 Pressure 15
1.3.5 Molecular Mean Free Path 17
1.3.6 Number of Impacts with the Chamber Wall 19

1.3.7 Time to Form a Monolayer 20
1.3.8 Thermal Transpiration 20
1.3.9 Coefficient of Thermal Conductivity 21
1.3.10 Coefficient of Diffusion 21
1.4 Throughput, Pumping Speed, Evacuation Rate, Outgassing Rate, and Leak Rate 22
1.5 Gas Flow 25
1.5.1 Nature of Gas Flow 25
1.5.2 Turbulent Flow 27
1.5.3 Viscous, Streamline, or Laminar Flow 28
1.5.4 Molecular Flow 29
1.5.5 Flow Relationships 29
1.6 Conductance 32
1.6.1 Conductance 32
1.6.2 Conductances in Parallel 33
1.6.3 Conductances in Series 33
1.7 Flow Calculations 35
1.7.1 Equations for Viscous Flow 35
1.7.2 Equations for Molecular Flow 37
1.7.3 Knudsen's Formulation 37
1.7.4 Clausing Factors 38
1.8 Surface Physics and Its Relation to Vacuum Science 40
1.8.1 Physical Adsorption or "Adsorption" 40
1.8.2 Chemisorption 42
1.8.3 Sticking Coefficient 43
1.8.4 Surface Area 44
1.8.5 Surface Adsorption Isotherms 45
1.8.6 Capillary Action 47
1.8.7 Condensation 48
1.8.8 Desorption Phenomena 49
1.8.9 Thermal Desorption 50
1.8.10 Photoactivation 52
1.8.11 Ultrasonic Desorption 53
1.8.12 Electron- and Ion-Stimulated Desorption 53
1.8.13 Gas Release from Surfaces 54
References 55

CHAPTER 1.1

Vacuum Nomenclature and Definitions

1.1.1

BASIC DEFINITION

The term *vacuum* is generally used to denote a volume or region of space in which the pressure is significantly less than 760 torr. In the traditional measurement system, normal pressure is expressed in millimeters of a column of mercury, and 760 millimeters of mercury is equal to 1 standard atmosphere. The traditional unit of pressure is the torr, which approximately equals 1 millimeter of mercury. A perfect or absolute vacuum, which implies a space that is entirely devoid of matter, is practically unrealizable. For practical purposes, however, and in accordance with the definition proposed by the American Vacuum Society, the term *vacuum* is generally used to denote a space filled with a gas at less than atmospheric pressure [1].

In the metric, or meter-kilogram-second (MKS), system, the unit of pressure is the pascal. In general, however, the torr still remains one of the most widely used units of pressure. Table 1 lists conversion factors between some of the most generally used vacuum units.

1.1.2

PRESSURE REGIONS OF VACUUM

Measuring a system's pressure is the traditional way to classify the degree of vacuum. Nowadays, the general term *vacuum* refers to a region that consists of about

ISBN 0-12-325065-7
$25.00

Table 1

Conversion Factors for Pressure Units

Unit	Torr	Pascal	Dyne cm^{-2}	Bar	Atmosphere (standard)
1 torr (0°C)	1	1.333×10^{2}	1.333×10^{3}	1.333×10^{-3}	1.3158×10^{-3}
1 pascal (newton m^{-2})	7.5006×10^{-3}	1	10	1.0×10^{-5}	9.8692×10^{-6}
1 dyne cm^{-2}	7.5006×10^{-4}	0.1	1	1	9.8692×10^{-7}
1 bar	7.5006×10^{2}	1.0×10^{5}	1.0×10^{6}	1.0×10^{6}	0.98692
1 atmosphere (standard)	760	1.0133×10^{5}	1.0133×10^{6}	1.0133	1
1 pound (force) $inch^{-2}$	5.1715×10^{1}	6.8948×10^{3}	6.8948×10^{3}	6.8948×10^{-2}	6.8047×10^{-2}

19 orders of magnitude of pressure below 1 atmosphere. For convenience, this extended pressure range is generally divided into several regions that denote the *degree* of vacuum. This division of the pressure scale below atmosphere is somewhat arbitrary and is a convenient method of denoting the different physical phenomena that occur within the pressure ranges specified for each category. Many industrial applications of vacuum can be also be classified using these categories. Table 2 shows the accepted categories and the corresponding pressure ranges. This table also lists the type of pump generally used to achieve a specified pressure range, as well the typical vacuum gauge used for measurement.

To discuss the different physical phenomena associated with the various vacuum categories that are indicated in Table 2, it is useful to introduce other concepts and properties that characterize the degree of vacuum, such as *molecular density*, *mean free path*, and *time to form a monolayer*. These terms are defined as follows:

Molecular density — Average number of molecules per unit volume.

Mean free path — Average distance a molecule travels in a gas between two successive collisions with other molecules of the gas.

Time to form a monolayer — Time required for freshly cleaved surface to be covered by a layer of gas of one molecule thickness. This time is given by the ratio between the number of molecules required to form a compact monolayer (about 8×10^{14} molecules/cm^2) and the molecular incidence rate.

Table 2

Applications of Vacuum Techniques

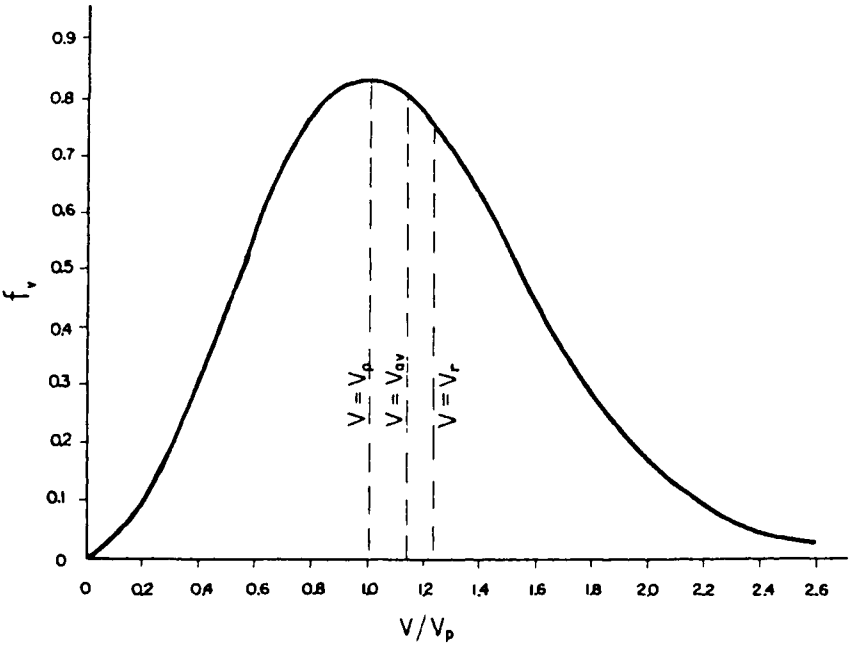
Physical Situation	Objective	Applications
Low pressure	Achieve pressure difference.	Holding, lifting, transport pneumatic, cleaners, filtering, forming
Low molecular density	Remove active atmospheric constituents.	Lamps (incandescent, fluorescent, electric discharge tubes), melting, sintering, packaging, encapsulation, leak detection
	Remove occluded or dissolved gas.	Drying, dehydration, concentration, freeze-drying, degassing, lyophylization, impregnation
	Decrease energy transfer.	Thermal insulation, electrical insulation, vacuum microbalance, space simulation
Large mean free path	Avoid collision.	Electron tubes, cathode ray tubes, television tubes, photocells, photomultiplier, X-ray tubes, accelerators, storage rings, mass spectrometers, isotope separators, electron microscopes, electron beam welding, heating, coating (evaporation, sputtering), molecular distillation
Long monolayer formation time	Clean surfaces.	Fraction, adhesion, emission studies, materials testing for space

Figure 1 shows the relationships among these various quantities as a function of pressure. Using the preceding definitions, it is possible to describe the different physical situations that characterize the different vacuum categories.

1.1.2.1 Low and Medium Vacuum

In the low-vacuum and medium-vacuum range, the number of molecules in a vacuum vessel in the gas phase are large compared to those covering the surface of the vessel. Thus the pumping of the space serves to remove molecules from the gas phase. This vacuum range extends from atmosphere to about 10^{-2} torr. Many industrial processes that need to outgas or dry materials and components use this region.

Fig. 1.



Maxwell-Boltzmann molecular velocity distribution curve.

1. $f_v = \frac{1}{n}\frac{dn}{dv} = \frac{4}{\pi^{1/2}}\left(\frac{m}{2kT}\right)^{3/2} v^2 \exp\left(-\frac{mv^2}{2kT}\right)$

2. $v_{max} = v_p = \frac{2kT}{m}$

3. $v_{mean} = v_{average} = \bar{v} = \sqrt{\frac{8kT}{\pi m}} = 1.13\, v_{max}$

4. Root mean square $v_{rms} = \sqrt{\frac{3kT}{m}} = 1.225\, v_p$

5. Mean energy $\bar{e} = \frac{3}{2}kT$

1.1.2.2 High Vacuum

The high-vacuum region corresponds to a state where the gas molecules are located mainly on the surfaces of the enclosure and the mean free path equals or exceeds the dimensions of the vacuum vessel. The particles travel in the vacuum enclosure without colliding with other molecules. Thus, in this vacuum region, under these conditions, the pumping consists of evacuating or capturing mole-

cules. The molecules leave the surface and individually reach the pump. This region is extensively used in the preparation and application of vacuum coatings, surface treatment, and modification. This region extends from 10^{-3} to 10^{-7} torr.

1.1.2.3 Ultra-High Vacuum (UHV)

Under ultra-high-vacuum (UHV) conditions, time to form a monolayer is equal to or longer than the usual time for most laboratory measurements. Thus clean surfaces can be prepared and their properties determined before an adsorbed gas layer is formed. This vacuum range extends from about 10^{-7} to 10^{-16} torr.

An indication of the extensive application of vacuum technology in many key industrial processes in a diverse range of industries is illustrated in Table 2, where common vacuum industrial processes are classified according to the degree of vacuum used.

CHAPTER 1.2

Gas Properties

1.2.1

DESCRIPTION OF VACUUM AS THE CHARACTERISTICS OF A LOW-PRESSURE GAS

The behavior and characteristics of gases are fundamental to vacuum systems. Even at the extremely low pressures typically encountered in vacuum technology, gases essentially still behave as gases. The necessity for creating a vacuum is usually related to the need to reduce the number density of gaseous molecules, or their surface collision rates. Behavior of gases in vacuum systems can be generally discussed in terms of the ideal gas laws; and some aspects of overall behavior of vacuum systems can be described by the static and dynamic properties of gases. The behavior of gases at low pressure and various aspects of gas flow are considered next.

1.2.2

CHARACTERISTICS OF A GAS—BASIC DEFINITIONS

A low-pressure sample of gas can be completely described if at least three of the four quantities that relate to it are known. These quantities are its *pressure*, *volume*, *temperature*, and the *amount* of gas in the sample.

Pressure: Pressure is defined as the force per unit area that a gas exerts on the walls of its container. In the MKS system of units, pressure is expressed as newton per square meter, or newton/m^2 or newton m^{-2}. The MKS unit is the pascal, where 1 torr = 133 pascal and 1 pascal = 7.5 torr.

Volume: The volume is simply a measure of the space a gas takes up; it is usually set by the dimensions of the enclosure. The MKS unit of volume is the m^3 but liters are extensively employed to refer to pumping rates, gas flow measure-

ISBN 0-12-325065-7
$25.00

ments, etc. The pumping speed of mechanical pumps is often expressed in cfm (cubic feet per minute).

Temperature: The temperature of a gas at pressure below 1 torr is determined mainly by the temperature of the surfaces with which the gas comes in contact. Typically the gas is at room temperature. In deriving the equations that describe the behavior of gases, the unit of temperature is K or Kelvin.

Amount of gas: The amount or mass of gas in a given sample is measured in molar gram units or moles.

Gram molecule or mole: Defined as that quantity of gas (or any substance) having a mass equal to its molecular weight in grams. A gram molecule contains 6×10^{23} molecules. One mole of any gas at 0°C and a pressure of 760 torr, occupies 22.4 liters of volume. The mass of 1 mole of gas is exactly equivalent to its molecular weight in grams.

Gram molecular volume: The volume occupied by a gram molecule of gas is a universal constant; it is found experimentally to be 22.414 liters at 760 torr and 0°C. As 1 mole of any gas, at a temperature of 0°C and a pressure of 760 torr occupies a volume of 22.4 liters, it is possible from this relationship to calculate the molecular density of any volume of gas if its temperature and pressure are known. For example, 1 cubic centimeter of air at 760 torr and 0°C contains 2.7×10^{19} molecules; whereas at a pressure of 1 torr and a temperature of 0°C, 1 cubic centimeter of air contains 3.54×10^{16} molecules.

1.2.3 GAS LAWS

The gas laws (Boyle's, Charles's, Gay-Lussac's) lead to relationships of the bulk physical quantities of the gas, such as pressure, volume, temperature, and the amount of gas to one another. These relationships describe the behavior of a given quantity of an "ideal" gas; an ideal gas is one where the volume of all the molecules is negligible compared to the volume of the gas, and the energy of attraction between the molecules is negligible compared to their mean thermal energy. This means that the sample of gas is dilute and is at a temperature that is not low enough to condense it. Gases that are ideal at room temperature include O_2, Ne, Ar, CO, H_2, N_2, and NO. A summary of the relationships that result from applications of the gas laws to an ideal gas is provided here:

Boyle's law: States that the product of pressure and volume, pV, is constant for a given mass of gas at constant temperature.

Charles's law: States that V/T is constant for a given mass of gas at a constant pressure, where V is the gas volume and T = the absolute temperature.

Avogadro's law: States that equal volumes of gas at any gas at the same tem-

perature and pressure contain the same number of molecules. From this law can be obtained an important relationship between the number of moles in a sample and the pressure the gas exerts.

General gas law: The general gas law relates all four quantities needed to describe the state of a gas. The general law states that

$$PV = nRT \tag{1}$$

for a given mass of gas, where R = universal gas constant (constant of proportionality) with a value of 62.4 torr-liter/mole°K, and n is the number of moles in volume V.

This law is known as the "ideal" gas law, because it is exactly true for ideal gases; most gases at reduced pressures behave as ideal gases.

Dalton's law of partial pressures: The total pressure exerted by a mixture of gases is equal to the sum of the partial pressures exerted by the individual components.

Partial pressure: The partial pressure exerted by any one component of a mixture of gases is the pressure exerted by that component if it occupied that volume alone.

Avogadro's law: Equal volumes of all ideal gases measured at the same temperature and pressure contain the same number of molecules.

Avogadro's number: The number of molecules in a gram molecule of gas or any substance is a universal constant and is 6.023×10^{23}.

Loschmidt's number: The number of molecules per cm^3 of gas at 760 torr and 0°C is a universal constant equal to 2.637×10^{19}.

For 1 mole at standard temperature and pressure (STP), P = 760 torr = 1,013,250 dynes/cm^2, V = 22.414 liters, and T = 273.2°K, whence R = 8.31×10^7 ergs per gram molecule or in thermal units R/J = 1.99 cal per °K (J = mechanical equivalent of heat = 4.182 joules cal^{-1}). In more tangible terms, therefore, 1.99 cal will raise the temperature of 1 mole of any ideal gas 1°K. Alternatively, having raised the temperature of 1 mole of any ideal gas by 1°K, the increase in energy of the gas amounts to 8.31 joules.

1.2.3.1 Nonideal Gases

Examples of some common nonideal gases are ammonia, ethane, benzene, CO_2, mercury vapor, NOSob, SO, and SO_2. The gas laws for any gas have to account for behavior of a gas at low temperature. Below a certain temperature, called the *critical temperature*, T_C, the gas begins to condense. Below this critical temperature, there is a vapor pressure of gas over the liquid condensate, called the *vapor pressure*. If the gas is condensed (V is decreased), the pressure will not change, but more gas will condense into the liquid phase. As the temperature is lowered, fewer molecules are present over the liquid and the vapor pressure is lower.

CHAPTER 1.3

Molecular Processes and Kinetic Theory

1.3.1

GENERAL DESCRIPTION

The gas laws describe the bulk behavior of gases but provide no insight as to why gases should behave in this fashion. Kinetic theory is an attempt to explain the behavior of gases in terms of the behavior of the individual molecules that make up the gas. The results of kinetic theory are in close agreement with the gas laws. The theory is based on the following assumptions:

1. A gas is composed of a large number of molecules. All the molecules of a given chemical substance are exactly alike.
2. The molecules are separated by distances that are large in comparison with their own dimensions.
3. The molecules are in constant state of random and chaotic motion. This motion is related to the temperature of the gas.
4. The molecules exert no force on each other or on the walls of their container except when they collide. The volume occupied by the molecules is negligible compared with the volume occupied by the gas.
5. The molecules behave as perfect elastic spheres.

These assumptions are clearly idealized, and no known gas behaves exactly in accordance with this set of assumptions. However, the theory based on these assumptions explains the behavior of real gases in very satisfactory manner.

The theory describes the motion of molecules that are in constant motion. The molecules are free to wander throughout any space available to them. The temper-

ISBN 0-12-325065-7
$25.00

ature of a gas is a measure of the kinetic energy of the particles. The pressure on the walls of a vessel containing a gas is caused by the impact of the gas molecules. If some of the gas is removed, the remaining molecules expand to fill the entire volume but as there are fewer impacts with the walls, the gas pressure drops. Some additional aspects of gas behavior, important to understanding certain characteristics of vacuum, can be calculated from the mathematical expressions developed from kinetic theory. Some of these mathematical relationships, derived from kinetic theory, are presented next; these relationships indicate the dependence of physical properties and processes on molecular quantities.

1.3.2 MOLECULAR MOTION

Kinetic theory postulates that gas molecules are in constant motion and that their velocities are temperature dependent. The distribution of velocities is given by the Maxwell-Boltzmann distribution, which states,

$$\frac{1}{n} \cdot \frac{dn}{dv} = 4\pi^2 \left(\sqrt{\frac{m}{2kT}} \right)^2 \frac{3}{2kT} \tag{2}$$

where m is the molecular velocity, k is the Boltzmann constant, and v the velocity.

This expression provides a statistical distribution of molecular velocities and states that if there are n molecules in the volume, there will be dn molecules having speeds of v and $v + dv$. Figure 2 shows a number of quantitative relationships that can be derived from the Maxwell-Boltzmann distribution:

- The most probable velocity:

$$v_{max} = \sqrt{\frac{2kT}{m}} \tag{3}$$

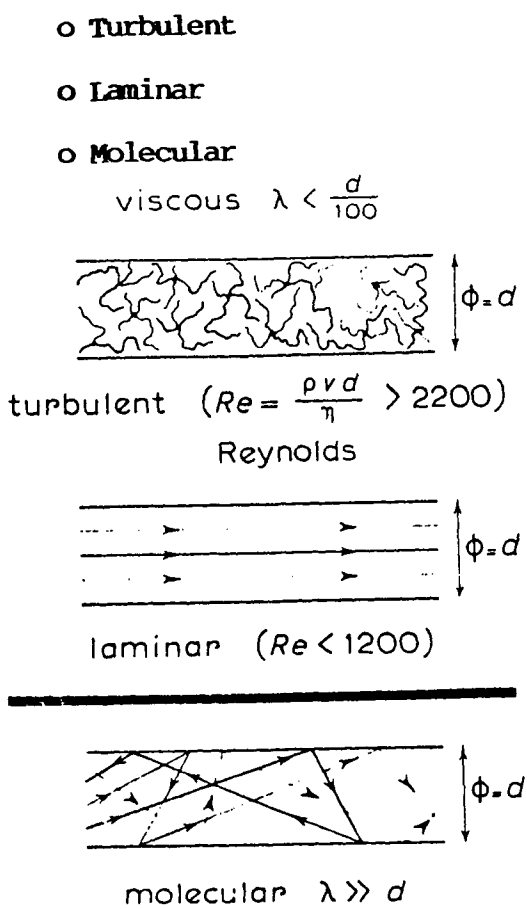
  where m is the molecular mass, k is the Boltzmann constant, and T the absolute temperature.
- The average velocity:

$$v_{av} = \sqrt{\frac{8kT}{\pi m}} = 1.13\, v_{max} \tag{4}$$

- The root mean square velocity v_{rms}:

$$v_{rms} = \sqrt{\frac{3kT}{m}} = 1.22\, v_{max} \tag{5}$$

Fig. 2.



Types of flow through cylinder tubes.

- At high pressures, the mean free path is much smaller than the diameter of the tube ($\lambda < 0.015d$). In this region, viscosity of gas plays an important role and the flow is called *viscous*.
- *Viscous* flow is divided into two regions: turbulent flow and laminar flow.

- Mean energy of a molecule:

$$E = \frac{3}{2}kT \tag{6}$$

The mean velocity is used in calculations of gas flow, and the rms velocity is used when discussing kinetic energy.

1.3.3 KINETIC THEORY DERIVATION OF THE GAS LAWS

1.3.3.1 Boyle's Law

Let δ be the gas density, thus

$$mn = \delta$$

where m is the molecular mass and n is molecular volume density. Hence, the pressure is given by

$$P = \frac{1}{3}\delta = \frac{1}{3}mnv^2 \tag{7}$$

Thus

$$PV = \frac{mv^2}{3}V \tag{8}$$

where m is the total mass of the gas. This relationship shows that at constant temperature, the pressure varies directly with density or conversely with volume, which is Boyle's law.

1.3.3.2 Charles and the Ideal Gas Law

The average energy of molecule is given by

$$E_{av} = \frac{mv^2}{2} \tag{9}$$

and therefore

$$P = \frac{2nE_{av}}{3} \tag{10}$$

That is, the pressure is equal to two-thirds of the total kinetic energy per unit volume. From the expression for v_{rms}, we have

$$v_{rms}{}^2 = \frac{3kT}{m} \tag{11}$$

Thus

$$P = nkT \tag{12}$$

where n is the number of molecules per unit volume. Substituting for $n = N/V$ we get

$$P = \frac{NkT}{V} \tag{13}$$

or

$$\frac{PV}{T} = N\mathrm{k} \tag{14}$$

which is the ideal gas law.

Consider equal volumes of any two different gases at same values of P and T. Since P and V are same for each gas, and $\frac{mv^2}{2}$ are the same at the same temperature, it follows that n must be the same for each gas. Thus we have "equal volumes of all gases at any given values of temperature and pressure contain equal number of molecules." This is, of course, Avogadro's hypothesis, which was stated around 1811. Introducing Avogadro's number, $N_A = 6.023 \times 10^{23}$/mole as

$$N_A = M/m \quad \text{and} \quad n = NA/V_m \tag{15}$$

where m is the molar mass and V_m is the molar volume.

$$PV_m = N_A RT \tag{16}$$

or

$$PV_m = \mathrm{R}T \tag{17}$$

where R is the universal gas constant, and T is temperature in °Kelvin.

1.3.4 PRESSURE

The following is a simplified derivation of the pressure exerted by a gas on the walls of its container as a result of molecular impacts. Since the molecular impacts are perfectly elastic, energy and momentum are conserved and each molecule is thrown back from the wall with the same speed as its incident speed.

Consider a molecule of mass m that approaches the enclosure wall with a velocity v and rebounds with the same velocity, thus

$$\text{Change in momentum} = 2mv$$

If n molecules strike a unit area of the wall in unit time, with average velocity v, then

Total impulse exerted on unit area of wall in unit time $= 2mnv$
Pressure = rate of change of momentum per unit area

Thus

$$P = 2mnv \tag{18}$$

To estimate P, we need to calculate v, the frequency of molecular impact of the unit area of the enclosure surface.

Let n = number of molecules in cube of unit volume. Each molecule can move in six possible directions. The velocity of each molecule is v. On average, $\frac{1}{6}$ of the molecules will move toward each face of the cube. Thus, for a gas with volume density n,

$(n/6)v$ molecules will cross a unit area in unit time

Thus

$$v = \frac{nv}{6} \tag{19}$$

Hence

$$P = 2mvn = 2mv \cdot \frac{nv}{6} \tag{20}$$

$$= \frac{1}{3}(mnv^2)$$

Thus

$$\text{Pressure} = P = \frac{1}{3} Nm\,(\bar{C})^2 \tag{21}$$

where P = pressure in MKS units (newton m^{-2})
N = number of molecules per unit volume
m = mass of molecule (g)
$(\bar{C})^2$ = mean square velocity (cm^2/s^{-2})

The average value of the kinetic energy of translation of molecules ($\frac{1}{2}mC^2$) is a direct measure of the absolute temperature of a gas, and therefore the average kinetic energy of molecules of all gases is the same at the same temperature.

1.3.5

MOLECULAR MEAN FREE PATH

The average distance a molecule moves before colliding with another (collisions with chamber walls being excluded) is called the *mean free path* (*mfp*) and is given by

$$\lambda = \frac{1}{\sqrt{2}} \cdot \frac{1}{(\pi n d_0^2)} = \frac{kT}{\sqrt{2}\pi d_0^2 P} \tag{22}$$

where d_0 = molecular diameter ($\sim 10^{-8}$ cm)
n = number of molecules per unit volume

Example:

At 760 torr, mfp for air at 20°C = 6.4×10^{-6} cm
At 10^{-6} torr, mfp for air at 20°C = 49 m

Derivation: Consider a molecule of diameter d_0, velocity v, moving in a gas of density n. The molecule moves distance vdt in time dt so the molecule will collide if there is another molecule of the same diameter within the volume:

$$\delta V = \pi d_0^2 v dt \tag{23}$$

Since n = molecules/cm^3, the volume associated with 1 molecule is $1/n$/cm^3.

Thus when δv becomes equal to $1/n$, a collision has occurred:

$$\frac{1}{n} = \pi d_0^2 v \tau_c \tag{24}$$

where τ_c is the average time between collisions.

The mean free path = distance traveled:

$$\lambda = v\tau_c = \frac{1}{\pi n d_0^2} \tag{25}$$

If we take into account the molecular velocity distribution, we get a more correct relation:

$$\lambda = \frac{1}{\sqrt{2}\pi n d_0^2} \tag{26}$$

Hence the *mfp* differs for various gases and is inversely proportional to the pressure. The mean free path must not be confused with the average distance be-

Table 3

An Overview of Important Points for Different Pressure Ranges

Pressure Range (torr) 10^2 1 100^{-2} 10^{-4} 10^{-6} 10^{-8} 10^{-10}	Type of Vacuum	Mean Free Path[a] (cm)	Number[a] of molecules per cm^3	Type of Pump	Type of Pressure Gauge
⟷	Rough or forevacuum	$\sim 10^{-5} - 10^{-1}$	$\sim 10^{19} - 10^{15}$	Mechanical oil-sealed, steam ejector, or sorption pumps	Liquid-filled U-tube manometer, membrane compression (spiral) gauge
⟷	Intermediate or booster vacuum	$\sim 10^{-2} - 10^{1}$	$\sim 10^{16} - 10^{13}$	Oil-ejector or oil-booster pumps	Thermocouple or alphatron gauges
⟷	High vacuum	1 cm to larger than vacuum vessel	$\sim 10^{-14} - 10^{10}$	Oil or Hg diffusion pumps	Ordinary ionization gauge
⟷	Ultra-high vacuum	Much larger than vacuum vessel	Less than 10^{10}	Ion, Ti-getter, cryo-, or rotary Roots-type pumps	Bayard-Alpert-type ion gauge

Note: Saturated water vapor pressures: at liquid N_2 (−190°C), less than 10^{-6} torr; at dry ice temperature (−78°C), about 10^{-3} torr; between +10 and +30°C, about 10–30 torr. Saturated Hg vapor pressures, at liquid N_2 or dry ice temperature, less than 10^{-6} torr; at or near room temperature, about 10^{-3} torr. Saturated CO_2 vapor.

[a]From simple kinetic theory.

tween molecules, which remains very small even at high vacuum, whereas the mean free path becomes very long due to the small size of molecules and the resulting unlikelihood of a collision occurring.

1.3.6 NUMBER OF IMPACTS WITH THE CHAMBER WALL

The impact frequency is a quantity of considerable interest in vacuum technology is the number of molecules that strike a surface in unit time. From kinetic theory, it can be shown that rate of bombardment of each unit surface area is

$$\Gamma = \frac{nv_{av}}{4} = \frac{n}{2}\left(\frac{2kT}{\pi m}\right)^{1/2} \tag{27}$$

The volume density is given by

$$n = 9.656 \times 10^{18}\frac{P}{T} \tag{28}$$

and

$$v_{av} = 1.45 \times 10^{4}\left(\frac{T}{M}\right)^{1/2} \tag{29}$$

thus

$$\Gamma = 3.5 \times 10^{22}\, P/(MT)^{1/2} \text{ molecules}^{-1}/\text{cm}^{-2} \tag{30}$$

where P is in torr
M is in g
T is in °K
P is pressure in torr
M is molecular weight

At a pressure of 760 torr and 20°C, for nitrogen molecules there are 2.93×10^{23} impacts s^{-1}/cm^{-2}. At a pressure of 10^{-6} torr and 20°C, for nitrogen molecules there are 3.87×10^{14} impacts s^{-1}/cm^{-2}.

1.3.7

TIME TO FORM A MONOLAYER

The number of molecular impacts per second is a difficult quantity to comprehend. It can be expressed in a more useful way as the time required for the surface to become covered by a single layer of molecules, as for instance in surface physics studies where a clean surface is prepared under high vacuum.

At a pressure of 760 torr and 20°C, the time to form a monolayer of nitrogen molecules is 3×10^{-9} seconds. At a pressure of 10^{-9} torr and 20°C, the time to form a monolayer of nitrogen molecules is about 2000 seconds. These values have been calculated assuming an accommodation coefficient of 1, that is, assuming that an incident molecule sticks on first impact.

1.3.8

THERMAL TRANSPIRATION

For two chambers maintained at T_1 and T_2 °K and separated by a porous partition, equilibrium chamber pressures are established such that

$$\frac{P_1}{P_2} = \sqrt{\frac{T_1}{T_2}} \tag{31}$$

Thermal transpiration is a process that occurs at all pressures, the porous partition being necessary to ensure that at one part of the system the mean free path of the gas is large compared with system dimensions, in this case the dimensions of the pores of the porous material. In high-vacuum systems, this condition exists without the presence of the porous partition, and thermal transpiration requires that the pressure (P_1) registered by a gauge at room temperature (T_1) be corrected to obtain the pressure (P_2) existing at points maintained at a different temperature (T_2).

Coefficient of viscosity: If η = coefficient of viscosity, then

$$\eta \propto m\bar{v} \tag{32}$$

where m = mass of molecule (g)
$\bar{v}$ = arithmetic average molecular velocity = $14{,}551(T/M)^{1/2}$ cm s^{-1}

Hence

$$\eta \propto T^{1/2}$$

It is surprising that the η is independent of pressure. This is verified by experiment at normal pressures where molecular collisions are frequent, that is, where

the assumption of the kinetic theory holds good. At low pressures, less than 10^{-3} torr, when the mean free path is 5 cm or more the molecules can travel directly from a moving surface to the walls of the vacuum chamber. Thus the momentum exchange between a moving surface and the walls or the drag on the moving surface depends on the number of molecules present (the pressure), and viscosity gauges are constructed to measure pressures below 10^{-3} torr.

1.3.9 COEFFICIENT OF THERMAL CONDUCTIVITY

Thermal conductivity is the transfer of thermal energy, whereas viscosity is the transfer of mechanical energy through a gas. Kinetic theory gives an expression for the coefficient of thermal conductivity C_T similar to that for viscosity. In fact,

$$C_T = \eta C_v \tag{33}$$

where C_v = the specific heat of the gas at constant volume and is constant over wide ranges of pressures and temperature. With the same reasoning as applied to viscosity, it is possible to construct a thermal conductivity gauge to measure below 10^{-1} torr. Because of their simple construction, versatile control, and detection circuit techniques, thermal conductivity gauges are widely used to measure pressures below 10 torr.

1.3.10 COEFFICIENT OF DIFFUSION

The rate at which molecules of gas 1 diffuse into the molecules of gas 2 is given by the following formula:

$$\frac{dp_1}{dt} = D\frac{d^2p_1}{dx^2} \tag{34}$$

where D = coefficient of diffusion (cm^2s^{-1}) of gas 1 through gas 2 and depends on the concentrating gradient dp_1/dx. It can be shown that

$$D = \text{constant} \times \frac{T^{3/2}}{p} \cdot \frac{\sqrt{(1/M_1 + 1/M_2)}}{(d_1 + d_2)^2} \tag{35}$$

By putting $M_1 = M_2 = 28$, $d_1 = d_2 = 3.75 \times 10^{-8}$ cm, a factor for the self-diffusion for nitrogen is obtained.

CHAPTER 1.4

Throughput, Pumping Speed, Evacuation Rate, Outgassing Rate, and Leak Rate

The process of evacuation is in effect the removal of mass from the vacuum vessel: Thus the rate of mass removal (the mass flow) determines the rate at which pressure falls. The relationship between the throughput, Q, and the corresponding volumetric flow, S, are related to the gas pressure as follows:

If the vessel initially contains N molecules of mass m_i, the rate of change of the total mass m is given by

$$\frac{dm}{dt} = \frac{d(Nm_i)}{dt} \tag{36}$$

Substituting for N from the equation of state ($PV = NkT$),

$$\frac{dm}{dt} = \frac{d}{dt}\left(\frac{m_i PV}{kT}\right) \tag{37}$$

In practice, the type of gas and the temperature can be considered to be constant during the evacuation, so that

$$\frac{dm}{dt} = \frac{m_i}{kT} \cdot \frac{d(PV)}{dt} \tag{38}$$

ISBN 0-12-325065-7
$25.00

Both the pressure and volume are easily measured quantities in vacuum vessels, so that it is convenient to define a gas flow rate, or throughput, Q, as

$$Q = \frac{d}{dt}(PV)$$

Q can then be measured in practical cases and is related to the mass flow rate by

$$\frac{dm}{d} = \frac{m_i}{kT} \cdot Q \tag{39}$$

provided always that m and T are constant. Q is expressed in units of Pa $m^2 s^{-1}$ and is often referred to as *gas throughput.*

The number of molecules per unit volume is proportional to the pressure and the volumetric flow rate, and is usually defined as the product:

$$Q = SP \tag{40}$$

This equation is one of the fundamental equations of vacuum physics.

Pumping speed: The speed of a pump is defined as the volumetric rate of flow into the pump (Ls^{-1}) where the volume is measured at the pressure at the pump inlet (torr). To be correct, therefore, a speed should be quoted as 500 Ls^{-1} at 10^{-5} torr. It is common practice, however, to refer only to Ls^{-1} (diffusion pumps and sputter-ion pumps) or L min^{-1} (for mechanical pumps). Hence a particular diffusion pump that has a speed of 500 Ls^{-1} at 10^{-5} torr will have the same speed of 500 Ls^{-1} at 10^{-8} torr. In the first case the pump pumps 500 Ls^{-1} of "1000^{-5} torr gas," and in the second case it pumps 500 Ls^{-1} of "10^{-8} torr gas." The amount of gas pumped (the throughput) has fallen by a factor of 10^3, but the pumping speed or volumetric flow rate remains the same.

Leak rates: A leak rate is the throughput of a small hole from atmospheric pressure to the vacuum space, and is conveniently measured in torr-Ls^{-1}. An older unit of leak rate is the liter micron s^{-1}.

Outgassing rates: An outgassing rate of a system is the rate at which gas is desorbed from system walls to be pumped away by the vacuum pump, and is expressed as a throughput. It is dependent on the surface area of the system from which desorption is taking place. The outgassing rate of a surface is therefore expressed as throughput per unit area, or torr-Ls^{-1} cm^{-2}. Typically a clean, unbaked, stainless steel surface has an outgassing rate, after one to two hours pumping, of 10^{-8} to 10^{-9} torr-Ls^{-1} cm^{-2}. Outgassing rates are particularly influenced by surface history such as exposure to moist atmosphere, baking, etc.

Permeability: Construction materials, especially glasses and elastomers, are permeable to the light constituents of the atmosphere. Permeability rates are expressed in the same terms as outgassing rates.

Evacuation rate: The pressure in a vacuum system falls as gas is removed, and the rate of gas removal is the throughput of the vacuum pumps. Assuming the pumping speed to remain constant, the throughput falls as the pressure falls, and is in fact exponential, given by

$$P_2 = P_1 \exp\left(-\frac{St}{V}\right) \tag{41}$$

where P_2 = pressure (torr) at time t (seconds), P_1 = pressure (torr) at time $t = 0$, V = system volume (liters) and S = pumping speed (Ls^{-1}).

Thus the pressure falls to half its original value in 0.69 V/S seconds, to quarter that value in twice as long, to $\frac{1}{8}$ that value in twice as long again, and so on. At low pressures the evacuation rate falls off as a result of leakage and outgassing from system walls. Pumpdown times to pressures below 10^{-3} torr are seldom computed but quoted as a result of experience with similar systems or extrapolation of experimental data on smaller systems.

CHAPTER 1.5

Gas Flow

1.5.1

NATURE OF GAS FLOW

The production and maintenance of vacuum involves flow of gas from the vacuum vessel through pipelines and pumps to the atmosphere. Hence mechanisms that control and affect the nature of gas flow are of great importance. Typically, gas flow in vacuum systems can be divided into three main categories or regimes:

- Turbulent flow
- Viscous or laminar flow
- Molecular flow

Many factors are responsible for determing the gas flow regime, including

- Magnitude of the flow rates
- The pressure drop at either ends of the pipe or duct
- The surface and geometrical properties of the pipe or duct
- The nature and characteristics of the gases being pumped

Quantitative expressions that relate factors such as duct geometry and nature of gas to the type and quality are often complex. However, formulas and expressions for number of commonly encountered geometrical configurations have been developed, and these enable gas flow rate to be calculated with reasonable accuracy and ease. Calculation of gas flow rates under for various geometries and different conditions are presented in the following sections. Detailed derivations of many of the formulas can be found in references. To simplify calculations, the three just-listed flow categories are usually considered independently. All the three regimes of gas flow are encountered when a vacuum system, connected by pipes

ISBN 0-12-325065-7
$25.00

to a vacuum pump, is pumped down from atmosphere to high-vacuum impact system characteristics such as pumpdown time.

1.5.1.1 General Description

At high pressures and very high flow rates, the mean free path of the molecules is very small compared with the dimensions of the pipe or vacuum vessel so that the flow of the gas is limited by the viscosity of the gas. Gas flow under these conditions is called *viscous*, and flow can be either *turbulent* or *laminar*. When the velocity of the gas exceeds certain values, the flow is turbulent, the flowing gas layers are not parallel, and their direction is influenced by any obstacle in the way. In spaces between the layers, spaces of lower pressures (cavities) appear. At lower gas velocities, the viscous flow becomes laminar—that is, the layers of gas flow are parallel, their velocity increasing from the walls toward the axis of the pipe. As the pressure is reduced, the mean free path of the molecules becomes similar to the dimensions of the vessel and flow is governed by a combination of the gas viscosity and by molecular phenomena. These conditions give rise to a type of flow known as *intermediate flow*. As the pressure continues to decrease, the mean free path becomes larger than the vessel dimensions and flow depends only on molecular collisions with the vessel walls; hence the flow under these conditions is referred to as *molecular*. The molecular motion in the various flow regimes is illustrated in Figure 2. The nature of gas flow (whether turbulent, laminar, intermediate, or molecular) is determined by the values of two dimensionless parameters, namely, the Reynolds and Knudsen numbers. The limit between the turbulent and laminar or viscous flow is defined by the Reynolds number, while those between laminar, intermediate, and molecular flow are determined by the Knudsen number.

The Reynolds number is a dimensionless quantity expressed by

$$Re = \rho v D/\eta \tag{42}$$

where ρ is the density of the gas, v is the velocity, η is the gas viscosity, and D is the diameter of the tube.

For $Re > 2100$, flow is entirely turbulent.
For $Re < 1100$, flow is entirely laminar.

In the region $1200 < Re < 2200$, the flow can be turbulent or viscous depending on the roughness of the surface of the tube and geometrical factors of the tube. As flow velocity increases, the Reynolds number increases at high pressures, and the gas, instead of flowing smoothly, develops eddies and oscillations.

Viscous flow occurs in the region bounded by a Reynolds number lower than 1200 and a Knudsen number lower than 0.001.

The Knudsen number, K_n, is also a dimensionless quantity, given by

$$K_n = \frac{\lambda}{D} \tag{43}$$

When the Knudsen number is equal to or greater than the pipe diameter, say $K_n > 1$, the flow properties are determined by gas–wall collisions and the flow is molecular. More detailed descriptions of the three main flow regimes are given next.

1.5.2 TURBULENT FLOW

Turbulent flow of gas occurs at high pressure gradients, is characterized by eddies and vortices in the gas stream, and is rarely encountered in most vacuum applications. It does briefly occur, however, on the onset of pumping where the gas pressure and velocity are sufficiently high and the gas flow in the vacuum vessel and in the connecting pipes is very chaotic; the flow pattern lacks order and is characterized by eddies that appear and disappear. The motion of gas molecules undergoing turbulent flow is complex and lacks order; the gas swirls and eddies, and the individual particles of the gas may have velocities and directions that are quite different from the average velocity and the overall flow direction. During turbulent flow conditions the gas pressure and velocity of flow at any point in the system fluctuates about a mean point. *Except in very special cases (for example, in very large vacuum systems), the duration of turbulent flow is short compared with viscous flow and molecular flow.* In most production thin film coating systems and special "soft pumpdown" procedures are employed to reduce the initial pumping speed, to reduce turbulence in the chamber and thus minimize dust or particle generation and contamination of the substrates. Although it is irregular, the turbulent flow can nonetheless be characterized by the laws of probability [2]. The existence of turbulent flow is determined by the value of the dimensionless Reynolds number. If the Reynolds number is larger than 2100, the flow will always be entirely turbulent; for example, for air at room temperature, flowing through a circular pipe, the flow is turbulent if

$$\frac{PF}{d} > 5 \times 10^5 \tag{44}$$

where F is the flow rate through the pipe in L/sec; P is the mean air pressure in millitorr, and d is pipe diameter in inches.

In many situations, however, turbulent flow can usually be neglected as it gen-

erally occurs only briefly (in the pipe between the mechanical pump and the vacuum system), during the initial stage of pumpdown from atmosphere.

1.5.3

VISCOUS, STREAMLINE, OR LAMINAR FLOW

This type of flow is much simpler than turbulent flow and occurs at moderate pressure gradients when $\lambda < (d/100)$ (where λ is the mean free path and d is the pipe diameter). Viscous flow is smooth and orderly; every particle passing a point follows the same path as the preceding particle passing that point. Flow rates are proportional to the pressure gradient, and the viscosity of the gas. Flow lines are straight lines or gradual curves unlike the case with turbulent flow. That mean free path of the molecules is small compared to the dimensions of the duct during this type of flow, so that collisions are predominant in determining the characteristics of flow and the flow rates will be affected by the viscosity of the gas. Viscous laminar flow often occurs in backing lines to diffusion pumps. Viscous flow conditions are reached as continued pumping reduces the pressure and the Reynolds number decreases below 2200. Eddies cease to appear, and the energy resulting from the pressure gradient is used to maintain a steady flow. The gas velocity and pressure become uniform with time, and the flow becomes streamlined; that is, the lines of flow are smooth and continuous and curve gently in the neighborhood of bends and other irregularities in the pipe. Near the wall the gas is almost at rest, but progressing toward the center the layers of gas slide more quickly over each other until the velocity reaches a maximum at the center. Viscosity is important in determining the amount of gas passing per second through the pipe under these conditions. This type of flow is referred to as either "*viscous*" flow or "*laminar*" flow. It is governed by Poiseuille's law, which states that

$$\frac{Q}{(P_1 - P_2)} = \mathrm{K}\frac{d^4 P}{\eta L} \tag{45}$$

where L = length of the tube of diameter d
Q = product of volume flow and corresponding pressure
η = gas viscosity
K = numerical constant

The pressure gradient causes adjacent layers of the gas to exert a pressure on each other in the direction of the negative pressure gradient, setting the gas in motion as a whole. Viscous flow can only occur when the mean free path is small compared with the pipe diameter. Only those molecules near the pipe walls collide with the walls, and since these will represent only a small fraction of the total number molecules present, the nature of the walls does not have an important effect on the flow rate.

1.5.4 MOLECULAR FLOW

If the gas pressure is lowered even further (while still maintaining a pressure gradient along the pipe), the mean free path of the gas molecules increases and approaches the pipe diameter and the nature of the flow changes. The laminar nature of the flow disappears because molecules now collide with walls of the pipe rather than with each other. When the pressure is low enough, the molecules move about inside the pipe independently of each other and the flow is said to be "molecular." The pressure along the tube no longer acts as a driving force pushing the gas along the tube in a stream. At these low pressures the molecules move in random directions, and there is net transfer of gas from the high- to the low-pressure region simply because there are more molecules per unit volume.

Molecular flow is characterized by molecular collisions with the tube walls rather than with other gas molecules. Flow rates are proportional to the difference in pressures across the component and the reciprocal of the square root of the molecular weight of the gas. The dependence of flow rate on viscosity begins to decrease, because intermolecular collisions are less frequent. At pressures sufficiently low for the mean free path to be several times greater than the vessel diameter or the duct, the molecules migrate through a system freely and independently of each other. This is called *free-molecule flow* or simply *molecular flow*, and the gas flow rates are affected mainly by collisions of molecules with the tube walls. Molecular flow occurs at high vacuum when $\lambda > (d/3)$. As the mechanical pump removes gas from the *pipeline* between the pump and vacuum chamber, the gas is pushed toward the pump by means of collisions from its upstream neighbors. This describes the phenomenon of viscous flow and indicates its proportionality to pressure gradient. When a high-vacuum pump moves gas directly from a *chamber*, the pump waits for molecules to wander into the pump entrance and once in, to prevent them returning: typical efficiencies in this respect for high-vacuum pumps are about 40%. During molecular flow, where intermolecular collisions are rare, molecules can move away from the pump independently of molecules moving toward the pump.

1.5.5 FLOW RELATIONSHIPS

1.5.5.1 Volume Flow Rate, S, and Throughput, Q

The gas flow rate is an important aspect of the behavior of gases in a vacuum system. It determines the time required to reach the operating pressure; it may also

determine the magnitude of gas leaks (gas leaking into the system from the outside) and outgassing (gas produced somewhere within the system) that can be tolerated without the pressure in the vessel rising above the desired operating pressure. The rate at which the gas flows through a system can be expressed in two ways: as a volumetric flow rate and as a mass flow rate.

1.5.5.2 Volumetric Flow Rate

Volumetric flow rate is usually given the designation S and is the volume of gas flowing past a given point in a system per unit time. The volumetric rate measured at the pump entrance is called the pump speed, S_p.

1.5.5.3 Throughput Q

Throughput Q is related to the mass flow rate and is given the designation Q. It is proportional to the number of moles and hence to the mass of gas flowing past a given point per unit time. The throughput and the volumetric flow rate are related by the equation

$$Q = SP$$

where Q is the throughput and is the mass flow rate measured in torr liter/sec (Ls^{-1}) and S is the volumetric flow rate in liter/sec or Ls^{-1}, measured at P, the pressure in torr.

By definition, the throughput, Q, is constant for all points in a vacuum system that is at equilibrium. Mass flow rate is measured in grams per second or the number of molecules per second. Since pressure (torr) is a measure of the molecular concentration (number of molecules per liter) and the volumetric flow rate is given as Ls^{-1}, then the product torr-Ls^{-1} is proportional to the number of molecules passing any point in a vacuum system per second and hence the throughput. The rate at which gas flows through the vacuum system depends on the pump speed, the geometrical shape and the dimensions of the passages, and the type of flow. Thus a volumetric flow rate of 1000 Ls^{-1} at a pressure of 10^{-3} torr gives a throughput of 1.0 torr $\cdot$ Ls^{-1}. This quantity 1.0 torr $\cdot$ Ls^{-1} = (1.0/760) atmosphere Ls^{-1}. For air, 1 mole (29 g) occupies 22.4 L, that is, 1 atmosphere liter = 29/22 g. Table 4 lists throughput conversion factors.

Table 4

Throughput Conversion Factors

Throughput Unit	Equivalent torr-Ls^{-1} Rate
1 lusec	0.001
1 clusec	0.01
1 cm^3 (*NTP*) s^{-1}	0.76
1 cm^3 (*NTP*) min^{-1}	0.0127
1 ft^3 (*NTP*) min^{-1}	358
1 meter^3 (*NTP*) hour^{-1}	211
1 ft^3 (*NTP*) hour^{-1}	5.97
1 watt[a]	7.5
1 gs^{-1} at 20°C	1.7×10^4 molecular weight

[a]In the MKS system where the unit of force is the newton and the unit of pressure is the newton per meter^2, the unit of throughput is newton meter s^{-1} which equates to 1 joule · s^{-1} or 1 watt.

CHAPTER 1.6

Conductance

1.6.1 CONDUCTANCE

Any pipe or duct offers a certain resistance to gas flow of any type. This resistance causes a pressure drop along the pipe. If F is the volume flow rate of gas flowing per sec across any cross section of the pipe and P is the pressure at the section, the quantity of gas passing per sec is

$$Q = FP \tag{47}$$

The resistance R (or impedance) of a length of pipe is defined by

$$R = \frac{P_1 - P_2}{Q} \tag{48}$$

where P_1 and P_2 are the pressures at ends of the pipe.

Thus R has units of s/L. The reciprocal of R is called *conductance* and is designated by the symbol C.

$$C = \frac{Q}{\Delta P} = \text{quantity of flow/pressure difference} \tag{49}$$

Thus the conductance of an orifice, pipe or vacuum component is a measure of the throughput for a given pressure drop across the pipe or component and is therefore expressed as torr-Ls^{-1}/torr^{-1}, which becomes Ls^{-1}. The conductance varies according to the mode of gas flow and the nature of the gas. Hence the net transfer of gas through a component connected to a high-vacuum pump is proportional to the pressure difference across the component. For a component with a conductance C Ls^{-1} and pressures P_1 and P_2 torr existing at its ends, the net gas

ISBN 0-12-325065-7
$25.00

flow through the component is C ($P_1 - P_2$) torr-Ls^{-1}, which is the difference in mass flow rate from the high-pressure end (C_{P_1} torr-Ls^{-1}) and the backflow from the low-pressure end (C_{P_2} torr-Ls^{-1}).

1.6.2 CONDUCTANCES IN PARALLEL

Consider two pipes having conductances C_1 and C_2. The mass flow in each pipe is given by

$$Q_1 = C_1 P \text{ and } Q_2 = C_2 P \tag{50}$$

The total flow is

$$Q_1 \text{ and } Q_2 = (C_1 + C_2)\, P \tag{51}$$

If the two pipes were replaced by a single pipe of conductance, C_T, which would give the same flow, then

$$Q_T = C_T P \tag{52}$$

where C_T is the total conductance and the subscript $_T$ signifies total conductance or flow for the element.

For series connection of components, pipes, or orifices, the total conductance is given by

$$C_T = C_1 + C_2 + C_3 \text{ etc.} \tag{53}$$

1.6.3 CONDUCTANCES IN SERIES

Consider two conductances C_1 and C_2 in series; consisting, for example, of a pipe joined at one end to another pipe of different diameter. Let P_1 and P_2 be the pressure drops along pipes 1 and 2, respectively. Let Q_1 and Q_2 be the mass flow rates. Let P_T be the total pressure drop in series, and Q_T be the corresponding mean flow rate.

Thus

$$P_1 = \frac{Q_1}{C_1}, \quad P_2 = \frac{Q_2}{C_2}, \quad \text{and } \Delta P_T = \frac{Q_1}{C_T} \tag{54}$$

But

$$\Delta P_T = \Delta P_1 + \Delta P_2 \tag{55}$$

Furthermore, the mass flow rate must be same throughout the system, otherwise gas would accumulate in the system. Hence,

$$Q_1 = Q_2 = Q_T \tag{56}$$

Thus

$$\frac{Q_T}{C_T} = \frac{Q_1}{C_1} + \frac{Q_2}{C_2} \tag{57}$$

$$\frac{Q_T}{C_T} = \frac{Q_T}{C_1} + \frac{Q_T}{C_2}$$

Hence, for conductances in series,

$$\frac{1}{C_T} = \frac{1}{C_1} + \frac{1}{C_2} + \frac{1}{C_3} \text{ etc.} \tag{58}$$

For a pump of speed S Ls^{-1} connected to a chamber by means of a component of conductance C Ls^{-1}, the effective speed at the chamber, S_e, is given by

$$\frac{1}{S_e} = \frac{1}{S} + \frac{1}{C} \tag{59}$$

CHAPTER 1.7

Flow Calculations

The conductance of a duct is a measure of its ability to transport gas and is expressed in units of volume transported per unit time. The quantitative expressions used to calculate conductance of an element under different circumstances are fairly complex and depend on the type of flow as well as on geometrical and surface-related factors. Calculations of conductance and the corresponding gas flow rate for turbulent flow are difficult to treat analytically. Viscous flow is also somewhat difficult to treat quantitatively, because it depends not only on the shape of the duct but also on the gas pressure. Fortunately, at most pressures of interest to high-vacuum situations, the flow is molecular rather than viscous. Considerable effort has been expended for developing analytical techniques for determining flow under viscous and molecular regimes.

1.7.1 EQUATIONS FOR VISCOUS FLOW

Generally, quantitative expressions that have been developed for calculating conductance and corresponding flow rate under viscous flow conditions are those for fairly simple geometrical configurations, such as circular tubes or rectangular ducts. These expressions are used, for example, to calculate the time required to evacuate a vessel or volume of some sort, through a pipe that is usually circular or rectangular in cross section.

1.7.1.1 Circular Tube

The mass flow rate through a straight tube of circular cross section under viscous flow conditions, is determined by Poiseuille's equation, namely,

ISBN 0-12-325065-7
$25.00

$$\frac{Q}{P_1 - P_2} = \mathrm{K}\frac{d^4 P}{\eta L} \tag{60}$$

where d is the diameter of the tube
L is the length of the tube
η is the viscosity of the gas
P is the average of P_1 and P_2, the pressure at the opposite ends of the tube

For dry air at 20°C, this equation becomes

$$Q = \frac{750\, d^4\, \bar{P}}{L}(P_1 - P_2) \tag{61}$$

where Q is the mass flow rate in torr-L/sec, d is the tube diameter in inches, L is the tube length in centimeters, and P is the pressure in torr.

The expression for conductance for a circular pipe for air at 20°C is

$$C = \frac{2.94 P d^4}{L} \text{ L/sec} \tag{62}$$

1.7.1.2 Rectangular Duct

For the rectangular duct, let a = long side and b = short side. The Poiseuille equation for the rectangular duct for air is at 20°C is

$$C = \frac{30a2b2KP}{L} \text{ L/sec} \tag{63}$$

where K is a shape factor whose value depends on b/a. As can be seen, the conductance of the rectangular slit increases rapidly as the cross section changes from slit to square.

As in the case of round pipe, the expression for C leads to a relation for the volume flow in terms of the pressure drop along the duct.

$$C = \frac{PK}{\Delta P} \tag{64}$$

where

$$F = \frac{CP}{\Delta P} \tag{64}$$

Thus

$$K = \frac{30a^2b^2K}{L} \cdot \Delta P \text{ L/sec} \tag{65}$$

1.7.2 EQUATIONS FOR MOLECULAR FLOW

At low pressures, intermolecular collisions are less frequent than wall collisions, so the latter determine the gaseous flow characteristics through the channel. Specifically, two aspects determine the conductance of a duct during molecular flow:

1. Rate at which molecules enter the duct
2. Probability that the molecules are transited through the system

The first depends on the entrance area of the system, while the latter is determined by the subsequent series of reflections from the walls, which result in the molecule eventually being transmitted through the duct or reflected back into original volume.

Consider first the case of very thin aperture plate, for which the area A is more important in determining conductance than the wall area or wall conditions.

The volume of gas traveling from one side of the aperture to the other side per unit time — the aperture conductance — is

$$C_a = \frac{1}{4} A v_{av} \tag{66}$$

when the molecules have a Maxwell-Boltzmann velocity distribution. Conductance values depend on the molecular mass and kinetic temperature. The case where wall collisions are more important than the conductance of an aperture is considered next.

1.7.3 KNUDSEN'S FORMULATION

The conductance C_T of a length of long tube of length L with uniform cross-sectional area A and perimeter H, was calculated by Knudsen to be:

$$C_T = \frac{4}{3} \frac{A^2}{HL} v_{av} \tag{67}$$

The assumptions for obtaining the general result are

1. Length of tube is much greater than the diameter.
2. Direction of reflected molecules is independent of the incident direction.
3. Reflected molecules are distributed equally per unit angle (cosine law for reflection from a Lambertian surface).

Relationships derived from the general equation are given in Table 5 for simple geometries. Assumption 1 indicates that the effect of aperture is insignificant, and the conductance value is given by the preceding equation is for molecules that are well within the tube and are sufficiently removed from the aperture so that it is of no consequence. A rough attempt to correct this deficiency is to include a series conductance of the entrance aperture. Weissler and Carlson [5] gives a formula for a tube of perimeter H, area A, and length L:

$$C = \left(1 + \frac{3}{16} \cdot \frac{LH}{A}\right)^{-1} C_a \tag{68}$$

1.7.4 CLAUSING FACTORS

The conductance of a long tube is approximately related to the conductance of the entrance aperture through the factor $[1 + (3/16)(LH/A)]^{-1} C_a$. This factor can be interpreted as the probability that a molecule incident on the aperature will be transmitted through the tube and leave at the other end.

It is convenient to discuss conductance in terms of the aperture conductance and corresponding probability of passage, $P_{1\to2}$,—the Clausing factor—so that

$$C = C_a \cdot P_{1\to2} = (1/4) v_{av} A_1 P_{1\to2} \tag{69}$$

Because conductance is independent of direction,

$$A_1 P_{1\to2} = A_2 P_{2\to1} \tag{70}$$

Examples: The throughput in pressure–volume units per unit time (torr · cm^3 · s^{-1}) through a long tube is given approximately by

$$Q = \left(\frac{\pi d^3}{12L}\right) n_a (P_1 - P_2) \tag{71}$$

where d is the tube diameter in cm
L is the tube length in cm
n_a is the average velocity of a molecule in cm s^{-1}
P_1 and P_2 are the pressures (in torr) at the opposite ends of the tube

Approximate values for some probabilities of passage are accurate to within $\pm 10\%$. A variety of techniques, which include analytical methods [6], Monte Carlo calculations, and variation methods [16]. Carlson lists different geometries that have been investigated and cites the corresponding references. Numerical examples can be found in Carlson [22].

CHAPTER 1.8

Surface Physics and Its Relation to Vacuum Science

The interaction of gas molecules with a solid surface is the basis for vacuum technology [6–10]. Vacuum pumps and pumping, gas transport through a physical system, system pumpdown, and material behavior under rarefied gas conditions all involve the interaction of gas molecules with a solid or liquid surface. The surface of a material is generally considered the topmost atomic layer. This may be the terminating layer of a single crystal material or a more complex structure of another material. At atmospheric pressure, gas–solid interfacial interactions are characterized by chemical reaction kinetics, wettability through condensation, diffusion, and other well-known mechanisms. As the pressure is decreased, the number of molecules available to interact with a surface is decreased until pressures are so low that the number of molecules colliding with a surface is best described by statistics. This is where the conventional gas laws break down. Some of the phenomena mentioned earlier are surveyed in this section.

1.8.1

PHYSICAL ADSORPTION OR "ADSORPTION"

When the surface of a material is exposed to an ultra-high-vacuum ambient, gas molecules tend to "stick" to clean surfaces. Equilibrium is rapid, and the phenomenon is entirely reversible. When a molecule reacts at the surface to form a strong bond, this is known as *chemisorption* and is discussed briefly in this section. For a given pressure and temperature, the kinetic theory of gases, assuming Brownian

ISBN 0-12-325065-7
$25.00

motion of the molecules, can be used to compute the number of molecules hitting the surface. More simply, from PV (unit volume) = MRT, the basic gas law, the flux striking a unit area of a surface is defined as

$$F = N_{av}P(2\pi MRT)^{-1/2} \text{ (\#/cm}^2\text{/s)} \tag{73}$$

where N_{av} is Avogadro's number, M is the molecular weight, P is pressure in torr, T is the temperature in Kelvin, and R is the universal gas constant. This can be reduced to

$$F = 3.51 \times 10^{22} P(\text{torr}) [M(\text{g/mole}) T(^\circ\text{K})]^{-1/2} \tag{74}$$

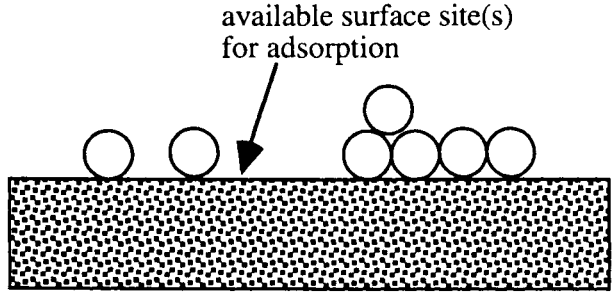
The flux of gas molecules, F, that come in contact with a surface is roughly $0.5 \times 10^{22} P$ (torr) for air. At 10^{-6} torr, $F \approx 10^{14}$ molecules/cm^2 sec impacting a surface.

When a gas molecule hits a surface and sticks, it is said to be *adsorbed* if it does not return to the gaseous state; that is, the molecule is trapped on the surface. The surface can be viewed classically, as a number of "billiard ball-type molecules" on the surface (Figure 3).

For this to occur, a long-range force must be acting on gas molecules, causing them to group together on the surface through a mutual attractive force called the *van der Waals force*. These forces, known as *dispersion forces*, are electrical in nature [7]. In this case, atoms tend to group together through the mutual attractive force (the van der Waals force, f) and was shown by de Boer to have a dependence: $f \propto 1/d^3$, where d is the distance between interacting molecules and a surface. This comes about through the interaction of distributed charges and dipole moment interaction. The force is always attractive and is described by Adamson in some detail [6].

Due to thermal vibration, adsorbed molecules will reside on the surface for a finite time depending on the energy of the thermal oscillation (temperature). This

Fig. 3.



A billiard ball diagram of molecules adsorbing on a flat surface.

is known as the *heat of adsorption*, H_{ads}. The residence time, τ, of a molecule is where

$$\tau = \tau_0 e^{\Delta H_{ads}/RT} \tag{75}$$

τ_0 is typically in the range of 10^{-12} to 10^{-13} sec. The fractional surface occupation of the molecules is θ:

$$\theta(\text{number/cm}^2) \propto F\tau$$
$$= N_{av}P(2\pi MRT)^{-1/2} \quad (\text{molecules/cm}^2/\text{sec})\tau_0 e^{\Delta H_{ads}/RT}\,(\text{sec}) \tag{76}$$

This assumes that all the molecules that hit an available site stick to it.

Multilayer physical adsorption occurs in many material systems. It was shown to exist at cryogenic temperatures by Gomer and his co-workers [10]. Visual evidence was observed using the field ion microscope. Molecules absorbed at very low temperatures (<70°K) were mobile and could laterally diffuse. A surface diffusion coefficient is defined from the motion of the "boundary" and given by

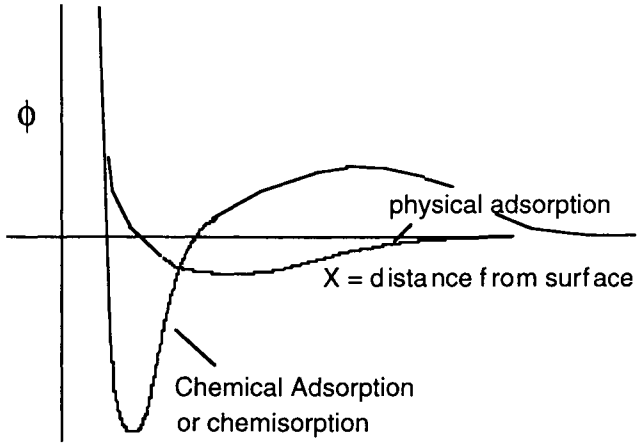
$$D_s = a^2\nu \exp(-E/kT) \tag{77}$$

where a = jump length ($\sim 3\text{Å}$), ν is the jump frequency ($\sim 10^{12}$/sec), and E is the activation energy for surface diffusion.

1.8.2 CHEMISORPTION

Physisorption phenomena are reversible phenomena; that is, with the addition of some stimulus such as thermal energy, the gas species or molecule will desorb from the surface. Using Somorjai's notation [7], this can be written chemically as $A \rightleftharpoons A^*$ where A^* is a nondissociated adsorbed molecule. When chemisorption occurs, A^* undergoes a one-directional reaction to form a new species on the surface. This requires a specific energy, and has a specific rate constant associated with the process. When the bonding energy between the absorbed species and the surface is large, the molecule or atom is considered to be chemisorbed to the surface. Isotherms are generally not as important in characterizing the surface, because only the first monolayer absorption is important. Figure 4 shows a typical energy diagram as a function of the distance from a surface. Molecules or atoms physically adsorbed are in a shallow potential well, as shown. Chemically adsorbed molecules reside in the deeper potential well. To break this bond, a significant amount of excess energy is needed to raise the molecule to an energy state greater than the barrier energy to free it from the surface.

Fig. 4.



Energy coordinate diagram of the difference between physisorption and chemisorption.

1.8.3 STICKING COEFFICIENT

The sticking coefficient for adsorption, S, is defined as the the probability of a reactive gas species interacting with the surface as a function of surface coverage, that is, $S \approx f(1 - \theta)$, where f is related to the bond strength between the adatom and the surface and θ is the surface coverage. The sticking coefficient is also a function of the surface structure and the temperature (substrate and gas). This includes the surface morphology and defect structure from an atomic and macro scale. Somorjai defines the rate of adsorption for a noninteracting molecule as the product of the flux, F, and the sticking coefficient, S.

At 10^{-6} torr, the gas will form a monolayer on the reactive surface in approximately 1 second of exposure. Assuming the residual gas pressure is 10^{-10} torr, and no other processes are functional, a monolayer will accumulate on the surface in $\sim 10^4$ seconds. This would permit "clean measurements" for ~ 100 seconds, assuming a θ of 0.01 would not appreciably perturb the measurement being made. Structural measurements, typically performed by scattering low-energy electrons from the clean single-crystal surface (low-energy electron diffraction or LEED), would require the surface to remain clean during the entire duration of the experiment. This requires special vacuum fixturing and the use of a very clean vacuum system (ultra-high vacuum). However, many surface chemical analytical

methods do not sample *only* the top surface monolayer of a material. This phenomenon will be discussed later in this book.

1.8.4 SURFACE AREA

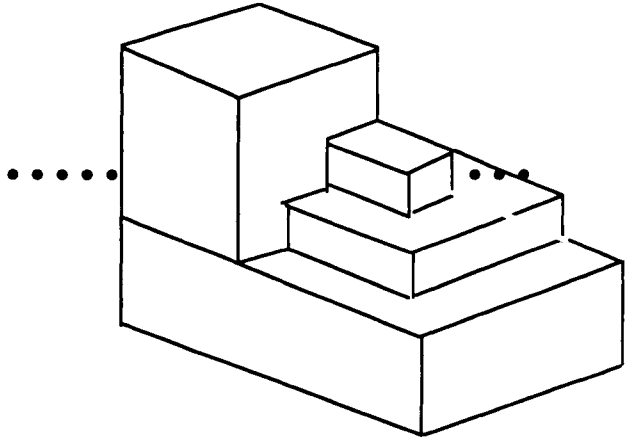
The surface of a material is generally thought of as simply the physical surface area due to the physical size of an object, that is, the integral over x and y (in this coordinate system). On the atomic level, surfaces are not flat. In fact, even single-crystal surfaces contain numerous defects such as screw dislocations at the surface, giving rise to atomic steps as shown in Figure 5.

Polishing a crystal surface may result in multiple steps as shown in Figure 5. The result of this is to increase the total surface area. Other steps are formed by the presence of chemically attached atoms caused by, for example, oxidation of the surface. These nonuniform regions may be many atom layers thick.

Machining a surface also produces macroscopic defects in the material surface. Lathes produce grooves, due to the tool size, as do many other procedures. Some of the defects can have a "high-aspect ratio"; that is, narrow, deep cuts in the surface can be produced. Many of the grooves or machining marks can act as capillaries.

The surface area of irregular surfaces or powders can be determined using models such as the BET (Brunauer-Emmett-Teller) isotherm [12]. This provides a measure of the heat of adsorption and the volume adsorbed. This is discussed more fully in the next section.

Fig. 5.



A small portion of a surface showing shelves and steps that have a profound effect on the total surface area.

1.8.5

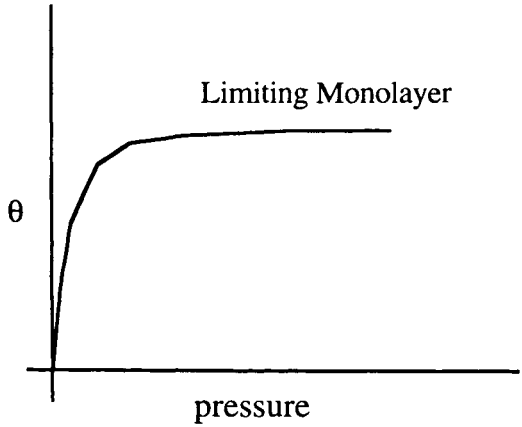
SURFACE ADSORPTION ISOTHERMS

Adsorption isotherms describe macroscopic effects of absorption of gases on a solid surface, that is, the effect of relative pressure on an adsorbed surface volume. Isotherms can be used to measure surface area as well as thermodynamic properties of the gas–surface interaction. Absorption isotherms include the effects of surface roughness or porosity. In the following, condensation will be discussed first and then the effects of the material on which condensation is occurring.

A metallic surface consists of an interfacial layer between the gas phase of the residual gases in the vacuum system. Portions of the gas phase consist of condensible gases such as water vapor. Water vapor is only one of the condensible molecules in air. When water vapor condenses on the surface of the metal, from a macroscopic point of view, a thin layer of water is captured as part of the gas–metal "interfacial" region. The water volume reaches an equilibrium state with the residual gas partial pressure of water. When the system is further pumped, the water on the surface tends to return to the gaseous state. As the temperature is changed, the relationship between the water vapor pressure and the amount of water on the surface of the metal also changes.

A plot of the surface coverage, θ, as a function of pressure or partial pressure of the gas species at a fixed temperature, is known as an *isotherm.* θ represents the number of available sites for adsorption and only takes into account vertical interaction. Figure 6 shows a typical isotherm for water vapor on a nonreactive smooth gold surface. As the partial pressure of water vapor is increased, the

Fig. 6.



Typical Langmuir isotherm, type I.

surface coverage increases until a plateau level is reached. This plateau level is known as a monolayer coverage of the surface and is descriptive of the Langmuir isotherm [6]. This is rarely found in nature but is used in plotting experimental data. Continuing to increase pressure eventually permits a second, third, etc. layer to condense.

The Langmuir isotherm is represented mathematically by

$$\theta = bP/(1 + bP) \tag{78}$$

where P is the pressure
θ is the fractional surface coverage
b is a constant $= k_1/k_2$ where $k_1 = (1/\tau_0) \exp(-Q/kT)$ and
$k_2 = N\sigma_0/(2\pi mkT)^{1/2}$

where T = temperature °K
Q = activation energy eV
k = Boltzmann constant
N = number of sites per unit area
σ_0 = site area
P = pressure (MKS units)
m = weight of a molecule
τ_0 = absorption time

More realistic models have been developed that include multilayers on a surface. The Brunauer-Emmett-Teller or BET isotherm is one of these more sophisticated models [12]. This takes into account vertical atomic interactions but neglects horizontal ones.

Mathematically, the BET isotherm is

$$\frac{P}{(P_0 - P)} \cdot \frac{1}{\theta} = \frac{1}{C\theta_m} + \frac{C - 1}{C\theta_m} \cdot \frac{P}{P_0} \tag{79}$$

where P_0 is the saturation or condensation pressure, θ is the surface coverage, θ_m is the monolayer surface coverage, $C \propto \exp(Q/kT)$, and Q is the heat of adsorption. A typical BET plot is where

$$\frac{P}{(P_0 - P)} \cdot \frac{1}{\theta} \tag{80}$$

plots linearly as a function of

$$\frac{P}{P_0} \tag{81}$$

From the slope and intercept, the heat of adsorption and θ_m, the monolayer capacity in $\mu g/cm^2$ can be computed. With the heat of adsorption, going back to the

first subsection of this section, it is now possible to compute the residence time of molecules on a surface and is useful in obtaining a fundamental understanding of the surface on which adsorption occurs. The reader is referred to the literature for a more detailed discussion of isotherms and their use in surface physics [6–10,12].

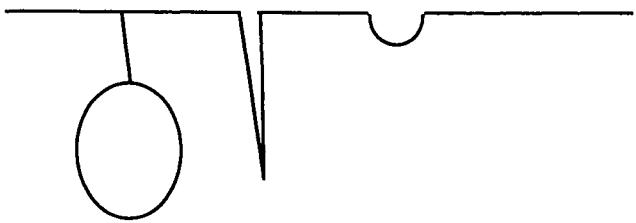
1.8.6 CAPILLARY ACTION

Surface porosity can occur in many forms as shown in Figure 7. Exposure of the pores to atmosphere or other gaseous environments may result in the condensation of some species in the mechanical defect structure of a surface. Under the proper conditions, internal voids, pores, and surface sites can be filled with the condensate through capillary action. Capillary action occurs when the condensate wets the substrate. The equation of Young and Laplace predicts a pressure differential across the gas-to-liquid interface. Therefore, pores will tend to fill with condensate as a result of the pressure differential. A good review of this theory is presented in reference [6].

As pumping is performed to evacuate a system, at a fixed temperature, the pressure in the vacuum vessel will lag the theoretical pumping capacity as a result of trapped condensates, mainly water vapor, evolving from the various surface features. This is shown in Figure 8 on the decreasing pressure isotherm.

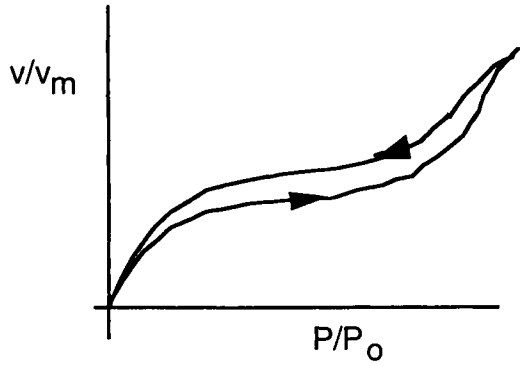
Figure 8 shows a isotherm where the volume of adsorption is plotted as a function of pressure. As the pressure is decreased, the curve maintains a high adsorbed volume until, at a much lower pressure, the adsorbed volume drops. During increasing the pressure, the adsorbed volume is lower than the adsorbed volume at the same pressure. This hysteresis is indicative of the presence of pores in a surface. The pores during the decreasing pressure cycle tend to retain a higher adsorbed volume of adsorbate than during the increasing pressure cycle. The difference can be used to compute the total pore volume.

Fig. 7.



Some typical features that can act as pores on a material surface.

Fig. 8.



Isotherm of a typical porous surface shows hysteresis between pumpdown and up to air showing capillarity.

Pores can result from poor welds in metal vacuum systems and act as virtual leaks. This results in a hysteresis in the isotherm. Using the BET model for isotherms, total porosity can be determined from the area enclosed in the hysteresis loop. Ultimately, the total surface area, which is significantly larger than the geometrical area of the surface, can be determined (see Parts 4 and 5 of this book).

In addition to pores at the surface, nonreactive gases can diffuse into the material surface. This occurs with "Fickian" kinetics [13]. With decreasing pressure, the gas will tend to diffuse out of the surface but lag the theoretical pumping curve.

In Part 4, the details of surface treatment and how it affects surface morphology are considered in detail. Perhaps the largest source of gas in a vacuum system after the initial pumpdown to the region of molecular flow (10^{-6} torr and below) is from the system itself. Metal surfaces are not planar but consist of machine marks, welds, and many other defects that may increase the effective surface area of the system. If a surface profilometer is drawn across a metal surface and a three-dimensional map made of the topography, the area of the sample is 1×1 mm^2. It is clear that the actual area must be computed from the peaks and valleys from machining marks, etc. This may increase the effective surface area by a factor of more than 3. However, this is only a small change in the total system volume. It is important, then, to discuss why this surface affects vacuum system performance.

1.8.7

CONDENSATION

Gas interaction with a surface is a material, thermal, and morphological phenomenon. Condensation implies that the gascous species develops a multilayer thick

film that has undergone a phase transition from the gaseous state to either liquid or solid state on a substrate [18]. A condensation coefficient can be associated with this phenomenon. This is defined as the probability that a gas molecule will stick to the condensate on a surface. Experimentally, condensation can readily occur on a substrate that is held at a temperature well below the boiling point of the condensate, such as during the evaporative condensation of hot metal on a cold substrate. The hot metal freezes on impact, forming a film. This phenomenon forms the physical basis for thin-film deposition and will be discussed later in this book. The condensed phase will increase in thickness if the vapor pressure of the condensate is above an equilibrium pressure or evaporate if the condensate is below this equilibrium pressure.

When the substrate temperature is above the boiling point of the condensate, the sticking coefficient can be small. This is of importance in pumping vacuum systems and reaching ultimate ultra-high-vacuum limits. Surface condensation of gases on surfaces is also of importance in pumping where the pumps are dry pumps based on surface adsorption, UHV pumps, and cryogenic pumps (see later discussions in this book).

Perhaps the phenomenon best known and studied in the vacuum technique field is condensation of water vapor on surfaces as it relates to system evacuation. Water condensation initially occurs by hydrating surface oxides on the system metal components. After the surface is hydrated, water adsorption can be macroscopically defined and measured using the BET isotherm where P_0 = a pressure of ~20 torr (2.67 kPa). Many layers of water molecules can form a film on the surface. When this occurs, capillarity can causes water to fill pores on metal surfaces, thereby loading the metal surface with water. Desorption from a saturated system requires considerably longer pumpdown times than a system in which water had not condensed and the metal surfaces are smooth. During pumpdown, the system pressure may reach a plateau at P_0. This is due to the desorption of water into the vapor state and subsequent pumpout of this vapor. After pumping the vapor from the system volume from the saturated system walls, the remaining residual gases are pumped away. A detailed review of the effects of water vapor in vacuum systems is presented in reference [18].

1.8.8 DESORPTION PHENOMENON

The long-range bonds found in physical absorption can be overcome by adding energy to the adsorbed molecule. The net effect may be to lower the total rate of adsorption or to fully remove the adsorbed species from the surface. As in the case of adsorption, desorption can be described in terms of rate kinetics and has

been measured for a number of material systems [17]. In practical terms, desorption from material surfaces may limit the pumpdown rate of a vacuum system and limit its ultimate pressure.

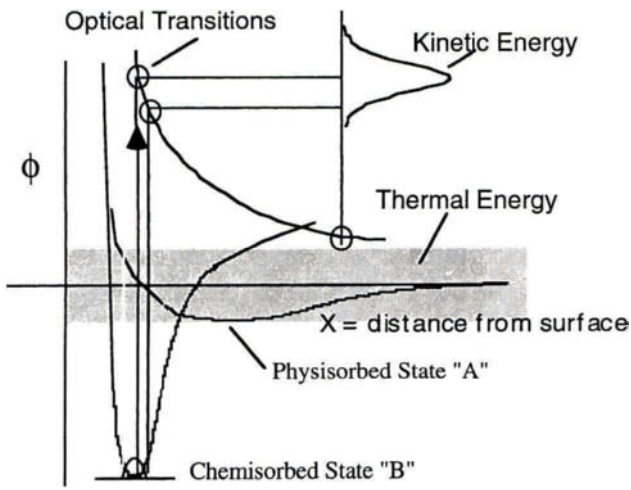
1.8.9 THERMAL DESORPTION

If a molecule is physically adsorbed (if a chemical bond has not been formed between the surface and the absorbate), then the molecule maintains its gas-phase structural features. Its state can be shown by an energy coordinate diagram. By heating the surface, the molecule can be desorbed or freed from the surface. The thermal energy added to the system is shown as a shaded region in Figure 9.

By adding energy through photon absorption (increasing the absorbate temperature), the molecule becomes excited and vibrates. The vibrational energy may be large enough to produce a transition to a nonbonding (repulsive) state at the surface. In this excited state, the molecule will have a finite probability of leaving the surface, as shown in Figure 9 (state A). An example is shown in Figure 9 where the energy is great enough to release the physisorbed molecule but not large enough to break a chemical bond formed (state B).

Surface adatoms or molecules can also be present as a result of bulk and surface diffusion, and permeation. In diffusion, gas can be entrained in metal fixturing and diffuse by grain boundary diffusion or bulk diffusion to the surface. Bulk diffusion is controlled by Fick's laws [13] and generally is slow, and grain boundary diffusion is rapid and has a low activation energy. The bulk diffusion coeffi-

Fig. 9.



Energy coordinate diagrams of thermal and optical transitions at a surface, after Madey [14].

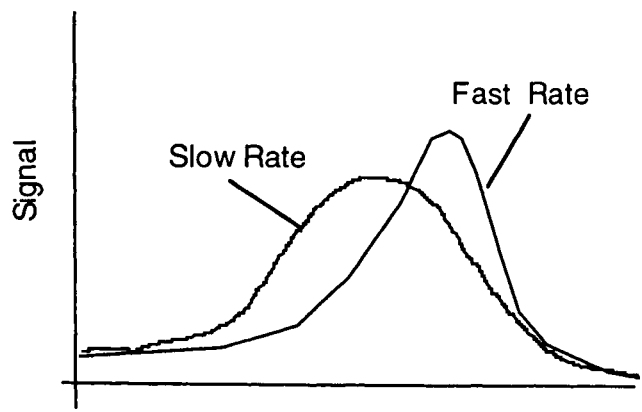
cient, D, is typically in the range of $\sim 10^{-14} - 10^{-19}$ cm^2/sec. From Fick's first law, the concentration $C \propto erfc\ (4 \times Dt)^{-0.5}$ where $D = D_0 \exp(E_D/kT)$, that is, D is thermally activated with an activation energy of E_D, where t is time, D is the diffusion coefficient, k is the Boltzmann constant, and x is distance. From this, estimates of the time required to form a monolayer on a surface can be obtained. If the species diffusing from the bulk is volatile—that is, has a high vapor pressure—then this species will contribute to the pump's gas load.

O'Hanlon describes permeation as gas diffusion from the outer wall to the inner wall of a vacuum system. This process is slow and occurs at a steady rate that is a function of temperature. This process is similar to and is described by Fick's laws [13].

Thermally stimulated desorption can be used as a measure of the desorption energy assuming that the process is a thermally activated process, where the desorption rate is $R \propto \exp(-E_d/kT)$ and E_d is the activation energy due to desorption. The energy levels of the desorbed species are an indication of the physical binding energy of the species to the surface. Experiments have been performed by measuring the pressure in a sealed vessel or a vessel pumped at a constant rate. In the later case $p \propto p_0 \exp(-E_d/kT)$, where p is the pressure measured with a gauge. This can be analyzed in terms of the residence time, τ. Details of this theory have been published by King [16].

In more practical terms, thermally stimulated desorption occurs when a vacuum system is heated after initial pumpdown. Figure 10 shows a sketch of a typical heating curve and the measured pressure in a system under constant pumping conditions. If all the gas is desorbed in each experiment, the areas under the curves should be identical. The peak temperature is related to the heat of desorption, E_d.

Fig. 10.



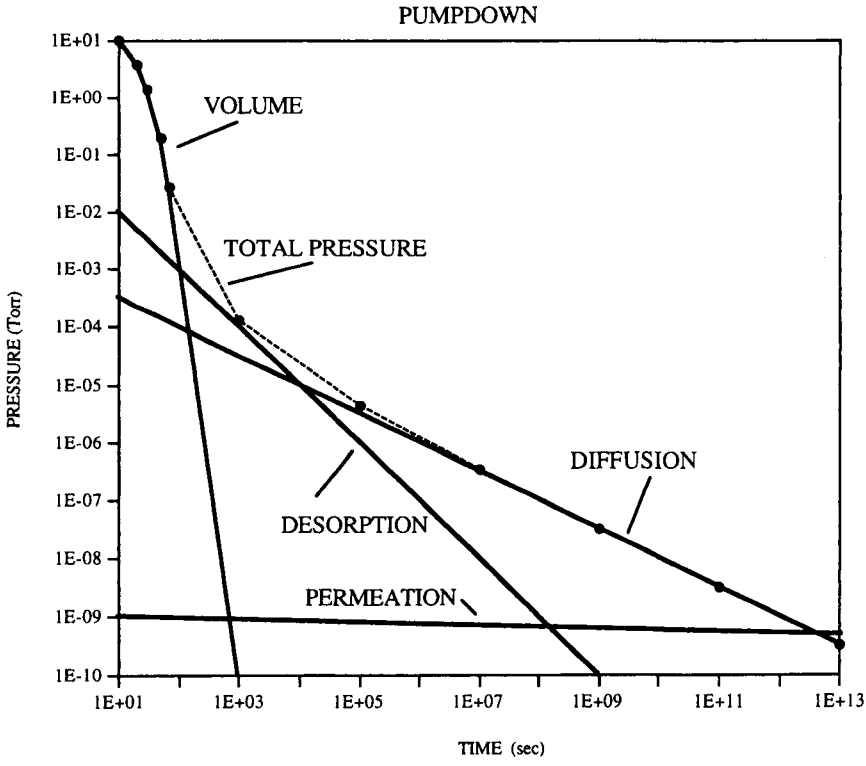
Thermal desorption at two heating rates, assuming a simple contaminant on a substrate. The signal in this case is pressure. The pumping rate can also affect the desorption curves.

Thermal desorption is common used to clean vacuum systems and can be observed when thermally degassing a system prior to operation. Near room temperature the thermal desorption rate is inversely proportional to time, t^{-1} [1]. This phenomenon has been extensively studied and is found to be very important in the application of vacuum technology.

1.8.10 PHOTOACTIVATION

When an atom is physically or chemically adsorbed on a surface, it can be released by the addition of energy through photon absorption. Referring to Figure 11, a vertical photon absorption process excites the atom without the addition of heat.

Fig. 11.



A sketch of the pumpdown regimes as a function of time, after Fig. 6.5 from O'Hanlon [1].

If the energy is large enough, the absorbate will be excited and an ion, excited atom, or neutral atom will produced. The physical phenomenon is described in terms of a model that predicts the formation of a repulsive state formation. The repulsive electronic state converts to motion, kinetic energy, and the molecule or atom leaves the surface as shown in Figure 11 and is free to move into the vacuum.

Photon-stimulated desorption (PSD) has become a surface science method for investigating the effect of photons on surfaces. Energy threshold measurements provide direct information relating to the bond-breaking process. Madey [15] describes in his review article that the threshold energy for neutral desorption can be as low as 5 eV (deep UV). At higher energies ($\sim$15 eV), one expects ion desorption to occur [19].

1.8.11 ULTRASONIC DESORPTION

Ultrasound-stimulated desorption is for practical purposes used to clean metal vacuum systems while under vacuum. The ultrasonic power is coupled to the system through impedance matching barium titanate crystals to the metal and thickness of the structure. This method of causing desorption is based on energy absorbtion by the absorbate causing localized heating, thereby releasing molecules through the addition of what is believed to be thermal energy. Water vapor is the most common species to be desorbed from metal surfaces using this method.

1.8.12 ELECTRON- AND ION-STIMULATED DESORPTION

This phenomenon has also been developed into a surface analytical technique known as electron-stimulated desorption (ESD) where specific species and/or ions are studied by low-energy electron diffraction as a surface-sensitive tool or by directly detecting the desorbed species leaving the surface using, for example, mass spectrometry. Low-energy electrons ($<$500 eV) are used to bombard monolayer-coated ultra-clean surfaces. Both chemisorbed and physisorbed molecules can be released from the surface as ions or neutrals. This method is sensitive to the bond strength and surface structure on an atomic scale. The cross section for the formation of neutrals and ions by electron bombardment is in some cases rather large (10^{-2} molecules/incident electron), consequently the signal expected from the surface can be large.

Electron-stimulated gas desorption or atomic desorption was discovered during the manufacture of electron vacuum tubes and was used in part to clean the metal structures in the glass or metal envelope. The methodology was developed as a production cleaning technique involving both thermal and electron bombardment desorption. This was performed with little physical understanding of the physics and chemistry of the desorption process. For example, electron-stimulated cleaning is still used in ultra-high-vacuum nude pressure gauges where this is the only method of removing residuals from the metal structure.

Electron beam bombardment as well as photon absorption by polymers can result in decomposition of the material in addition to electron-induced desorption of surface species [17,19]. These materials are generally not used where they might be exposed to high-energy particle environments. Energy near the bond strength of a typical carbon-to-carbon bond is needed to break the bond. Under appropriate conditions, the material can decompose in a dramatic fashion through coulombic explosion [17]. In essentially the same fashion, electron bombardment can break bonds, resulting in fragmentation of the molecule that can be released into the vacuum.

Ion-stimulated desorption of adsorbed species is similar to electron beam-induced desorption. Ion yield is larger than electron-stimulated desorption yields as a result of the increased particle mass. This class of desorption event is related to physical sputtering when the energy is great enough (see later discussions in this book). Typically, the energy must be above the threshold energy for sputtering (~10 electron volts, eV) to produce desorption. When the energy is higher, molecular fragments are produced and can be released from the absorbate. This mode is what is typically found in a static SIMS (secondary ion mass spectrometry) experiment employed in surface analysis of materials (see Part 5 of this book). Ion bombardment is common when ion gauges are degassed to produce a pressure rise in the vicinity of the gauge. Oxygen glow discharge cleaning is commonly used for substrate, vacuum chamber, and fixture cleaning. In this case oxygen ions bombard the surface along with neutral bombardment.

1.8.13 GAS RELEASE FROM SURFACES

The mechanisms listed — vaporization, thermal desorption, photoactivation, electron- and ion-stimulated desorption, and outgassing through diffusion (or permeation) — all produce a gas load on a vacuum system being pumped from atmospheric pressure. O'Hanlon describes the rate-limiting steps in pumping a vacuum chamber as volume, surface desorption, diffusion, and permeation. They occur in sequence. Figure 11 shows a sketch of O'Hanlon's Figure 6.5 [1].

REFERENCES

1. J. F. O'Hanlon, *A User's Guide to Vacuum Technology* (Wiley Interscience, New York, 1980).
2. A. H. Turnbull, R. S. Barton, and J. C. Riviere, *An Introduction to Vacuum Technique* (Wiley, New York, 1962).
3. L. Ward and J. B. Bunn, *Introduction to the Theory and Practice of High Vacuum Technology* (Butterworth's, London, 1967).
4. G. F. Weston, *Ultra High Vacuum Practice* (Butterworth's, London, 1985).
5. G. L. Weissler and R. W. Carlson, *Vacuum Physics and Technology* (Academic Press, 1979).
6. A. W. Adamson, *Physical Chemistry of Surfaces* (Wiley Interscience, New York, 1994).
7. G. A. Somorjai, *Introduction to Surface Chemistry and Catalysis* (Wiley Interscience, New York, 1994).
8. G. L.Weissler and R. W. Carlson (eds.), *Vacuum Physics and Technology*, Vol. 14 of *Methods of Experimental Physics* (Academic Press, San Diego, CA, 1979).
9. S. Roy Morrison, *The Chemical Physics of Surfaces* (Plenum Press, New York, 1978).
10. R. Gomer, *Field Emission and Field Ionization* (Harvard Univ. Press, Cambridge, MA, 1961).
11. J. H. de Boer, *The Dynamical Character of Adsorption* (Oxford Univ. Press, New York, 1953).
12. S. Brunauer, *The Adsorption of Gases and Vapors*, Vol. 1. (Princeton Univ. Press, Princeton, NJ: 1945); S. Brunauer, P. H. Emmett, and E. Teller, *J. Am. Chem. Soc.*, **60**, (1938) 390.
13. For example, see, J. Crank, *The Mathematics of Diffusion* (Oxford Univ. Press, London, 1970); D. Gupta in *Diffusion Phenomena in Thin Films and Microelectronic Materials*, edited by D. Gupta and P. S. Ho (Noyes, Park Ridge, NJ, 1988), Chap. 1.
14. T. E. Madey, Radiation damage in auger electron spectroscopy and x-ray photoelectron spectroscopy, in *Analytical Electron Microscopy*, edited by D. C. Joy (San Francisco Press, San Francisco, 1987), pp. 345–349.
15. T. E. Madey and John T. Yates, Jr., *J. Vac. Sci. Technol.*, **8** (1971) 525.
16. D. A. King, *Surf. Sci.*, **47** (1975) 384; M. W. Roberts and C. S. McKee, *Chemistry of the Metal–Gas Interface* (Clarendon Press, Oxford, 1978).
17. T. A. Carlson, The coulomb explosion and recent methods for studying molecular decomposition, in *Desorption Induced by Electronic Transitions DIET*, edited by N. H. Tolk, M. M. Traum, J. C. Tully, and T. E. Madey (Springer-Verlag, Berlin, 1983).
18. A. Berman, Water vapor in vacuum systems, *Vacuum*, **47** (1996) 327–332.
19. T. Kelen, *Polymer Degradation* (Van Nostrand Reinhold, New York, 1983), Chap. 7.

This Page Intentionally Left Blank

PART 2

Creation of Vacuum

2.1 Technology of Vacuum Pumps — An Overview 59
- 2.1.1 Vacuum Pump Function Basics 59
- 2.1.2 Gas Transport: Throughput 61
- 2.1.3 Performance Parameters 62
- 2.1.4 Pumping Speed 64
- 2.1.5 Pumpdown Time 65
- 2.1.6 Ultimate Pressure 69
- 2.1.7 Forevacuum and High-Vacuum Pumping 71
- 2.1.8 Pump System Relationships 73
- 2.1.9 Crossover from Rough to High-Vacuum Pumping 78
- 2.1.10 Pumping System Design 79

References 83

2.2 Diaphragm Pumps 84
- 2.2.1 Introduction: Basics and Operating Principle 84
- 2.2.2 State-of-the-Art Design and Manufacturing 87
- 2.2.3 Performance and Technical Data 91
- 2.2.4 Modular Concept for Specific Application Setups: Standalone Operation 92
- 2.2.5 Diaphragm Pumps as Backing and Auxiliary Pumps in Vacuum Systems 93

References 96

2.3 Vacuum Blowers 97
- 2.3.1 Introduction 97
- 2.3.2 Equipment Description 97
- 2.3.3 Blower Operating Principle 100
- 2.3.4 Blower Pumping Efficiency 101
- 2.3.5 Blower Pumping Speed Calculations 103
- 2.3.6 Power Requirements 104
- 2.3.7 Temperature Considerations 106
- 2.3.8 Flow and Compression Ratio Control Mechanisms 108
- 2.3.9 Liquid-Sealed Blowers 112
- 2.3.10 Selected System Arrangements 112

2.4 Vapor Jet Pumps (Diffusion Pumps) 116
2.4.1 Basic Pumping Mechanism 117
2.4.2 Pumping Speed 122
2.4.3 Throughput 127
2.4.4 Tolerable Forepressure 128
2.4.5 Ultimate Pressure 132
2.4.6 Backstreaming 137
2.4.7 Other Performance Aspects 144
References 148
2.5 Cryogenic Pumps 149
2.5.1 Introduction 149
2.5.2 Cryopump Basics 156
2.5.3 Advanced Control Systems 167
2.5.4 Cryopump Process Applications 173
2.5.5 Cryogenic Pumps Specifically for Water Vapor 177
2.5.6 Comparison of Cryopumps to Other Types of Pumps 179
2.5.7 Future Developments 181
References 181
2.6 Turbomolecular Pumps 183
2.6.1 Turbomolecular Pumps (TMP) 183
2.6.2 Molecular Drag Pumps (MDP) 195
2.6.3 Combination of Pumps (TMP + MDP) 197
2.6.4 Evaluation of Combinations of Backing Pumps and TMPs, Etc. 200
2.6.5 The Use of TMP in Applications: Specific Effects and Demands 208
2.6.6 Avoiding Operational Mistakes 211
References 212
2.7 Pumps for Ultra-High Vacuum Applications 214
2.7.1 System Design for Ultra-High Vacuum 215
2.7.2 The Selection of Pumps for Ultra-High Vacuum Applications 216
2.7.3 Sputter-Ion Pumps 220
2.7.4 Getter Pumps 242
References 252

CHAPTER 2.1

Technology of Vacuum Pumps—An Overview

Marsbed Hablanian *(Retired)*
Varian Associates

2.1.1 VACUUM PUMP FUNCTION BASICS

Under normal environmental conditions, creation of vacuum requires a removal of gases, contained in atmospheric air, from a given vessel or a chamber. Thus, high-vacuum technology involves the rarefaction of atmosphere by as much as 15 orders of magnitude (from 10^3 to 10^{-12} mbar). This would be relatively simple if the chamber walls did not have any internal gas content and any adsorbed gas at the surfaces. The continuous evolution of occluded and adsorbed gases usually necessitates continuous pumping even if the object is only to maintain a given high vacuum level inside a chamber, such as the space between the double walls of a thermos bottle. Even in this case, after evacuation the sealed space must contain some residual pumping means, such as chemically active evaporated metallic films.

There are a few basic ways to remove air (or another gas) from the space of a vacuum chamber.

1. Gas can be removed by an action of positive displacement devices, such as reciprocating pistons, rotary pistons, rotary vanes, gears, cams, meshed screws, orbiting scrolls, and so on.

ISBN 0-12-325065-7
$25.00

2. Another way is by partial displacement or momentum transfer, such as fast-moving blades (fan or turbine), or a high-velocity jet stream of a vapor or other motive fluid (ejector).

Pumps based on these methods are often called *throughput* pumps, because usually the gas evacuated from the vacuum chamber is compressed and exhausted to the atmosphere. Pumps based on the following removal methods are called *capture* pumps.

3. Vacuum can also be created by refrigeration to temperatures at which the gas either becomes solid or can be adsorbed on the surface of a suitable porous material.
4. Gas can be removed from space by chemical reaction, which produces a solid residue.
5. Finally, the gas can be ionized and expelled from space either by the action of suitable electromagnetic fields or embedded (implanted) into a solid surface inside the pump due to the high velocity achieved by acceleration in a high-voltage field.

The word "pump," usually associated with liquids, is only of historical significance in high-vacuum technology. The displacement devices could be called more descriptively "rarefied gas compressors" and capture devices, perhaps, "condensers."

Generally, it is not easy to design a single pumping device that would function effectively under conditions of density changes involving 15 orders of magnitude. All pumping devices have pressure or density ranges in which they have reasonable performance, size, weight, noise level, cost, etc. Typically, at least two different pumping devices are used to obtain high vacuum. They can be conveniently designated as coarse vacuum pumps and high- (or fine) vacuum pumps. Coarse pumps are used to remove the bulk of the gas from a chamber (the gas that was in the space or volume), and high-vacuum pumps are used for maintaining vacuum by removing the gas evolving from the surfaces. Very roughly speaking, these two types are associated with the conditions of viscous and molecular flow.

Compared to the ordinary compression work near atmospheric pressures, high-vacuum gas flow conditions are characterized by very high volumetric flows, extremely small mass flows, rather high pressure ratios, and very small pressure differences. The onset of molecular flow produces additional complications. First, under molecular flow conditions, pressure differences are very small (although significant density gradients usually are present). Thus, individual molecules do not travel into the pump following the direction of reducing pressure, but enter the pump as a result of their normal thermal motion. Therefore, when molecules rebound from a surface they do not "know" which is the downstream direction, and they have an equal chance to move toward or away from the pump. An additional complication of molecular flow is the separation of flow behavior of different gases in a mixture. Because of the absence of gas to gas collisions, different

gases may be pumped with different pumping effects; that is, different volumetric flow rates and different compression ratios.

The absence of pressure gradients produces lower values for maximum possible volumetric flow through a restriction, such as an orifice. In viscous flow the volumetric flow rate is nearly twice as high as that for molecular flow (for air, near 20 L/s cm^2, versus 11.6 L/s cm^2 at room temperature). In general, if the orifice diameter is less than or equal to the mean free path, the 11.6 L/s figure applies [1]. Also, relatively very high pressure ratios may exist in molecular flow across a rather short flow duct. For example, an ordinary elbow may produce 50% reduction in volumetric flow compared to a few percent in viscous flow. Therefore, the conductance of ducts in high-vacuum system design becomes important. The concept of conductance (instead of resistance or friction coefficients in conventional fluid mechanics) is described in detail in Section 2.4 and Part 1.

The units of conductance are usually given in "liters per second" (L/s) for simplicity. However, the correct units are "torr L/s per torr" or, in other words, torr L/s of throughput per torr of pressure difference.

2.1.2 GAS TRANSPORT: THROUGHPUT

The actual amount of gas flowing through a duct or into a pump cannot be determined by volumetric flow unless its density, δ, is also known. The more important quantity is the mass flow expressed in units such as grams per second. Usually, the conversion from volume flow to mass flow can be easily made using the well-known gas law relationships

$$\begin{aligned} PV &= \mathrm{R}T \\ Q &= V \cdot \delta \end{aligned} \tag{1}$$

where P = pressure, R = gas constant, T = absolute temperature, Q = mass flow, and δ = density.

However, these relationships may lose their usual meaning when the gas approaches molecular beam conditions, that may exist on occasions in high-vacuum systems. It is not a simple matter to convert flow in grams per second into liters per second for a gas that does not have isotropic molecular velocity distribution (non-Maxwellian gas). For this reason, great precision and accuracy is often not possible for considerations of design of most high-vacuum systems.

Because mass flows are very low, it is customary to use "throughput" instead of mass flow. The assumption is made that heat transfer effects are negligible and the gas quickly accommodates to the temperature of the vessel or pumping ducts.

When temperature is constant, the "mass flow" can be measured in units of throughput, torr liters per second. Throughput then is simply the product of pressure and the volume flow of a gas ($Q = Sp$) for any given location (cross-sectional plane) in the flow passages.

Most analytical discussion of molecular flow found in textbooks on vacuum technology is associated with noncondensable gases and steady-state flow condition. Surface effects that govern the transient flow of condensable gases have little effect on O_2, N_2, etc., at normal ambient temperatures and thus are often neglected. (A detailed discussion is presented by Lewin [2].) Usually vapors of higher molecular weight are troublesome, but water vapor and even such "volatile" substances as acetone or alcohol are difficult to work with at room temperature. For example, if we attempt to measure the pumping speed of a high-vacuum pump for water vapor, the measurements of the flow rate must be done by carefully considering the available surface area for evaporation from the liquid, heat transfer necessary for evaporation, the temperature of the tubes leading the vapor into the pump, etc. Even then it is difficult to separate the amount of water vapor actually pumped by the pump and that which is adsorbed inside the cold inlet areas. For example, pumping-fluid vapor in a baked system may take a week to be noticed by a high-vacuum gauge with a 4-cm-long, 2-cm-diameter entrance tube, or a mass spectrometer can retain a strong "memory" of acetone for two or three weeks if its ion source is 30 cm away from the pumps (2-cm-diameter tube). In such cases, baking is indispensable.

In molecular flow environment, gas molecules rarely collide and, therefore, move independently of each other. An outgassing or evaporating molecule can move as easily away from the pump as well as into it. There exists a density (number of molecules) gradient toward the pump, but for condensable species, even this cannot be assured, as temperature distribution may influence the flow more than the presence of a pump.

It is often assumed that residual collisions between gas (or vapor) molecules are completely negligible in high vacuum. This is not always true. Sometimes the remaining collisions may be the predominant method by which certain molecules pass through barriers such as cryogenic traps. Also, in some cases, introduction of one gas into the system may affect flow conditions of another. For example, in turbomolecular pumps, an improvement of density ratio for helium can be obtained by introducing a certain amount of argon into the pump [3].

2.I.3 PERFORMANCE PARAMETERS

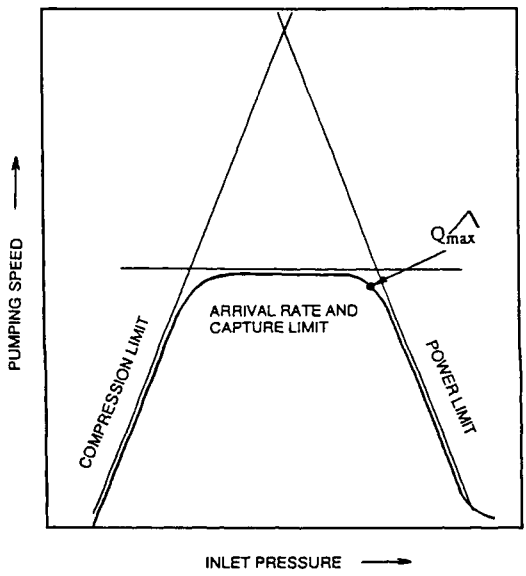
Performance of a vacuum pump or a compressor can be characterized, measured, and displayed in several ways, depending on the interest in a particular quality.

Basic parameters are volumetric flow (known as pumping speed in high-vacuum technology), mass flow (usually displayed as pressure–volume flow: throughput), compression ratio, and pressure difference (between discharge and inlet). The relationships between these parameters can be shown in several ways. For example, we can plot pressure versus flow or flow versus pressure, where the flow can be either volumetric or mass flow and the pressure term can be in the form of a pressure ratio or pressure difference. The emphasis can be on inlet pressure or discharge pressure, etc.

The most common traditional representation of performance has been in the form of pumping speed versus inlet pressure. This is a somewhat unfortunate choice because it is a reverse of the common practice in other engineering fields. It tends to produce a conceptual confusion between the dependent and independent variables, and it does not clearly distinguish between transient and steady-state range of the operation. High-vacuum pumps are usually rated by their pumping speed, the maximum permissible discharge pressure, the best obtainable vacuum level (lowest possible absolute pressure), maximum obtainable compression ration, maximum mass flow capacity (throughput), and some secondary characteristics related to the possibility of system contamination emanating from the pump.

When the performance of a vacuum pump (or a compressor) is presented in the form of a plot of volumetric flow (pumping speed) versus inlet pressure, the basic limitations of the performance may be represented as shown in Figure 1. If this

Fig. 1.



Basic limitations of pump performance.

graph were simply tilted 90 degrees to show pressure versus pumping speed, we should be immediately suspicious of this double-valued function, which does not resemble the usual representations of pump and compressor performance found in the fields of mechanical engineering, fluid mechanics, or pneumatic machinery, where the power limit section of the curve is normally omitted because it is not really a part of intended performance; in other words, the performance range has reached its power limit. However, in high-vacuum practice, the initiation of crossover from roughing to high-vacuum pumps can be made in this region as long as the maximum flow capacity of the high-vacuum pump is not exceeded.

2.1.4 PUMPING SPEED

It is customary in high-vacuum technology to express gas flow in liters per second. This is convenient in basic computations if the temperature is assumed to be constant. There exist two distinct geometries in which the flow parameters are associated ($Q = pS$). One can be called orifice geometry (Knudsen geometry); the other, pipe flow geometry. The two are not identical. In the case of the pipe flow, the pressure and speed are associated with the same cross-sectional plane: In the case of the orifice, the pressure is measured upstream from the orifice. The orifice geometry is shown in Figure 2. The definition of orifice conductance, C, is derived from

$$Q = C(P_1 - P_2) \tag{2}$$

where P_1 and P_2 are upstream and downstream pressures and Q is throughput. If P_2 is much smaller than P_1, we may speak of the orifice speed

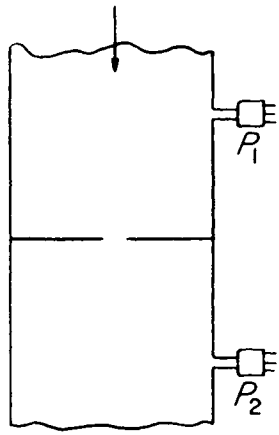
$$S = C = Q/P_1 \quad \text{when} \quad P_2 << P_1 \tag{3}$$

Note that in this case, the P_1 is not associated with the vicinity of the orifice, but with an upstream location where the gas conditions may be considered to be isotropic.

As derived from kinetic theory of gases and confirmed by experiments, that for air at room temperature, each square centimeter area of an orifice (having "zero" thickness) will have a maximum pumping speed of approximately 11.6 L/sec. When computing the number of molecules bombarding a small (compared to the chamber) unit area at a given pressure (measured far away), we can arrive at the same amount of 11.6 L/sec. Note that this 11.6 L/sec is a conversion from number of molecules and is associated with the remote, upstream density.

The two geometries just mentioned occur in practice. Often the geometry represents a mixture of the two. For pipe flow and for a sequence of matched flow elements connected in series, the association of pumping speed at a cross section is more convenient. For a chamber with a small pumping port, the orifice association is more helpful. This duality of approaches is responsible for the difference

Fig. 2.



Basic geometry defining the concept of orifice conductance.

between the older pumping speed measurement standards (American Vacuum Society 4/1 and International Standards Organization 1607). The ISO standard and the new AVS Recommended Practice tend toward the orifice treatment and give 10–15% lower pumping speed values for a pump compared to the older AVS standard. The two can be related, depending on the ratio of pumping port to chamber diameter [4].

The net pumping speed of the pumping system or the pumping stack can be determined by the use of the same standard measuring dome. To relate this measurement to the vacuum chamber in order to obtain an effective speed for the chamber, we should take into account a pressure difference existing between the center of the chamber and the inlet into the pumping port. In practice, such considerations are not of great importance, and simple approximate calculations can be made by adding conductance in series as long as no abrupt changes in geometry are present.

In cases where pronounced anisotropy exists, pumping speed consideration can become meaningless. Consider an imaginary molecular beam with parallel molecular trajectories (without collisions) entering directly into a pump. A gauge mounted on the wall in the vicinity of this beam may not know of its existence. In such cases, the meaning of $Q = pS$ is not obvious. Situations of this nature may arise in practice.

2.1.5

PUMPDOWN TIME

One of the basic considerations in designing high-vacuum systems is the time required to evacuate a vessel to a given pressure. Usually, this ranges from a few

minutes to a few hours. However, in some cases, the desired time can be a few seconds or a few days. The prediction of time and pressure relationship by theoretical or experimental methods is very difficult because of the uncertainty associated with gas evolution rates from inner surfaces of the vessel (outgassing or virtual leak).

Gas evolution from a given material depends on temperature, surface finish, previous history of exposure to a variety of atmospheric conditions, and cleaning methods. Thus, even approximate prediction of evacuation time can be extremely difficult when outgassing becomes significant. Using well-known relationships applicable to higher pressures, the designer can predict a pressure–time curve from atmospheric pressure to about 10^{-2} torr. From steady-state (long time) outgassing data and from characteristics of pumping devices, the final (ultimate) pressures can sometimes be predicted. The region of greatest interest often is between these two points. The following discussion attempts to develop at least a qualitative appreciation of significant effects.

2.1.5.1 Constant Speed Case

In the pressure region between 760 and 10 torr, the pumping speed of the forevacuum pump is usually nearly constant. A 10% reduction at 10 torr compared to the speed near atmospheric pressure is common. Referring to Figure 3, the evacuation process can be represented by the following relationship:

$$-V dp/dt = Sp \tag{4}$$

where V is the volume of the vessel, dp/dt the rate of change of pressure with respect to time, and S the pumping speed. The physical meaning of Equation (4) can be understood by noticing that the left side represents the amount of gas leaving the chamber (minus sign indicates pressure decrease), and the right side shows the gas entering the pumping duct. It is important to note that a solution of this differential equation will depend on the location where the pressure is measured, p_1, p_2, or p_3. Otherwise, to be valid, the equation must have the same pressure value on both sides of the equality sign. Thus, the equation can be used as an abstract definition of a pumping speed for the given chamber, pumping system, and gauge location.

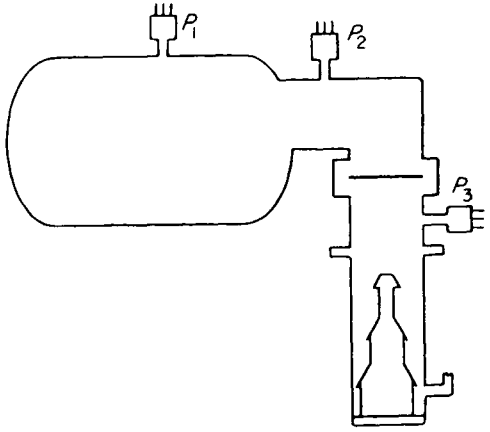
If this pumping speed is plotted versus the inlet pressure using log-linear coordinates, a straight line will be obtained in sections where speed is constant. The slope of this line represents the pumping speed, as related to the pressure reduction rate in the chamber ($S = \Delta \ln p/\Delta t$).

The evacuation time from Equation (4) for constant volume and speed is

$$t = (V/S) \ln (P_0/P), \quad \text{or} \quad t = 2.3(V/S) \log(P_0/P) \tag{5}$$

where P_0 is the initial pressure and P the final pressure. This equation gives adequately accurate values after a few seconds from the start of evacuation and until approximately 10 torr is reached. At lower pressures the outgassing becomes

Fig. 3.



A chamber with a pumping system showing three typical locations of vacuum gauges.

significant and the evacuation period is extended. Generally, it is recommended to multiply the values obtained from Equation (5) by about 1.5 between 10 and 0.5 torr, by 2 between 0.5 and 5×10^{-2} torr, and by 4 between 5×10^{-2} and 1×10^{-3} torr in order to accommodate the outgassing load.

For rapid pumping of small volumes (such as locks), the equation will not give accurate results due to complications of geometry, conductance and volume of pumping ducts, and transient effects within the pump itself. If the desired pumping time is less than 10 sec, experimental measurements may be advisable.

2.1.5.2 Constant Throughput Case

When constant throughput is used in Equation (4), the solution becomes

$$t = (V/Q)(P_0 - P) \tag{6}$$

This is not an interesting case technologically, but it occurs in a narrow pressure region, 10^{-1}–10^{-3} torr, in some pumping systems. This period is usually short, for example, less than 10 sec for a typical bell jar system.

2.1.5.3 In the Presence of a Leak

Equation (3) disregards a possibility of leaks and desorption gas loads. If $Q_\sim = Sp_\sim$ is the gas load after "infinite" pumping time, adding this term into Equation (3) leads to

$$-V dp/dt + Q_\sim = Sp_\sim \tag{7}$$

with a corresponding solution

$$t = (V/S)\ln(P_0 - P_\sim)/(P - P_\sim) \tag{8}$$

The gas load after a long pumping time $Q_\sim$ may be due to an actual leak, outgassing of almost constant rate, a finite permeability of the vessel walls, or any combination of these factors.

If, after a long period of pumping, the external leak is much greater than the surface outgassing, the pressure established in the chamber is simply given by the ratio of the leakage throughput divided by the pumping speed

$$p = Q_{\text{leak}}/\text{S torr L/s / L/s} \tag{9}$$

2.1.5.4 In the Presence of Outgassing

For qualitative purposes, the outgassing rate of a surface in high vacuum can be represented as

$$Q = Q_0 \exp(-t/\tau) \tag{10}$$

where Q_0 is an initial outgassing rate, t the time, and τ is associated with the slope (decay) of the outgassing function relative to time and is assumed to be a constant [5]. If necessary, for proper matching of experimental outgassing curves, two superimposed exponential terms can be used, such as $Ae^{-at} + Be^{-bt}$. Q is the total outgassing rate and can be expressed as qA, where A is the area of the outgassing surface. The general differential equation describing the evacuation of a vessel can be then written as follows:

$$-V dp/dt + Q_\sim + Q_0 \exp(-t/\tau) = Sp \tag{11}$$

The expression for the evacuation time becomes

$$t = \frac{V}{S} \ln \frac{(P_0 - P_\sim) - Q_0/(S - V/\tau)}{(P_1 - P_\sim) - [Q_0/(S - V/\tau)] \exp(-t/\tau)} \tag{12}$$

The expression for pressure decay relative to time will be

$$P = (P_0 - P_\sim)\exp(-St/V) + [Q_0/(S - V/\tau)][\exp(t/\tau) - \exp(-St/V)] + P_\sim \tag{13}$$

The first parts of the last two equations can be recognized as solutions when outgassing is disregarded.

In the high-vacuum region there does not exist a simple inverse relationship between pumping speed and time of evacuation, which is obtained at higher pressures. Using Equation (13) it can be demonstrated that for a common bell jar system, the pumping speed may have to be increased six or seven times in order to

cut the time in half. However, when process outgassing occurs, the large pump will maintain a lower process pressure inversely proportional to speed.

The outgassing rates of various materials can vary by many orders of magnitude. Adsorbing species such as water may exist in ordinary surfaces in the amounts of the order of 10 to 100 monolayers. The behavior of such films is highly dependent on temperature and binding energies involved. Outgassing rates of many substances may be considered to vary exponentially with temperature. The most disturbing species are those with desorption energies from 15 to 25 kcal/mole. Below that value, the outgassing proceeds rapidly. Above 25 kcal/mole, the desorption of the material in the vacuum space is likely to give partial pressures below 10^{-11} torr. The ordinary conductance relationships cannot be applied to gases in the temperature regimes in which they are adsorbed.

Returning to equation (13), it may be observed that in many common systems, for the part of the evacuation in the high-vacuum region,

$$P_{\sim} << P_0 \quad \text{and} \quad V/\tau << S$$

and the expression for the pressure reduction can be simplified to

$$P - P_{\sim} = P_0 e^{-(S/V)t} + (Q_0/S)(e^{-t/\tau} - e^{-(S/V)t}) \tag{14}$$

When the time of interest is typically between 1 and 10 hours of continuous pumping, we may say that $e^{-St/V} << e^{-t/\tau}$, and the solution simplifies further to

$$P = (Q_0/S)e^{-t/\tau} + P_{\sim} \tag{15}$$

This illustrates that the main effect on the rate of pressure reduction is due to the decay of the outgassing rate in respect to time. Q_0 is the initial outgassing rate, which can be associated with some significant starting time. One possibility is to use the time where the pressure versus time curve (on a log-linear plot) begins to deviate from a straight line. The onset of this deviation depends on the specific outgassing rate and the surface area exposed to vacuum. Thus, the evacuation process can be divided into two parts: the rough vacuum period, where the major influence is due to volume, and the high-vacuum period, where the dependence is mainly due to the internal surface area.

2.1.6 ULTIMATE PRESSURE

The ultimate pressure for a high-vacuum pump in a system is reached when further pumping will result in negligible reduction of pressure. The total ultimate pressure consists of partial pressures of various gases present in the system. In positive displacement pumps such as forevacuum pumps, the pumping speed for

various gases is approximately the same. However, under molecular flow conditions in a high-vacuum pump each species of gas is pumped with a different pumping speed, and the ultimate pressure is given by the sum of the individual partial pressures obtained by dividing the gas load (throughput) of the gas by its pumping speed. In addition, the ultimate pressure includes the component arising in the pump such as pumping fluid or lubricant vapors. Generally, the ultimate pressure of a pump is obtained when an equilibrium is established between the amount of gas flowing into the pump and the amount flowing in reverse (backstreaming).

2.1.6.1 Pump Limit Versus System Limit

A clear distinction should be made between an ultimate pressure attained by a pump alone and that attained when connected to a system. System gas loads often are orders of magnitude higher than the internal pump gas evolution, so the system may never reach the ultimate pressure capability of the pump. Normally, pump performance data are given for a minimum volume and surface area at the inlet of the pump. In addition to the amount of gas loads, the conductance of the inlet ducts will also influence the system ultimate pressure.

Ultimate pressure in a system is rarely an end in itself. In some work, a minimum partial pressure of one species, such as oil, may be critical, while much higher pressures of other species may be far less important. If certain work is to be performed inside the vessel after evacuation, it usually results in additional gas evolution. Therefore, if a process or an experiment is to be performed at a given pressure, the ultimate system pressure should be lower. A convenient and very approximate rule for work that produces low gas evolution is to specify ultimate system pressure 10 times lower than the desired process environment.

In general, the inlet pressure of a high-vacuum pump can be expressed as a summation of effects of various gas flows and compression ratio limits,

$$P_1 = \left(\Sigma \frac{Q_i}{S_i}\right)_{\text{ext}} + \left(\Sigma \frac{Q_i}{S_i}\right)_{\text{int}} + \Sigma \frac{P_{2i}}{K_i} \tag{16}$$

where P_1 is the inlet pressure, Q_i is the gas flow rate (throughput) for a particular gas, P_{2i} is the partial pressure of the particular gas at the discharge end of the pump, and K_i is the limiting compression ratio for the particular gas; the first term represents external gas loads; the second term, gas loads from sources inside the inlet sections of the pump itself. What is generally taken to be the ultimate pressure of the pump (not the system!) is given by the last two terms of the summation. However, it cannot be emphasized too strongly that the manufacturers and users of vacuum pumps traditionally have ignored the last term for nearly

100 years. This is certainly true for mechanical pumps and vapor jet pumps (so-called diffusion pumps). The compression ratio effects and limits are typically only discussed with the more recently developed turbomolecular pumps. For capture pumps, the compression ratio term must be replaced by a term associated with the number of desorbing previously pumped molecules.

Without knowing the gas composition and the compression ratio limits for each gas present in the system, it is generally not possible to know what determines the ultimate pressure in a vacuum system.

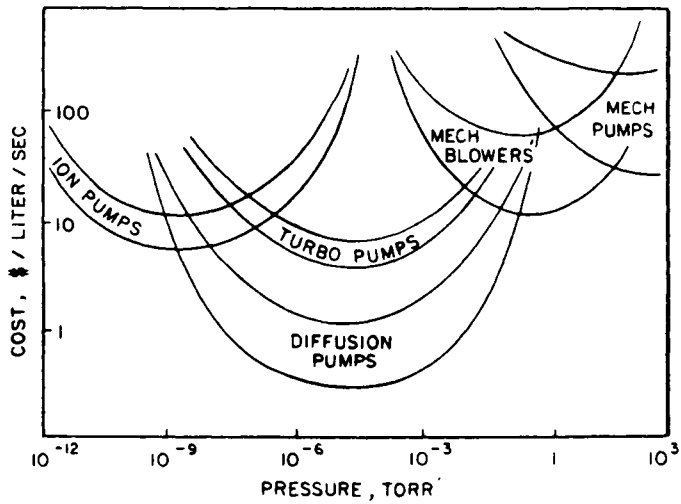
Because of the enormous range of densities involved in evacuating a vacuum chamber from atmospheric pressure to ultra-high vacuum, it is not possible, for practical reasons, to do the required pumping with one integral pump. Typically, at least two pumps are connected in series. To understand the interaction between those pumps in regard to its effects on inlet pressure, it is necessary to know not only the total pressure established in the conduit between the two pumps but also the partial pressures of various gases. Yet even if a partial pressure analyzer were used at the discharge side of the pump, it is not simple to make definitive conclusions from the known pump performance parameters.

Most high-vacuum pumps do not operate entirely in the molecular flow range. Therefore, gas interactions can occur in the downstream parts of the pump, which are not easily predictable from the primary performance data. Even if compression ratios are known for the particular pump and gas, the presence of a second gas can alter the expected compression limit values. Therefore, it is very important for a high-vacuum technologist to be aware of various effects not normally given in the technical bulletins or textbooks on high-vacuum technology, and to understand qualities and peculiarities of a particular pump and its construction details. The traditional display of pump performance does not represent the entire story. It does not, for example, give any information about the influence of the discharge pressure on the performance of the pump.

2.1.7 FOREVACUUM AND HIGH-VACUUM PUMPING

High-vacuum technology deals with a very wide range of pressure or particle density conditions. Usually, the process of evacuation begins at atmospheric pressure and then proceeds to high or ultra-high vacuum. In a multistage pumping system, the type of pumping mechanism employed is different at atmosphere and at high vacuum. Even if, in principle, a given pumping method could be used throughout the entire pressure range, such an attempt would be impractical in regard to size, weight, or cost of equipment, as illustrated in Figure 4.

Fig. 4.



A qualitative illustration of the pressure ranges in which the use of various pumps is most economical.

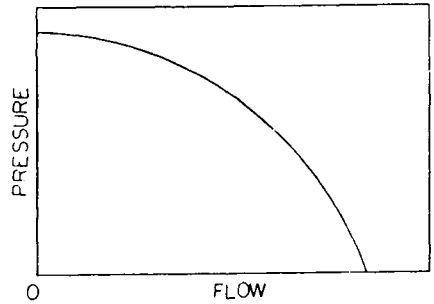
The most common pumping arrangement for production of high vacuum consists of a positive displacement mechanical pump for initial evacuation followed by a high-vacuum pump. Often parallel and series arrangements are used. For initial evacuation, the mechanical pumps are used alone, and for obtaining high vacuum, the two types of pumps are connected in series. In such a system, the gas enters the pumping train at high vacuum and is exhausted at atmospheric pressure by the last pump. In some cases, the device may be evacuated and sealed, terminating the pumping process; in others, the pumping is continuously applied to compensate for the gas evolution in the vacuum chamber.

In capture pumps (sorption pumps, and particularly in getter-ion pumps), the pumped gas is not exhausted to atmosphere. This has an obvious advantage of isolation from high-pressure environment and the disadvantage of limitation in gas load capacity or the necessity for periodic regeneration or replacement, in the case of getter-ion pumps.

The basic performance of pumps and compressors can be associated with flow and pressure factors. With appropriate allowance for size, power, pumped fluid characteristics, etc., all such devices generally behave according to the pressure flow graph shown in Figure 5.

For high-vacuum work, both high-volume flow rate and high-pressure ratio (inlet to discharge) are necessary. The flow rate is associated with the size of the pumping device. Thus, for a given size, the pressure ratio needs to be increased as much as possible. Mechanical vacuum pumps, vapor jet pumps, and turbomolecu-

Fig. 5.



The general qualitative relationship between pressure and flow for most compressors or pumps.

lar pumps produce pressure ratios of over 1 million to 1. This can be compared to industrial plant air compressors, automobile engines, or aircraft compressors not exceeding a compression ratio of 10 to 1. Obviously, high-vacuum pumps require very special designs, and familiarity with their design, construction, operation, and maintenance is an important ingredient of success in producing and using high-vacuum environments.

2.1.8

PUMP SYSTEM RELATIONSHIPS

There are two basic ways of reducing the pressure or time needed for evacuation of a chamber: increase the pumping speed or decrease the outgassing rate. In some cases, it may be less expensive to provide additional pumps rather than attempt to reduce gas evolution. This depends on the degree of vacuum and other process requirements. For very-low-pressure or ultra high vacuum systems, degassing is the only approach possible. Checking the following points is useful before specifying or designing a vacuum installation: The gas-handling capacity of the pumping system at a given pressure is usually more important than the ultimate pressure; the net pumping speed at the chamber is more significant than the nominal speed of the pump; increasing the number of pumps in the high vacuum region does not necessarily decrease the evacuation time in the same proportion; well-chosen pumps, working fluids, baffles, and traps can reduce or eliminate contamination problems: rapid pumping requirements increase system cost; long exposure of the chamber to atmospheric conditions will increase the subsequent outgassing rate; temperature history and distribution in the chamber before and during evacuation should not be neglected; the amount of surface area present in the vacuum chamber is more significant (at high vacuum) than the volume.

2.1.8.1 Surface-to-Volume Ratio

Generally, in high-vacuum conditions more gas is adsorbed on the walls of the vacuum chamber than present in the space. This, of course, depends on the size of the chamber. For small devices, the surface-to-volume ratio is higher, and small vacuum chambers tend to have a higher ratio of surface area to pumping speed. Therefore, generally 5-cm-diameter high-vacuum pump systems, for example, do not produce low ultimate pressure as easily as 10- or 20-cm-diameter pumping systems.

A similar comment may be made about other sources of gas. Pumping speed grows approximately with the square of the pump diameter, but the exposed area of gaskets, and therefore the associated outgassing load, increases closer to a linear rate. Thus it is relatively easy to produce high vacuum in large chambers (as high as 10 m in diameter). In such cases, the rough vacuum system may be the expensive part, because of its dependence on volume.

In laboratory system, pumps of 15–20 cm diameter are usually more convenient for low-ultimate-pressure work even if the high-speed requirements are not necessary. In normal practice, pumps are described by the size of the flange at the opening.

2.1.8.2 Evacuation or Process Pumping

Vacuum pumps are used for two somewhat distinct applications. The first involves evacuation of gas (usually air) from a vessel, the other maintaining process pressure with a given gas load. It is difficult to choose the correct pump size with any degree of precision for both applications when required pressure is in the high-vacuum region. In the first case, this is due to the uncertainty of surface outgassing rates, which depend on temperature, humidity during atmospheric exposure, and the presence of gas adsorbing deposits on the chamber walls. The water vapor is the usual cause of concern. In the second case, the process gas evolution is uncertain, due to variation of gas content in materials and the variations in the process parameters. For example, when electron beams are used for evaporation the power distribution in the focal spot may not be exactly repeatable. This may produce variations in the degree of heating of the surfaces in the vicinity of the electron gun.

As a very rough rule for evaporation, melting, welding, or sputtering of metals, at least 100 L/sec of net pumping speed should be provided for every kilowatt of power used for the process.

Generally, one can hardly make an error in the direction of using too large a pump. It is common to use 25-cm vapor jet pumps in a large bell jar system to

produce a net system speed exceeding 1000 L/sec. Thirty years ago, the net pumping speed at the base plate was only 500 L/sec. These differences are not important in improving system pumpdown time but they do help to maintain lower process pressures. As far as evacuation time is concerned, the difference between a 15- or 25-cm pump is hardly noticeable. When outgassing is taken into account, the evacuation time is not inversely proportional to the pumping speed, as might be expected from the experience at higher pressures. However, during rapid process gas evolution or introduction, the process pressure and the pumping speed are often related inversely.

In some cases, such as sputtering, maximum throughput becomes important, because it may be desirable, from the process viewpoint, to have a certain flow of argon through the system. A high pumping speed for the initial evacuation of the system will be of lesser importance. In some instances, a system is pumped with the pump valve fully open to a given base pressure and then throttled to obtain an argon pressure and flow rate as required by process specifications, without exceeding the tolerable throughput of the pumps. As the pump speed has been reduced, it is possible that the impurity level in the system has increased.

2.1.8.3 Pump Choice

A variety of mechanical pumps are used at the present time for the initial evacuation of a vacuum vessel. Most common are rotary vane and rotary piston positive displacement pumps, in which the narrow spaces between moving and stationary parts are sealed by oil. The presence of sealing and lubricating oil reduces the internal leakage from discharge side to the inlet side of the pump. The oil also fills the so-called dead space under the discharge valve. This permits very high compression ratios to be achieved in such pumps. They obtain inlet pressures of 10^{-2} torr in a single stage and better than 10^{-4} torr in two stages. When the presence of oil is undesirable, a variety of oil-free pumps can be used. These are usually limited to a lowest inlet pressure of about 10^{-2} torr. Such pumps can be multistaged piston pumps, multistaged Roots type (or other cam-type) blowers, orbiting scroll pumps, or screw pumps.

To achieve high vacuum after the initial evacuation, at the present time four different type of pumps are used: getter-ion (or sputter-ion) pumps, cryopumps, oil vapor jet pumps (or diffusion pumps), and turbine-type pumps of various designs.

1. *Getter-ion pumps* function by providing chemical reaction with freshly evaporated or sputtered metallic films (usually titanium) and by ionization and subsequent capture of the gas molecules in the surface layers of the electrodes. They are most suited for systems that are not frequently opened to

atmosphere. Because they are capture pumps, they cannot accept large gas loads. They are extremely useful for pumping of vacuum tubes and long-term experiments. They are clean and have no oil contamination.

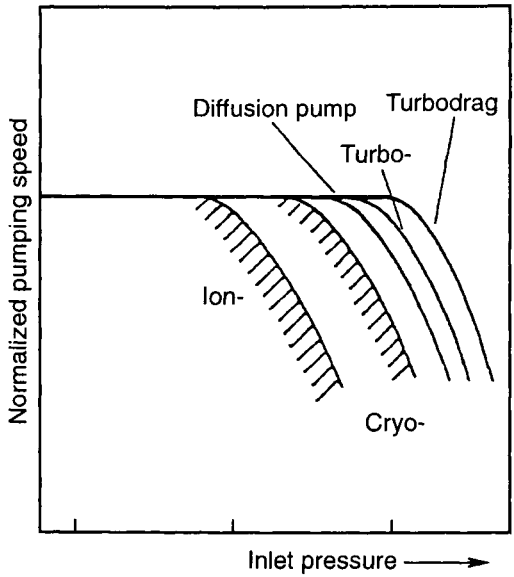
2. *Cryopumps*, as the name implies, immobilize the gas on cold surfaces by condensation and cryosorption. Cryopumps are oil free and have high pumping speeds, particularly for water. They are capture pumps and consequently require frequent regeneration to dispose of the captured gases. This can pose a problem when used with hazardous gases, because a significant accumulation of an explosive or toxic gas is difficult to deal with in an emergency shutdown.
3. *Vapor jet pumps* function by moving gas molecules from the inlet to the discharge side through collisions with a high-velocity stream of vapor (usually a variety of low-vapor-pressure oil-like substances). The advantage of a vapor jet pump is that it can pump a very wide variety of gases, including rare gases, such as helium, that are difficult to pump using cryopumps or ordinary turbopumps. Vapor jet pumps with very high pumping speeds can be obtained. The major disadvantages are that if improperly operated, oil vapor can backstream into the system. Cold traps are used to prevent this migration and have an advantage of increasing the pumping speed for water.
4. *Turbine-type pumps* produce the pumping action by collisions with a high-velocity moving surface (blade rotors or plain disks or cylinders). These pumps are generally clean and oil free, and provide high pumping speeds for most gases. Due to their precision construction, these pumps are particularly susceptible to particle contamination. A turbopump can be turned on or off faster than both the cryo and diffusion pumps. Hence its use may be preferred in applications sensitive to pumping system failures.

Figure 6 shows an approximate range of inlet pressures where such pumps can be used for continuous operation. To achieve oil-free pumping, cryopumps or grease-lubricated or magnetic bearing turbopumps backed by dry roughing pumps are normally employed.

The selection of high-vacuum pumps is not as straightforward as selection of mechanical pumps. The considerations for reaching a given pressure in a short time may not prove an adequate measure of the system's capability to handle steady gas load. Characteristics of pumping curves of high-vacuum pumps indicate the proper components. More frequently, gas loads are not known, and the selection of the pump must be based on experience with similar applications.

The vacuum equipment suppliers generally offer performance based on pumpdown characteristics of an empty chamber. This can be misleading if it is construed to be a measure of pumping efficiency under full load conditions. The vacuum equipment designers must relate performance to a set of conditions that are known and can be defined and duplicated, which may have misled some users to construe the pumpdown time to be the true indicator of process time.

Fig. 6.



Comparing overload regions for various pumps. The shading for ion and cryopumps reflects variations in mass flow.

The selection of high-vacuum pumps should be concerned with all factors involved in the system. These factors normally consist of pumping speed, forepressure tolerance, backstreaming rates, pressure at which maximum speed is reached, protection devices, ease of maintenance, backing pump capacity, roughing time, ultimate pressure of the roughing pump, baffles or cold traps employed, valve actuation, etc. Unless these factors are carefully weighed in the system design, the system may not perform satisfactorily as a production system. It is to the advantage of both the supplier and the user to carefully consider these factors before finalizing system design.

It is obvious from the previous discussion that pump overloading should be avoided in system operation. As a general rule, high-vacuum pumps should not be used at inlet pressures above the specified maximum value for prolonged periods of time. In a properly designed system with a conventional roughing and high-vacuum valve sequence, the period of operation above this value should be measured in seconds. As a "rule of thumb," if this period exceeds half a minute, the pump is being overloaded [5].

It is clear then that in applications requiring only 10^{-4} torr process pressure, the mechanical pump roughing period is normally much longer than the high-vacuum pump portion of the pumpdown cycle, or at least it should be, for economic reasons.

2.1.9

CROSSOVER FROM ROUGH TO HIGH-VACUUM PUMPS

During the evacuation of a vessel, the question arises regarding the proper time to crossover (switch) from rough pumping to the high-vacuum pump or, in other words, when the high-vacuum valve should be opened. Ideally, the answer should be: when the gas flow into the high-vacuum system is below the maximum throughput of the next pump. In practice, the transfer from roughing to high-vacuum pumping is normally made between 50 and 150 mtorr. Below this pressure region, the mechanical pumps rapidly lose their pumping effectiveness and the possibility of mechanical pump oil backstreaming increases. Although the throughput of the high-vacuum pump is nearly constant in the region of inlet pressures of 1–100 mtorr, the initial surge of air into the pump, when the high-vacuum valve is opened, will overload the high-vacuum pump temporarily. [6]

The general recommendation is to keep the period of inlet pressure exceeding approximately 1 mtorr (0.5 mtorr for large pumps) very short, a few seconds, if possible. Consider the following example. For constant throughput, the evacuation time (between 100 and 1 mtorr) is

$$t = (P_1 - P_2)V/Q$$

(See Equation (6), Section 2.1.5.)

For a common 45-cm-diameter bell jar, the volume is about 120 L and, using a high-vacuum pump with a maximum throughput of about 3 torr L/sec, we obtain

$$t = 120 \text{ L } 0.1 \text{ torr}/3 \text{ torr L/sec} = 4 \text{ sec}$$

If the high-vacuum valve were to open slowly to admit the gas into the high-vacuum pump at the maximum throughput rate, it would take only 4 sec to reach the stable pumping region. This transition occurs with a time constant

$$\tau = V/S$$

where V is the volume of the chamber and S the pumping speed of the high-vacuum pump. In typical vacuum chambers, this time constant is less than 1 second. On the other hand, the decay function associated with the outgassing rate has a time constant of minutes or hours. For purposes of this discussion, the outgassing rate can be assumed to have a quasi-steady-state constant value. Thus, we may say that

$$Q = PS_{\text{net}} = Q' = P'S'_{\text{net}} = P_2S_2 \tag{17}$$

where: P is pressure before crossover
S_{net} = the net rough pumping speed near the chamber
P' = pressure after crossover

S'_{net} = the net high-vacuum pump speed
P_2 = discharge pressure at the high-vacuum pump (which must never rise above the tolerable forepressure)
S_2 = backing speed at the outlet of the high-vacuum pump

The crossover pressure, then, is

$$P = Q_{max}/S_{net} \tag{18}$$

where Q_{max} is the maximum mass flow capacity of the high-vacuum pump (maximum throughput) and S_{net} is the net pumping speed of the rough pump at the chamber.

It may be expected that the effect of outgassing will increase when the pressure drops after crossover. Usually this is not significant because the interdiffusion of gases, at pressures considered here, is very high.

Traps are often employed to minimize oil backstreaming from the roughing pump to the system during roughing and to the high-vacuum pump during the "backing" stage. Traps are typically located in the foreline and provide a large area surface on which the backstreaming fluid is adsorbed. To be effective, foreline traps require regular maintenance. Adsorbed matter must be removed by baking or cleaning. An overloaded foreline trap simply moves the source of contamination from the roughing pump to the trap itself.

2.1.10 PUMPING SYSTEM DESIGN

High-vacuum pumps are used, together with mechanical pumps, in applications where system operating pressures of 10^{-3} torr and below are desired. The physical arrangement of system components depends on the characteristics of the process to be carried out, such as the pressure level, cycle time, cleanliness, etc. To some extent, the availability and compatibility of components influence the system design. In some instances, the economic aspects of component selection may determine the system layout. The following paragraphs illustrate briefly the most common component arrangements, referred to as valved and unvalved systems, and outline their major respective advantages and disadvantages. A recommended operating procedure to ensure minimum work chamber contamination is outlined for each type.

To furnish maximum effective pumping speed at the processing chamber, it is generally desirable to make the interconnecting ducting between the chamber and the pump inlet as large in diameter and as short in length as practical. The amount of baffling, trapping, or throttling required depends on the desired level of process gas flow, cleanliness in the chamber, the necessary reduction of the inherent back-

streaming characteristics of the pump, and the migration of the pumping or lubricating fluid, or reemission of previously pumped gases.

The size of roughing and foreline manifolding is governed by the capacity of the mechanical pump, the length of the line, and the lowest pressure region in which it is expected to function effectively. In addition, the size of the foreline is influenced by the forepressure and backing requirements of the high-vacuum pump under full-load operating conditions.

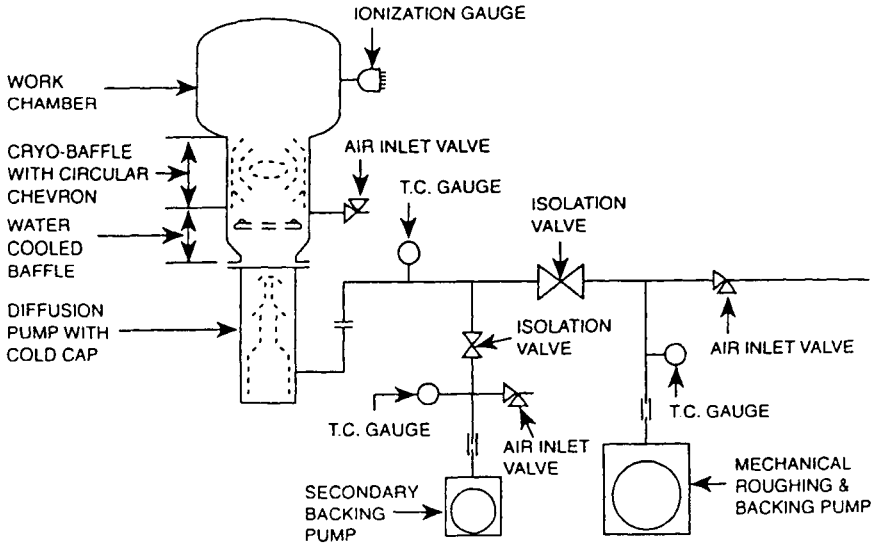
The need for a holding pump, which can be used to maintain the forepressure of the high-vacuum pump during system roughing, is determined by several factors, including the forepressure characteristics of the high-vacuum pump, time cycle requirements, manifold configuration, and leak-tightness. The selection of the type and capacity of mechanical pump will depend on the desired operating cycle, the throughput and forepressure requirements of the pump, and the proposed operating procedures for a system pumpdown.

Usually, in a high-vacuum system that is expected to be pumped to 10^{-6} torr in less than half an hour, the outgassing will be negligible (compared to the maximum throughput of the pump) during the initial period after the high-vacuum valve is opened. In practice, due to lack of precise valve control, the period between the time when the valve is opened and the time when inlet pressure of 1 mtorr is reached can be shorter than that computed using Equation (6). This is due to the expansion of air across the high-vacuum valve into the higher-vacuum space downstream. This downstream space (part of the valve, trap, baffle, upper part of the pump) can be significant compared to the chamber volume. On the other hand, when outgassing is severe the pump may be overloaded for longer periods of time. The same is true if the pump is too small for the chamber volume. This can be encountered in furnaces having porous insulating materials and in coaters where large drums of thin film plastic are present in the vacuum chamber. In such cases, it may be better to extend the rough pumping period before returning to the high-vacuum pump. When throughput is nearly constant, the pumping time will be the same whether the high-vacuum valve is only slightly cracked or fully opened. In general, the high-vacuum valves should be opened slowly; very slowly at the start to minimize turbulence and particulate generation. The motion of ordinary pneumatic valves can be controlled to some extent with the air inlet and exhaust adjustments. Special valve controls can be made either to maintain approximate constant throughput during initial opening or to have a two-position interrupted operation. When the valve is almost closed, it serves as a baffle.

2.1.10.1 Valveless Systems

Valveless vacuum pumping systems are generally considered for applications where the length of the pumpdown is of less importance, and process or test cy-

Fig. 7.



A vacuum system arrangement without a high-vacuum valve.

cles are of extended duration. They are found on large, baked, ultra-high vacuum chambers, due to a lack of availability and/or prohibitive cost of valves compatible with the proposed operating pressure levels. However, analytical systems operating in UHV are bakeable and use gate valves very successfully. The valveless system (Figure 7) generally offers higher effective pumping speed at the chamber and, in view of its prolonged operation at very low throughputs, lends itself to the use of a smaller holding pump. This pump is sized to handle the continuous throughput of the high-vacuum pump at low inlet pressures, 10^{-6} torr and below. The valveless system, however, also has a number of disadvantages. More complex operating procedures are necessary to ensure maximum cleanliness in the work chamber and minimum contamination from the pump.

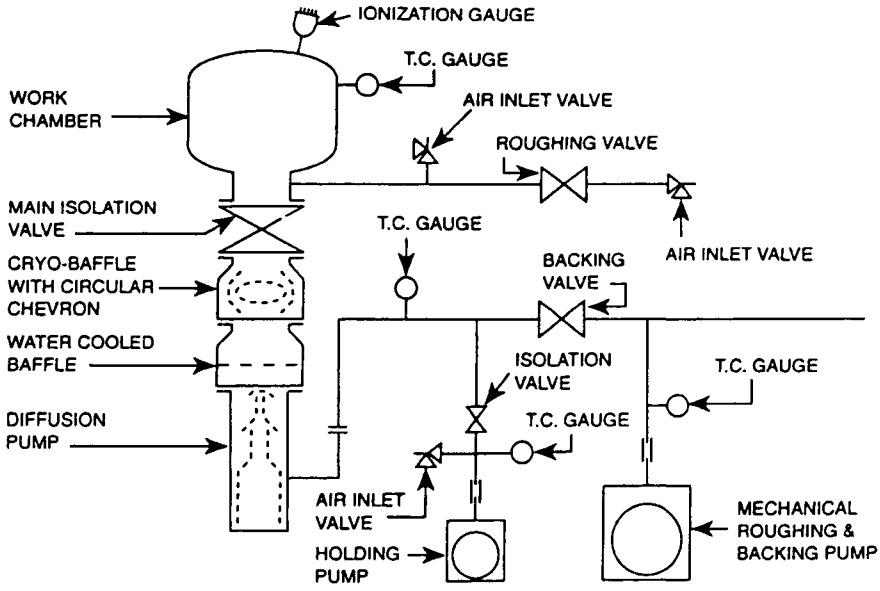
To minimize chamber contamination from vapor jet pumps, it is recommended to keep at least one of the baffles or traps operative while the vacuum chamber is being baked. When the inlet ducts are at the elevated temperature, the rate of pumping fluid migration into the system is accelerated. At temperatures near 200°C, the oil vapors are not condensed in the chamber and are subsequently pumped by the vapor jet pump. They can be adsorbed, and so, if cleanliness is essential, traps must be used during baking and especially during the cooldown, when a very thin (invisible) film of oil will condense in the chamber. For example, electron multipliers are very sensitive to such adsorption. This film formation is minimized by continuous operation of at least one of the baffles.

The vapor jet pump must be cooled down to a safe level before the chamber can be repressurized. This makes servicing cumbersome and time consuming. Failure of utilities necessitates immediate shutdown of the system to protect the workload and equipment. Operating costs of cryogenic baffles are higher, leak testing and leak hunting are less convenient and more time consuming.

2.1.10.2 Valved Pumping System

For applications involving rapid recycling, a fully valved pumping system is essential. This type of system is shown schematically in Figure 8. It permits isolation of the high vacuum from the work chamber at the conclusion of a pumpdown and prior to air admittance. The pump can, therefore, remain at operating temperature and pressure during periods when the chamber is at atmosphere and during the rough pumping portion of the cycle. The length of these periods may indicate the need for a holding pump. The main isolation valve also permits the continuous operation of cryobaffle between it and the high vacuum pump inlet. Neither this, nor rapid cycling can be realized without the valving indicated, in view of the cooldown and heat-up time lapse inherent in vapor jet pump operation, and

Fig. 8.



A schematic arrangement of a typical vacuum system with a high-vacuum valve.

cooldown and reheat time of the cryobaffle. Judicious operation of the main valve at the changeover phase from roughing to vapor jet pumping can significantly reduce the backstreaming of oil vapors to the work chamber. Valved systems are generally confined to operating pressures in the 10^{-8} torr range and above. Most large commercially available valves contribute too high a gas load to the system to allow operation at pressures lower than 10^{-8} torr.

Leak testing and leak hunting are considerably easier in valved systems, and repair procedures are also generally less time consuming than in unvalved systems. However, the following disadvantages are noted. Valved systems are initially more expensive, especially when large valves are involved. In addition, the use of valves inevitably adds to the system complicity and generally results in lower effective pumping speed at the chamber. For operation below 10^{-8}–10^{-9} torr and for use with large, baked, ultra high vacuum chambers, the availability and/or cost of valves may make their use prohibitive.

REFERENCES

1. S. Dushman and J. Lafferty, *Scientific Foundations of Vacuum Technique* (John Wiley & Sons, New York; London, 1962, p 90).
2. G. Lewin, *Fundamentals of Vacuum Science and Technology* (McGraw-Hill, New York, 1965).
3. F. J. Schittko and S. Schmidt, *Vak. Tech.*, **24** (1975) 4.
4. H. G. Nöller, *Vacuum* **13** (1963) 539.
5. M. H. Hablanian, *High Vacuum Technology* (Marcel Dekker, New York) 1990; 2nd edition 1997.
6. M. H. Hablanian, *J. Vac. Sci. & Technol.*, A10(4) (1992).

CHAPTER 2.2

Diaphragm Pumps

F. J. Eckle
VACUUBRAND GMBH + CO

2.2.1 INTRODUCTION: BASICS AND OPERATING PRINCIPLE

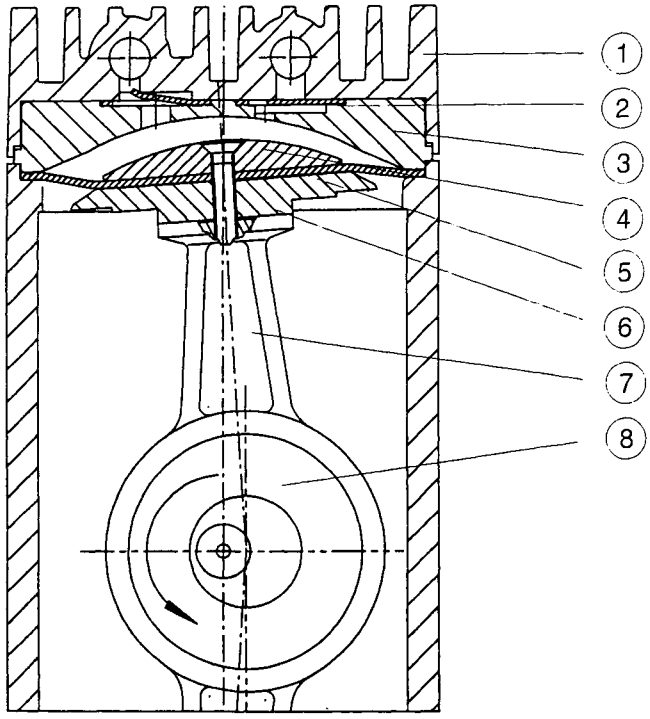
Modern diaphragm pumps have become well-established devices for the generation of rough and fine vacuum. Physical, technical, and economical challenges have led to diaphragm pumps covering ultimate vacua from 100 torr to 0.1 torr with pumping speeds up to 200 L/min. Diaphragm pumps, in contrast to other mechanical vacuum pumps, can be made of materials with high resistance against chemical attack. Therefore, one of their major applications as stand-alone pumps has become vacuum generation in "chemical laboratories." As backing or auxiliary pumps for modern oil-free high-vacuum pumps, they open the gateway to oil-free high and ultra-high vacuum. Numerous applications result in the coating and semiconductor industry, vacuum metallurgy, and analytical instruments business.

Figure 1 shows the setup of a modern diaphragm pump. The pumping chamber is defined as the volume between the cylinder head and the diaphragm, which is attached to the connecting rod by means of the diaphragm clamping disc. Due to the movement of the connecting rod, the size of the actual volume of the pumping chamber is altered periodically from expansion to compression. The connecting rod is attached to a crankshaft driven by the motor. The inlet and outlet valves are located between the head cover and housing cover. These are gas-flow-operated reed valves.

An enlargement of the pumping chamber of a rotational positive displacement

ISBN 0-12-325065-7
$25.00

Fig. 1.



Schematic section drawing of a diaphragm pump stage: (1) housing, (2) valves, (3) head cover, (4) diaphragm clamping disc, (5) diaphragm, (6) diaphragm supporting disc, (7) connecting rod, (8) eccentric bushing.

pump (e.g., rotary vane pump) can be accomplished by scaling the dimensions so that in fact the price per L/min decreases with increasing pumping speed. In contrast, dimensional scaling of a diaphragm pump would lead to technical problems—e.g., in manufacturing—and would reduce the lifetime of diaphragms. Therefore, demands for a higher pumping speed are met by connecting cylinder heads in parallel.

The application of diaphragm pumps is mainly restricted to laboratory scale due to dimensional and economical reasons.

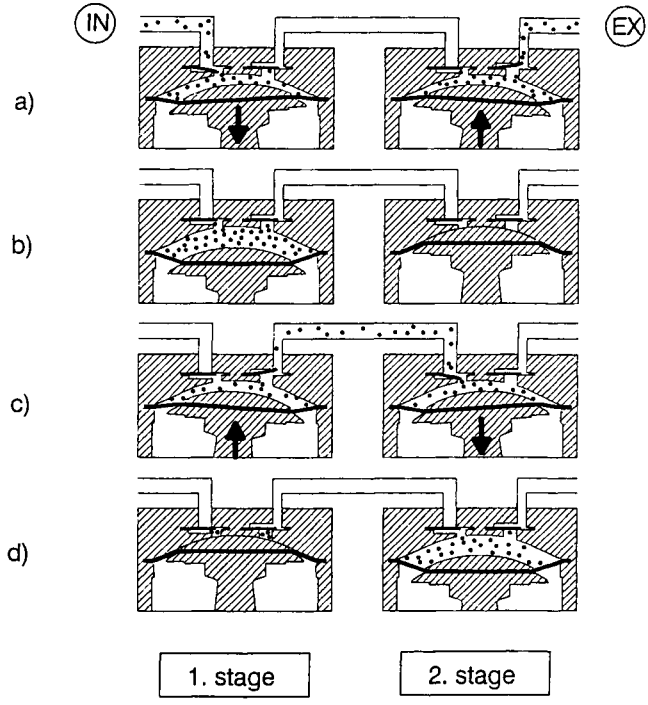
The compression ratio of a single stage of a diaphragm pump typically amounts to greater than 10. Single-stage diaphragm pumps therefore have an ultimate vacuum around 70 torr. For lower ultimate vacua, cylinder heads are connected in series. The ultimate vacuum that can be attained is limited by the performance of the reed valves. Depending on the dynamic response of the individual valve, correlated with its mass, geometry, modulus of elasticity, and the driving gas flow,

this limit is near 1 torr. Therefore, only serial connections of up to four stages are of practical use.

In multistage diaphragm pumps, the outlet of one stage is connected with the inlet of the next stage. The connecting rods of two stages connected to one crankshaft have in general a phase shift of half a cycle: During expansion in the first stage, the second one compresses, and vice versa. This design is convenient but not indispensable. The stages could be arranged as well in V-shape or line configuration.

The operating principle of a diaphragm pump, the opening and closing of the valves as well as the mechanism of gas flow, is shown in Figure 2. The inlet of the first stage is the inlet of the pump. Figure 2(a) displays the first stage during aspiration (expansion of gas). As the volume of the pumping chamber increases, the inlet valve is opened as a result of the pressure difference between the volume to be evacuated and the pumping chamber, and gas flows into the pump. The diaphragms have two momentarily stationary points, at maximum and at minimum pumping chamber volume, as shown in Figure 2(b). At these stationary points

Fig. 2.



The operating principle of a two-stage diaphragm pump: opening and closing of the valves, mechanism of gas flow from inlet to outlet during one pumping cycle.

there is no gas flow into the pump—the valves are in an indifferent equilibrium. The pressure in the volume to be evacuated corresponds to the pressure in the first stage, and the pressure in the second stage to that at the exhaust of the first stage.

The pumping process continues as displayed in Figure 2(c). The pumped amount of gas is transferred from the first stage to the second. The diaphragm in the first stage moves in the opposite direction and the inlet valve is closed. The outlet valve of the second stage is closed due to the atmospheric pressure. The pressure difference between the two stages keeps the outlet valve of the first stage and the inlet valve of the second stage open, and therefore the gas flows from the first into the second stage. Figure 2(d) represents the mirror image of Figure 2(b) with regard to the functional characteristics of the pump stages.

2.2.2 STATE-OF-THE-ART DESIGN AND MANUFACTURING

The enclosed dead volume in the pump chamber (minimum volume of the pumping chamber at the corresponding stationary point) limits the attainable ultimate vacuum of a single pumping stage. In addition to that limit, which is correlated with the theoretical compression ratio, in practice the ultimate vacuum is often limited by an external leakage or internal backstreaming due to a time delay in closing of the valves with regard to the corresponding stationary point.

The compression ratio is given as the ratio of the pressure at the outlet, to the pressure at the inlet of a pump stage:

$$k = p_0/p_i$$

The theoretical compression ratio, in the case of zero gas flow, is defined as

$$k = V_p/V_D$$

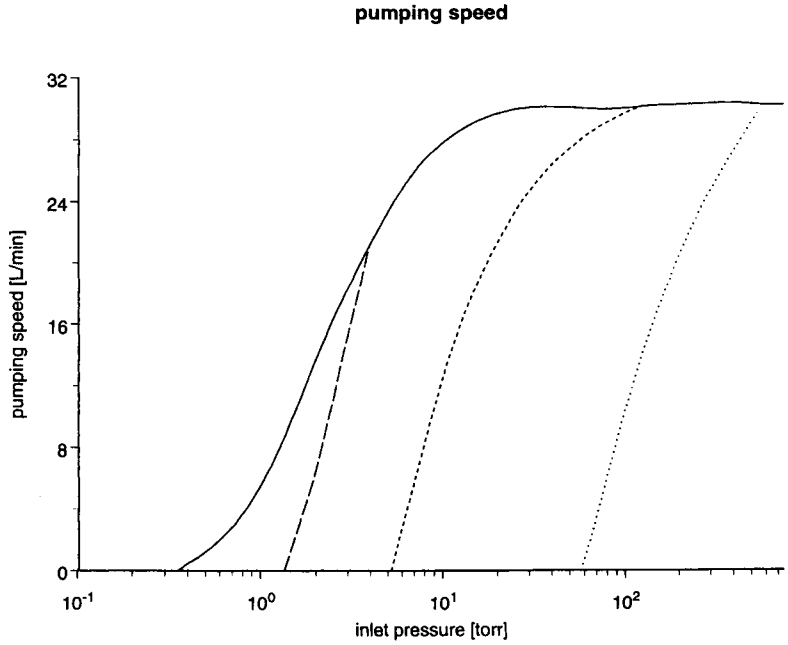
where V_p = volume of the pumping chamber
V_D = dead volume

Typical values of k range from 10 to 30.

Pumping speed per pump head is typically in the range of 5 to 40 L/min. By connecting up to 8 medium-sized pump heads in parallel, pumps with a pumping speed of up to 320 L/min are made.

Figure 3 shows pumping speed curves obtained for nitrogen using diaphragm pumps with one (dashed), two (dashed-dotted), three (long dashed), and four (solid) stages. These pumping speed curves were obtained according to the pump-down method (standard DIN 28432: acceptance specifications for diaphragm pumps). This method offers instrumental advantages and yields the same values

Fig. 3.



Pumping speed curves of diaphragm pumps with one (dashed) to four (solid) stages.

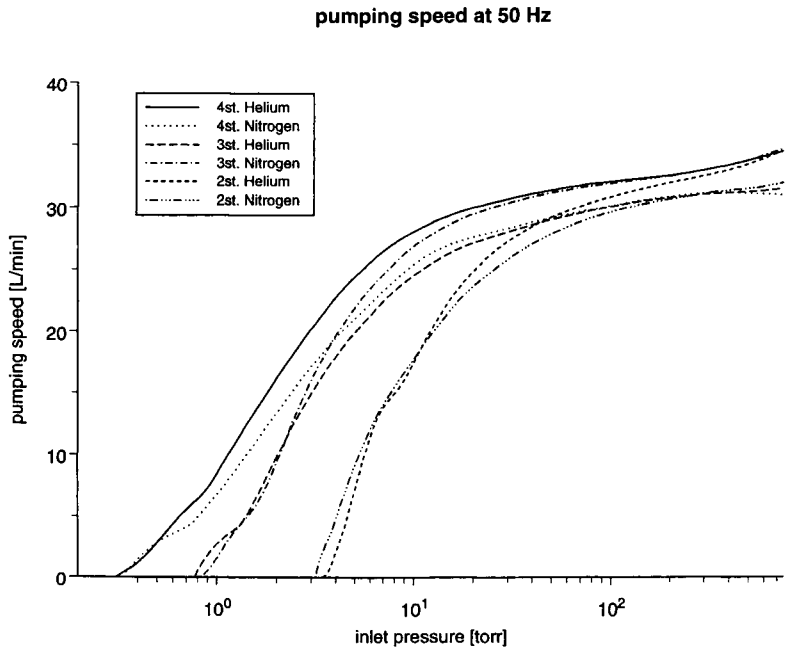
of the pumping speed when compared to results obtained with measurements according to the Pneurop [1] standard for oil-sealed rotary vane pumps or roots pumps within the experimental uncertainty of about 3%. [2]

Figure 4 displays the corresponding pumping speed curves for a four-stage, three-stage, and two-stage diaphragm pump for nitrogen and helium versus inlet pressure. The pumping speed of diaphragm pumps depends linearly on the displaced volume defined by the geometry of the chamber itself. In a first approximation, it is expected to be the same for all gases. In fact it is found to be dependent on the type of gas. The observed deviations can be related to influences of gas dynamic effects due to laminar and turbulent flow, and thermal effects due to cooling of the gas entering the pump as well as to the different thermal conductivities of the gases. [3]

A high external leakage rate can limit the ultimate vacuum of multistage diaphragm pumps. Leakage rates in the order of magnitude of 10^{-5} torr L/sec may be achieved by appropriate design but do not necessarily result in improvement of the ultimate vacuum, which is restricted by the dynamic response of the valves.

Being totally oil free and made of materials with high resistance against chemical attack, diaphragm pumps can be used for pumping aggressive and condens-

Fig. 4.



Pumping speed for a four-stage (4 st.), three-stage (3 st.) and two-stage (2 st.) diaphragm pump versus the inlet pressure for helium and nitrogen.

able solvents. Large amounts of condensates may cause mechanical damage and thus reduce the lifetime of diaphragms and valves. It is therefore recommended that a gas ballast be used on chemical versions of diaphragm pumps. The gas ballast reduces the formation of condensates and may purge formed condensate out of the pump.

The low noise level of diaphragm pumps, typically below 50 dBA, favors the diaphragm pump for use in laboratories.

Outstanding benefits, such as comparatively high pumping speed, low ultimate vacuum, long lifetime of diaphragms and valves, as well as easy maintenance and favorable prices of modern diaphragm pumps, are based on CAD design, calculations on stability and strength properties using the finite element method, CNC (computerized numerical control) manufacturing, and the use of modern materials.

Early diaphragm pumps had already been very much improved, compared to reciprocating piston pumps, in regard to leakage rate and dead volume. The ultimate vacuum was comparatively good. The major restriction was mainly the wear of the diaphragms caused by the rocking of the connecting rod. In addition, the stroke was comparatively small, and the pumping speed and ultimate vacuum per

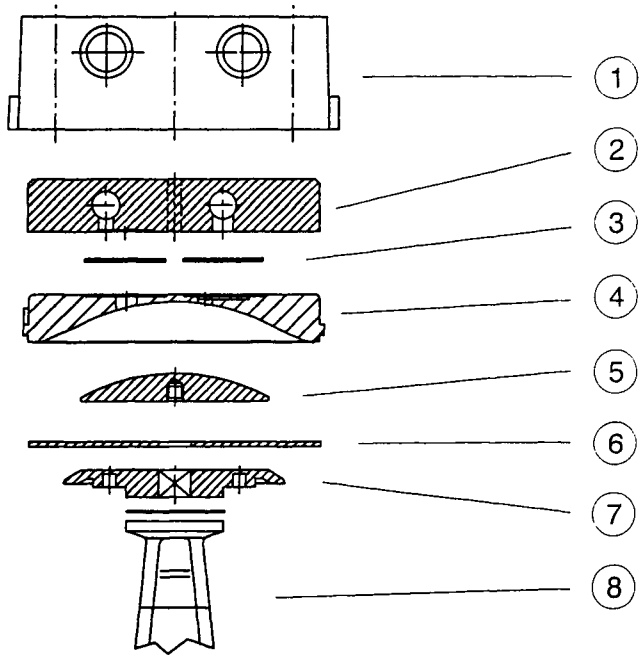
"cylinder" were thus limited. When replacing the rocking by a linear movement, the flexing work of the diaphragm is reduced, but the diaphragm is deformed elastically, which, for example, is not desirable with PTFE (Polytetrafluoroethylene, e.g. Teflon® of DuPont) diaphragms.

Current designs are mainly characterized by the use of comparatively long connecting rods, prolonging typical lifetimes of diaphragms from the range of several weeks to more than 5000 hours. This became possible due to improvements in modern elastomer materials, e.g., laminated, textile-reinforced materials. CAD design and CNC manufacturing allow smooth radia and transitions in radia at the head cover and diaphragm clamping discs, which are essential for a long diaphragm lifetime.

Modern intrinsically corrosion-resistant diaphragm pumps use PTFE-compound materials for parts coming into contact with vapors and gases and are therefore ideal vacuum pumps for the chemical laboratory.

Figure 5 shows the design of a pump head in a chemical version: the head cover itself is made of a ceramic-reinforced PTFE with high mechanical stability and

Fig. 5.



Exploded view of a pump head of the chemical pump version: (1) housing cover, (2) housing cover insert (reinforced carbon, PTFE strengthened), (3) valves, (4) head cover (ceramic-reinforced PTFE), (5) diaphragm clamping disc (special steel with E/CTFE (e.g., Halar® of Allied Chemicals) coating, (6) PTFE Perbunan sandwich diaphragm, (7) diaphragm support disc, (8) connecting rod.

high resistance to creeping. The housing cover insert itself is made of a carbon-reinforced PTFE with high thermal conductivity to remove the heat generated by compression.

The complete cylinder head is surrounded by an aluminum housing to contain the mechanical forces due to clamping of the parts. This construction ensures long-term stability against creeping of the PTFE material, the necessary small clearances, and proper clamping of diaphragms and valves for long-term use. In this design of a chemical diaphragm pump, the diaphragm itself is a PTFE/Perbunan sandwich flat diaphragm. The PTFE foil is an extruded foil of high density having the advantage of small gas permeation. This diaphragm is clamped by a polyfluoroelastomer-coated clamping disc. The setup reduces the deformation of the diaphragm significantly, compared to diaphragms with an embedded clamping disc. The valves are made of a polyfluoroelastomer (such as Kalrez® of DuPont) combining high elasticity compared to PTFE and low leakage rates with high resistance against chemical attack.

2.2.3 PERFORMANCE AND TECHNICAL DATA

The requirements resulting from applications for a vacuum system are usually defined by the type and the amount of gas to be pumped and the vacuum range or operating pressure. The corresponding pump features are the technical parameters of the individual pumps such as pumping speed, ultimate vacuum, resistance against chemical attack, and the capability of handling vapors.

The pumping speed of a one-stage diaphragm pump is defined by the maximum volume of the pumping chamber, the dead volume, and the number of revolutions per unit time. With every stroke of the connecting rod, the amount of gas G_1 is compressed to the pressure p_0 at the outlet of the stage (atmospheric pressure).

$$G_1 = V_p p_i$$

V_p denotes the maximum volume of the pumping chamber of the diaphragm pump and p_i the pressure in the volume to be evacuated.

This amount of gas does not correspond to the quantity of gas actually pumped because of the remaining dead volume V_D resulting in the quantity of gas G_2 expanding in the subsequent expansion stroke of the rod.

$$G_2 = V_D p_0$$

The actual flow of gas q_{pv} is therefore defined by

$$q_{pv} = (G_1 - G_2) = n(V_p p_i - V_D p_0)$$

where n = number of revolutions per time.

This gas flow defines the effective pumping speed of a one stage pump according to

$$S = q_p V / p_i$$

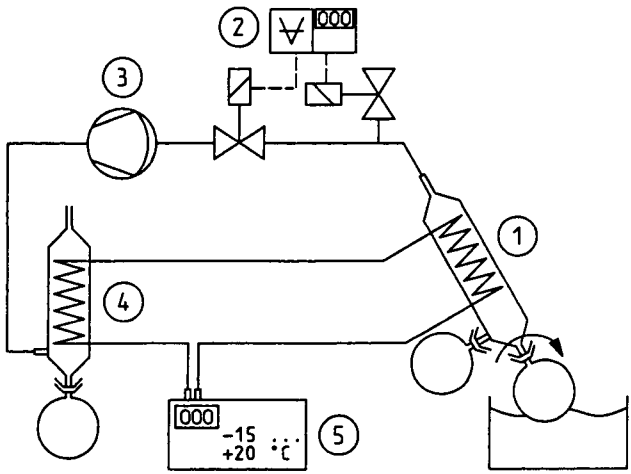
Backflow due to a time delayed opening and closing of the valves or due to internal leakages as well as external leakage is neglected. The pumping speed of a multistage pump is to be calculated by solving a system of equations to be formed by arranging the corresponding equations for each pump stage.

2.2.4

MODULAR CONCEPT FOR SPECIFIC APPLICATION SETUPS: STANDALONE OPERATION

Chemical versions of diaphragm pumps have outstanding advantages for vacuum generation in the chemical laboratory. Compared, for example, to the commonly used water aspirators, neither water is consumed nor waste water is produced. Due to the complete freedom from oil, no reactions of the pumped gases or vapors with the lubricant or sealing oil can occur and therefore, no additional accessories for protection of the pump such as cold traps are required. In addition, the

Fig. 6.



Schematic view of a chemical pumping unit consisting of (1) a rotary evaporator, (2) a vacuum controller, (3) a chemical diaphragm pump, (4) an exhaust waste vapor condensor, and (5) a circulating cooling unit.

diaphragm pump is nearly insensitive to condensation inside the pump. Protection of the environment is not only a question of not producing waste, but of avoiding the emission of volatile and harmful chemicals. By using a pumping unit consisting of a chemical diaphragm pump with vacuum controller and waste vapor condensor (as shown schematically in Figure 6) nearly 100% recovery of all solvents commonly used in the chemical laboratory can be achieved. According to the specific requirement the modularity of diaphragm pumps and their accessories allow the buildup of specific systems for any application.

2.2.5

DIAPHRAGM PUMPS AS BACKING AND AUXILIARY PUMPS IN VACUUM SYSTEMS

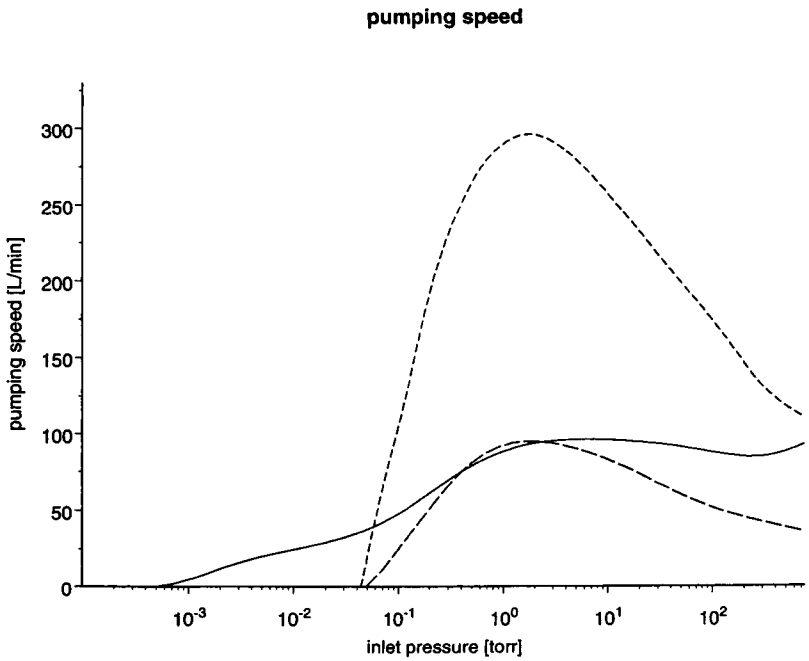
Diaphragm pumps are totally oil free and can be made of materials highly resistant against chemical attack. Their performance is subject to restriction in those cases in which high pumping speed or an ultimate vacuum below 0.1 torr is required. By connecting diaphragm pumps in series to booster pumps the positive features of both kinds of pumps can be combined to get vacuum systems with improved chemical resistance, low ultimate pressure, high pumping speed, and greatly reduced hydrocarbon background. Conventional dry vacuum pumps covering the fine and the rough vacuum range in many cases need additional expensive accessories to overcome specific problems that occur: compressing to atmospheric pressure and achieving ultimate vacuum below 0.1 torr with one pumping principle.

Diaphragm pumps produce satisfactory performance in the rough vacuum range in both corrosive and in clean applications. Depending on the actual requirements the diaphragm pump may be combined with Roots pumps as fine vacuum pumps, with turbomolecular pumps with additional wide pressure range or cryopumps to achieve high and ultra-high vacuum.

The performance of oil sealed rotary vane pumps is limited when pumping condensable or corrosive vapors. These inherent restrictions can be overcome by a rotary vane pump—diaphragm pump combination: The diaphragm pump unit compresses to atmospheric pressure, whereas the oil-sealed rotary vane unit compresses only to the inlet pressure of the diaphragm pump, typically 12 torr. Pumping speed and ultimate vacuum of this combination are determined by the performance of the rotary vane pump (Figure 7). The vacuum requirements for the diaphragm pump result from elaborated dewpoint considerations taking into account the vapor pressures of typical solvents to be pumped.

In almost all cases, a maximum inlet pressure at which no condensation occurs

Fig. 7.



Pumping speed curves for (1) a 100-L/min rotary vane diaphragm pump in series, (2) a 170 L/min Roots pump backed by a 33-L/min diaphragm pump, and (3) a 500-L/min Roots pump backed by a 100-L/min diaphragm pump, with an ultimate vacuum of 0.5 torr.

inside the oil sealed pump can be specified for typical chemicals (see Table 1). In addition, the permanent cleaning of the oil by an inherent vacuum distillation significantly reduces oil change intervals and consequently maintenance costs. Reduction of the oxygen partial pressure in the oil reservoir reduces corrosion.

Similar arguments hold for the combination of Roots pumps and diaphragm pumps, connecting the exhaust of the roots pump to the inlet of the diaphragm pump. The benefit of no oil in the swept volume allows the application of these types of pumping units not only in the chemical field, but also in applications where cleanliness of the residual vacuum is required. No gas cooling or additional resources are required for the Roots pump and all the advantages of the diaphragm pump such as high pumping speeds and oilfree vacuum are retained. Four stage diaphragm pumps in the 0.5 torr range allow the achievement of complete pumping unit ultimate vacua near 5.10^{-2} torr with resulting pumping speed curves as shown in Figure 7. The pumping unit can be used as a comparatively low priced backing unit for conventional turbopumps. A great reduction of hydrocarbons in

Table I

Vapor Pressure of Typical Solvents at the Lowest Temperature (45°C at the Housing) of the Oil-Sealed Part of the Rotary Vane/Diaphragm HYBRID Pump and the Corresponding Highest Inlet Pressure at Which the Dewpoint Is Not Reached

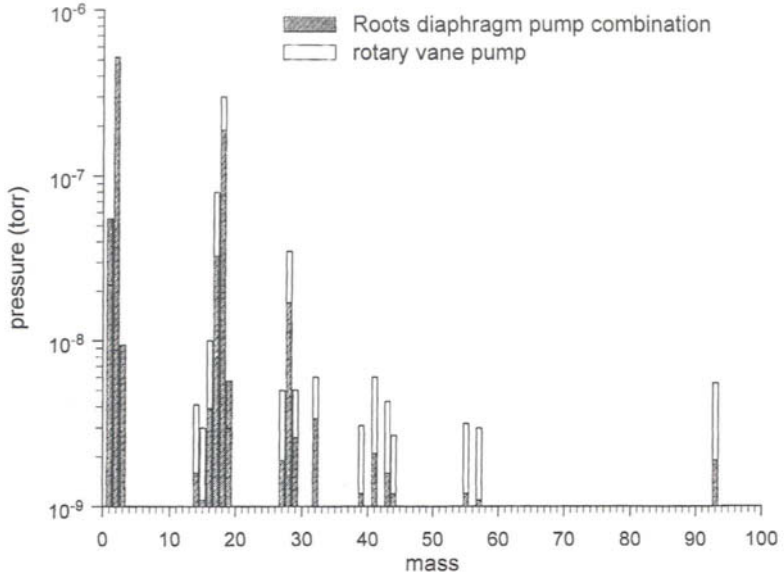
Medium	Vapor Pressure at 45°C [mbar]	Inlet Pressure [mbar]
Diethylether	>1000	380
Methylene chloride	1000	380
Acetone	600	153
Chloroform	570	144
n-Hexane	435	103
Methyl alcohol	415	96
Carbon tetrachloride	300	59
Cyclohexane	283	54
Ethyl acetate	268	50
Benzol	265	49
Acetonitril	230	39
Isopropanol	163	23
Dioxan	138	17
Toluene	100	10
n-Propanol	81	7
Acetic acid	60	4
Iso-butanol	51	3
n-butanol	33	1,9
Acetic anhydride	21	0,7
DMF	18	Condensation
Glycol	< 1	Condensation
Water	95	9

the residual gas of a conventional turbomolecular pump has been observed for the combination with a roots-diaphragm unit compared to a combination of the turbomolecular pump with an oil sealed rotary vane pump. The turbomolecular pump has been equipped with a permanent magnetic upper bearing and a lower pad lubricated ball bearing. Mass spectra (see Figure 8) have been recorded using a quadrupole mass analyzer.

Similar reductions can be attained using wide range hybrid or compound turbomolecular pumps with magnetic bearings. These combinations allow totally oil free vacua.

Investigations performed on the subject of backing a cryopump directly by a diaphragm pump revealed no problems in the operation of the cryopump. [4] The diaphragm pump can be used to preevacuate systems for cryopumps as well as for fast regeneration.

Fig. 8.



Residual gas analysis of a turbomolecular pump with a Roots diaphragm pump combination and a rotary vane pump as a backing pump. Backgrounds have been subtracted in both cases.

REFERENCES

1. PNEUROP Vacuum Pumps, Acceptance Specifications, No. 6602 Maschinenbau-Verlag, Frankfurt.
2. W. Jitschin et al. *Vacuum in Forschung und Praxis*, **3**, 183–193 (1995).
3. F. J. Eckle et al. *Vacuum*, **47**, 6–8, 799–801 (1996).
4. U. Timm et al. *Vacuum*, **46**, 8–10, 875–881 (1995).

CHAPTER 2.3

Vacuum Blowers

Jay Richman
Consultant, Stokes Vacuum Inc.

2.3.1 INTRODUCTION

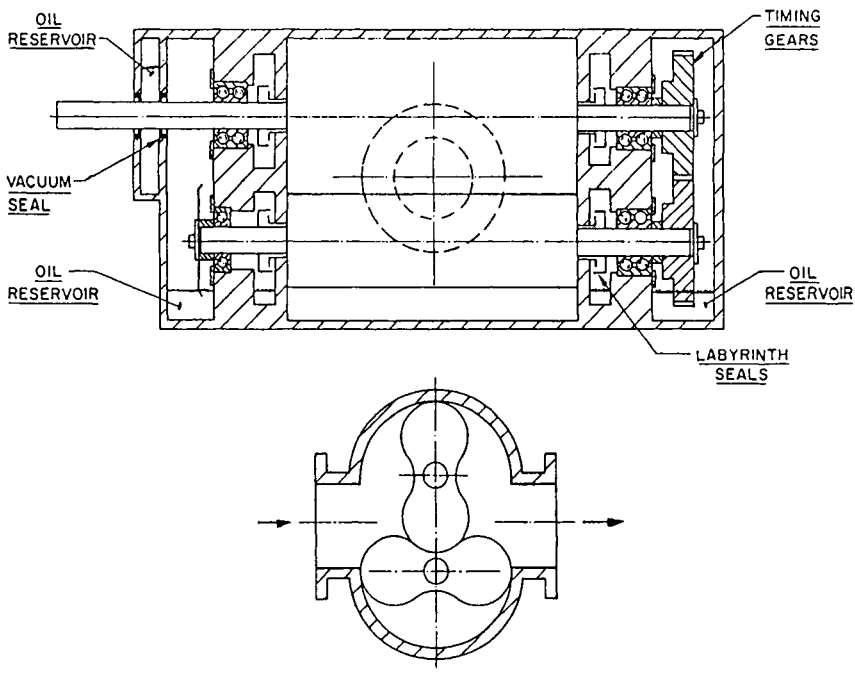
Roots-type blower systems comprised of blowers and backing pumps are almost universally used in vacuum systems where large gas loads are pumped in the intermediate vacuum range and/or volumes must be evacuated in a relatively short time. They are used as primary pumps at intermediate vacuum levels, for rough pumping to the vacuum level where high-vacuum pumps can be effectively employed, and, where required, as backing for high-vacuum pumps. In some systems, the blower system provides all three functions. This section is intended to provide the user with an understanding of the pumping principles, operating characteristics, and limitations of blower systems to assist in proper application and problem diagnostics.

2.3.2 EQUIPMENT DESCRIPTION

Roots-type blowers (mechanical boosters) are positive displacement machines that employ two identical counterrotating impellers, configured to run with close clearances between the impellers and between the impellers and the enclosing casing. The impellers may have two or three lobes with cycloidal, involute, or

ISBN 0-12-325065-7
$25.00

Fig. 1.



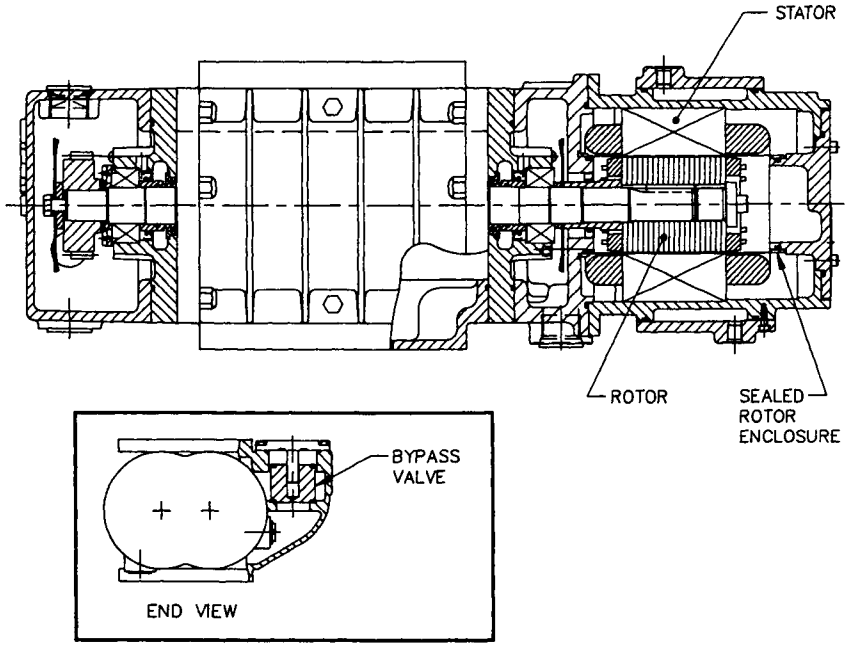
Vacuum seals may be lip or face type. Labyrinth and shaft seals can be replaced by four face seals to isolate the pumping chamber from the bearings and gears.

compound curve flank geometry. Figure 1 illustrates the typical construction of a two-lobe blower. The impellers are mounted on bearings, are well balanced, and normally run at rotational speeds in the range of 1800 to 3600 rpm. Lip or face type seals are used for the atmospheric shaft seals. For critical applications, blowers are available with hermetically sealed direct drive motors to replace the shaft seals. In one version of these motors, the rotor is mounted directly on the drive shaft and isolated from the stator by a hermetic enclosure (Figure 2).

Pumping action is obtained by the impellers trapping pockets of gas at inlet pressure between the impellers and casing, and transferring the gas to a higher-pressure discharge port (Figure 3). The blowers are valveless, run without sealing fluids (except in special cases), and close clearances are required to maintain pumping efficiency by minimizing backflow. Clearance between impellers is maintained by minimum backlash timing gears, and clearances between impeller ends and casing end plates are controlled by accurate positioning of the shafts in the bearings.

Fig. 2.

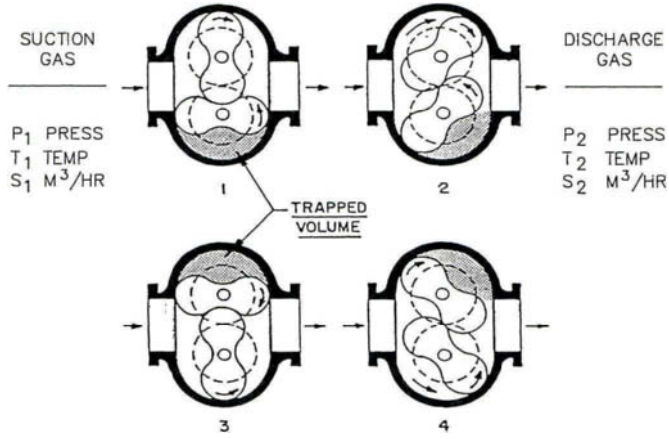
Sealess blower drive.



Rotor isolated from the stator and atmosphere by hermetically sealed, thin-walled shell. Blower has a bypass valve to facilitate starting from atmosphere. Courtesy: Stokes Vacuum Inc.

Fig. 3.

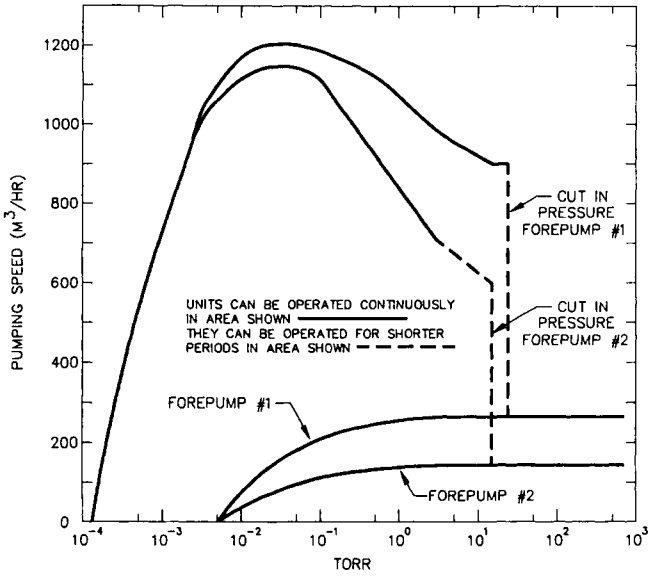
Blower pumping principle.



Two lobe impellers discharge four trapped volumes (per rotation of drive shaft) to higher-pressure region.

Fig. 4.

Typical blower-backing pump performance.



Pumping speed curves for the same blower backed by two different-capacity backing pumps.

Blower systems consist of one or more blowers in series with dry-running or liquid-sealed mechanical vacuum pumps that discharge to atmosphere. The blowers provide high pumping capacities at low pressure and extend the operating range of the backing pumps by a decade or more. Staging ratios (relative displacements of blower and backing pump) range from 2 to 10 or more, depending on operating pressure. The blowers also provide vapor compression that reduces the amount of vapor backstreaming from liquid-sealed backing pumps.

Blank-off pressures of two-stage blower-mechanical pump systems range from 1×10^{-4} to 1×10^{-2} torr depending to a large extent on backing pump blank-off. Capacities range from 170 m^3/hr to 30,000 m^3/hr (and higher). Figure 4 illustrates typical pumping speed relationship to pressure level of two systems that employ the same blower with different capacity backing pumps.

2.3.3

BLOWER OPERATING PRINCIPLE

As noted, pumping action is obtained by the transfer of low-pressure pockets of gas to a higher-pressure regime. At the instant that the pocket is exposed to the blower exit port, high-pressure gas diffuses into the displacement volume to equal-

ize pressure. Recompression is complete when the impeller has rotated 90 degrees and interstage volume between the blower and backing pump has been reduced to its original size. This action takes place four times per revolution of the drive shaft when two lobe impellers are employed. Volumetric displacement (maximum theoretical pumping speed), m^3/hr, is obtained by multiplying the four volumes displaced by the rotational speed per hour (rph).

2.3.4 BLOWER PUMPING EFFICIENCY

As noted, the blower has no valves and depends on close clearances to minimize backflow from discharge to inlet pressure. The amount of gas that flows back through the clearances reduces the net effective pumping speed. Efficiency is measured by the ratio of the actual volume of gas pumped to the gross displacement or theoretical maximum pumping speed.

$$E = \frac{S_E}{S_D} \tag{1}$$

$$E = 1 - \frac{P_2 S_{FB}}{P_1 S_D} \tag{2}$$

where E = blower pumping efficiency
S_D = blower displacement, m^3/hr
S_{FB} = total backflow, m^3/hr
S_E = net pumping speed, m^3/hr
P_2 = interstage pressure, torr
P_1 = blower suction pressure, torr

Total backflow, S_{FB}, (m^3/hr) is comprised of flow through clearances, carryback of gas on impeller surfaces, and reexpansion of minor volumes trapped between impellers.
Efficiency is influenced by the following factors.

2.3.4.1 Pressure Level and Compression Ratio

Impeller clearances vary with blower size but typically are on the order of 0.25 mm. At high pressures, with a compression ratio of 2 (or higher), backflow is sonic and turbulent and is a function of upstream pressure. As suction pressure decreases, flow may become viscous and backflow will be a product of the differential pressure times clearance conductance. With a further reduction in pressure,

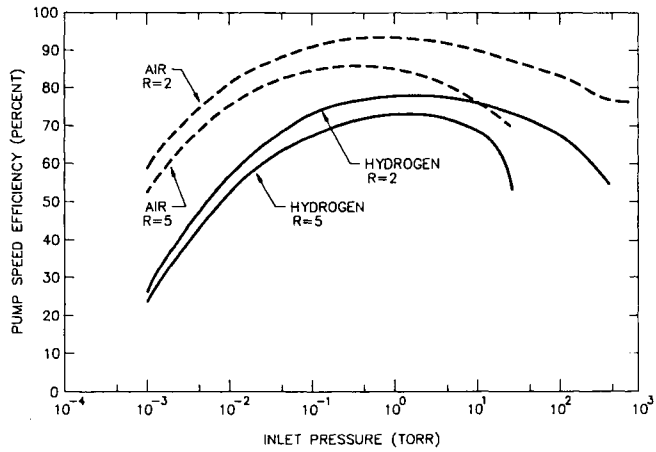
flow becomes transitional and finally molecular; flow resistance increases, conductance backflow becomes nearly constant and compression ratio increases. The results of this phenomenon can be observed in Figure 4, where, at pressures below 10 millitorr the pumping speed of the blower is essentially the same despite the difference in pumping speeds of the backing pumps.

2.3.4.2 Gas Species

Conductance through clearances varies not only with the pressure regime but also with the type of gas being pumped. In the turbulent and viscuous flow ranges, flow is inversely proportional to gas viscosity. As an example, the viscosity of air = 180 and hydrogen = 87 micropoise, and conductance and backflow for hydrogen would be twice that for air. In the molecular flow range, conductance is proportional to the square root of the ratio of the molecular weights, and conductance backflow for hydrogen would be 3.79 times that for air. The effect of the increased backflow (and compression ratio) on pumping efficiency when pumping air and hydrogen at two different compression ratios is illustrated in Figure 5.

Fig. 5.

Blower pumping speed efficiency curves.



Effect of staging ratio (R) and gas species (air and hydrogen) on pumping speed efficiency.

2.3.4.3 Rotational Speed

Theoretical pumping speed increases with rpm, but clearances and conductance backflow remain constant for a given compression ratio. As RPM increases, efficiency increases because backflow becomes smaller relative to displacement.

2.3.4.4 Gas Carryback

Gases adsorbed on impeller surfaces and trapped in minor impeller volumes are carried from the high-pressure interstage and released at low inlet pressure. The effect of carryback at pressures above 100 mtorr are minimal. Polishing the impeller surfaces can significantly reduce carryback.

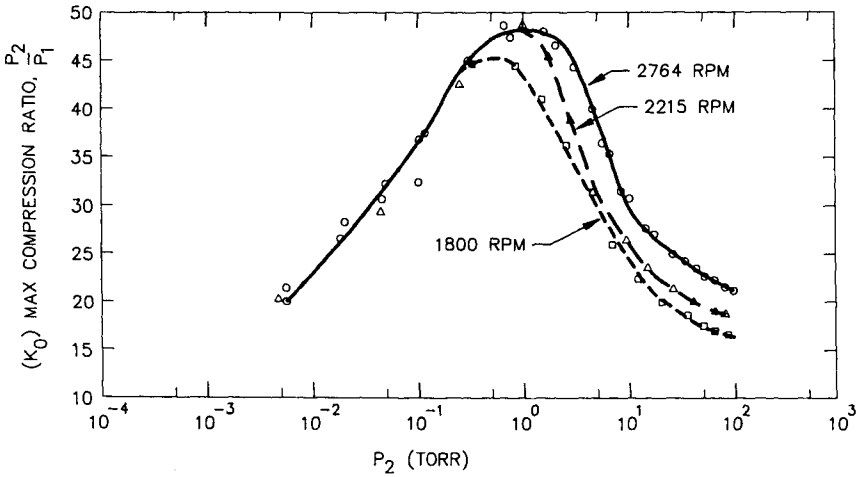
2.3.5 BLOWER PUMPING SPEED CALCULATIONS

The net effective pumping speed of the first-stage blower in a blower system is a function of blower efficiency and the efficiency of the backing pumps. In standard two-stage systems, the pumping speed of the system is usually published by the manufacturer. For nonstandard, two-stage, and multistage systems, the published pumping speed of the backing pump is the starting point for calculations. The pumping speed of the blower adjacent to the backing pump is calculated, and in multistage systems this progresses until the speed of the first-stage blower is determined.

The information usually supplied by the blower manufacturer for purposes of pumping speed calculations is a zero-flow maximum compression ratio (K_0) curve. Figure 6 illustrates typical compression ratio curves at three rotational

Fig. 6.

Typical K_0 zero flow compression ratios.



Data taken with three blower rotational speeds. Backflow (slip) decreases as a proportion of displacement with increased rotational speed.

speeds. These compression ratio curves are obtained by blanking off the blower inlet (zero flow) and, with the blower and backing pump running, admitting air or other gas into the interstage volume between blower and backing pump. The interstage and resulting blower suction pressures are measured, and a curve is plotted of K_0 (compression ratio P_2/P_1) as a function of the interstage pressure P_2. Following Equation (2) and letting $P_2/P_1 = K$

$$E = 1 - KS_{FB}/S_D \tag{3}$$

When flow at the inlet is zero, backflow is balanced by pumping capacity, net pumping speed is zero, efficiency is zero, and

$$K_0 = S_D/S_{FB} \tag{4}$$

The general equation for blower efficiency is derived from Equations (3) and (4).

$$E = 1 - K/K_0 \tag{5}$$

Relating the throughput of the backing pump to blower pumping speed:

$$S_2P_2 = (ES_D)P_1 \quad \text{(ignoring temperature effects)}$$

$$K = (ES_D)/S_2 \quad \text{and} \quad K = EK_D \tag{6}$$

$$(S_D/S_2 = K_D = \text{theoretical compression ratio})$$

Combining Equations (5) and (6),

$$E = \frac{K_0}{K_0 + K_D} \tag{7}$$

Sample pumping speed calculation:

S_D = blower displacement = 600 m^3/hr
S_2 = backing pump pumping speed at P_2 = 200 m^3/hr
P_2 = 1 torr
K_0 = 44 From Figure 6 at P_2 and 1800 rpm, for air
$K_D = S_D/S_2 = 600/200 = 3$ | $S_1 = ES_D = 0.94 * 600 = 563$ m^3/hr

$$E = \frac{K_0}{K_0 + K_D} = \frac{44}{44 + 3} = 0.94 \quad | \quad P_1 = \frac{S_2P_2}{S_1} = \frac{200 * 1}{563} = 3.6 * 10^{-1} \text{ torr}$$

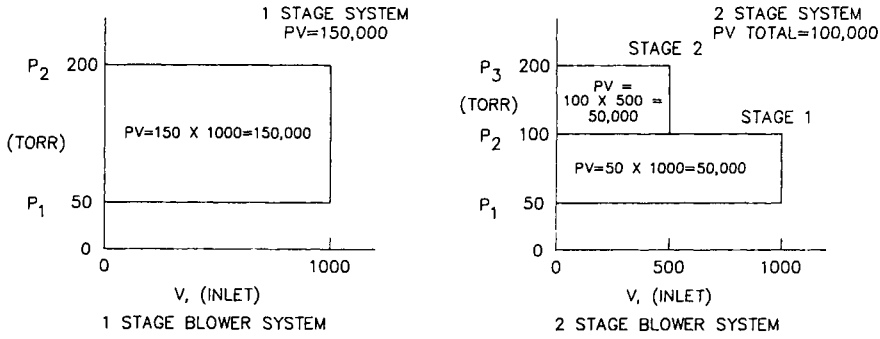
2.3.6

POWER REQUIREMENTS

Power requirements in blower pumping systems are somewhat unusual. Many systems run at low operating pressure for most of their operating cycle. Time at

Fig. 7.

Power comparison for single- and two-stage blower systems.



Staging ratio of 2 assumed for this illustration.

high pressure may be relatively short, and the large motor required to handle the high gas loads may be underutilized. Also, large motors running at low loads can cause power factor problems. To reduce initial and operating costs, systems have been designed with motors that have been much overloaded at peak loads. These have performed satisfactorily, but to prevent equipment damage safe time at overload must be carefully predetermined. High-temperature windings and thermal overload protection must be used.

Power savings can be obtained by the use of two blowers in series or a blower-backing pump combination to obtain a given pressure rise. Contributing to this phenomenon is the pressure–volume relationship of Roots-type blowers.

The rotary positive blower follows a "square card" diagram. That is, if pressure rise is plotted against the constant displacement volume, the resulting curve throughout one complete cycle would be a rectangle under ideal conditions. In Figure 7, with a single blower, and a pressure rise of 200 torr, $PV = 150{,}000$. If two stages are used with an assumed compression ratio of 2, PV would equal 100,000. Power consumption is a direct function of PV, and the savings would be significant. The cost of using two blowers in series may outweigh the power savings, but such systems have been built for factors other than economy and power savings have been achieved.

The power required by a blower is the sum of the power required for gas compression plus the power required to overcome the friction of bearings, seals, gears, etc. The values for friction losses must be obtained from the blower manufacturer. Power for compression is calculated as follows:

$$KW = 3.7 * 10^{-5} * S_D * P_1[(P_2/P_1) - 1)]$$

$$P_2/P_1 = r, \qquad P_1[r - 1] = P_2 - P_1 \qquad \text{and} \tag{8}$$

$$KW = 3.7 * 10^{-5} * S_D * (P_2 - P_1)$$

Note that power for compression is dependent on blower displacement and delta *P* but independent of gas species. In the absence of specific information, approximately 20% may be added to the calculated power for the power to overcome friction but it would be best to contact the manufacturer for this value. In addition to these power requirements, each blower has lower and (as might be expected) upper power limits. The lower limit is established by the power required to overcome friction and impeller inertia when accelerating the impellers from a standing start to operating speed with no gas load. The upper limit is dictated by the mechanical strength of the operating components: impellers, gears, shafts, bearings, etc. These data must also be obtained from the manufacturer.

2.3.7 TEMPERATURE CONSIDERATIONS

2.3.7.1 Maximum Allowable Temperature Rise

A maximum allowable continuous gas temperature rise is established by the blower manufacturers to prevent seizure from excessive differential expansion between the impellers and casing. The maximum rise allowed varies with blower size and manufacturer but a fair average may be 100°C, based on a maximum inlet temperature of 37°C. The differential expansion occurs because both the casing and impellers absorb significant amounts of the heat of compression, but the impeller is partially insulated by vacuum and expands at a greater rate than the casing. Gas temperature rise may be calculated as follows:

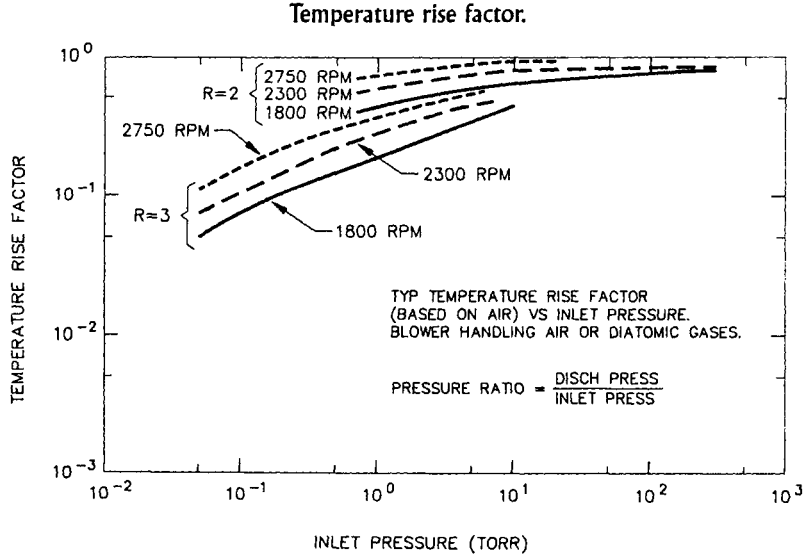
$$T_2 - T_1 = \frac{(k-1) * T_1}{k} * \frac{(r-1) * F}{E} \tag{9}$$

T_2 = degrees Kelvin discharge temperature
T_1 = degrees Kelvin inlet temperature
$r = P_2/P_1$
$k = c_p/c_v = 1.4$ air
F = temperature rise factor

The temperature rise factor, F, is obtained empirically and accounts for heat loss by radiation and convection. Figure 8 illustrates typical (varies with the blower) factors related to inlet pressure, compression ratio, and rotational speed. Note that as compression ratio decreases and rotational speed increases, the factor increases. These trends increase efficiency, which tends to decrease the theoretical temperature rise. However, the increased efficiency also results in less thermal losses, hence an increase in the factor.

When inlet temperatures are above 37°C, the maximum allowable temperature

Fig. 8.



Empirical data used to modify the theoretical temperature rise.

rise, T_a, must be derated as follows:

$$T_a \text{ corrected} = T_a - 0.66(T_{in} - 27°C) \qquad (10)$$

$$T_{in} = \text{degrees C inlet temperature}$$

Because thermal expansion is time dependent, in pumpdown situations it may be possible to exceed the maximum allowable temperature rise provided that the time from blower start pressure to the safe operating pressure does not exceed a certain limit. The blower manufacturer should be consulted for this limit. As a word of caution, if time between pumpdowns is very short and the cycle is repetitive, it would be best not to exceed the continuous allowable limit.

2.3.7.2 Maximum Allowable Discharge Temperature

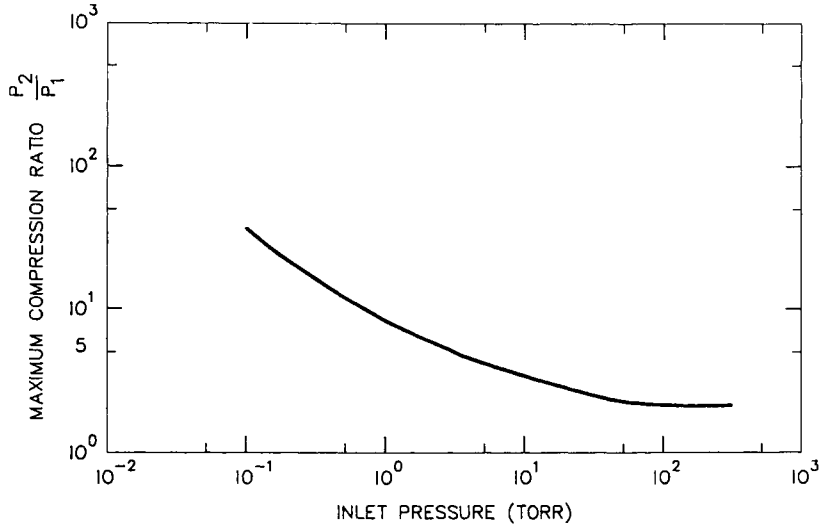
The maximum allowable discharge temperature is based on the limits established by materials of construction such as seals, lubricants, etc. The manufacturer's recommendations should be followed.

2.3.7.3 Maximum Allowable Compression Ratio

The curve in Figure 9 depicts a typical range of allowable pressure ratios versus inlet pressure based on the allowable temperature rise and average temperature

Fig. 9.

Maximum allowable compression ratio.



Typical values based on "average" temperature rise factor.

rise factors. It can be used for a first approximation to establish blower-backing pump relationships subject to a more accurate analysis.

2.3.7.4 Discharge Gas Cooling

Coolers that employ water-cooled surfaces in close proximity to the impellers at the discharge port may be used to cool the interstage gas that surges into the pumped gas pockets. The resulting cooling effect allows a higher compression ratio.

2.3.8 FLOW AND COMPRESSION RATIO CONTROL MECHANISMS

Several arrangements are available to keep power requirements low and the heat of compression within allowable limits.

2.3.8.1 Sequenced Start Systems

In sequenced start systems, the backing pumps are started first, and blowers are brought on line (starting with the first upstream blower) at safe operating pres-

sures by pressure switches or timers. In some variations, with several stages, two or more blowers may be switched into parallel at low pressure to provide more pumping speed.

2.3.8.2 Inlet Throttling

In inlet throttling systems, the blower and backing pump run at full speed continuously. A throttling valve at blower inlet controls blower inlet pressure to maintain an allowable temperature rise or, if power limited, the differential pressure. In the former case, over-temperature switches are still required. The valve is partially open at high pressure and may be fully open at low pressure.

2.3.8.3 Blower Rotating Speed Control

Blower rotating speed control systems maintain allowable compression ratios by gradually increasing blower rotating speed as the inlet pressure decreases. Following are some typical speed control systems.

1. Variable-speed AC or DC motors and drives.
2. Hydraulic or magnetic slip-type coupling between the blower and drive motor.

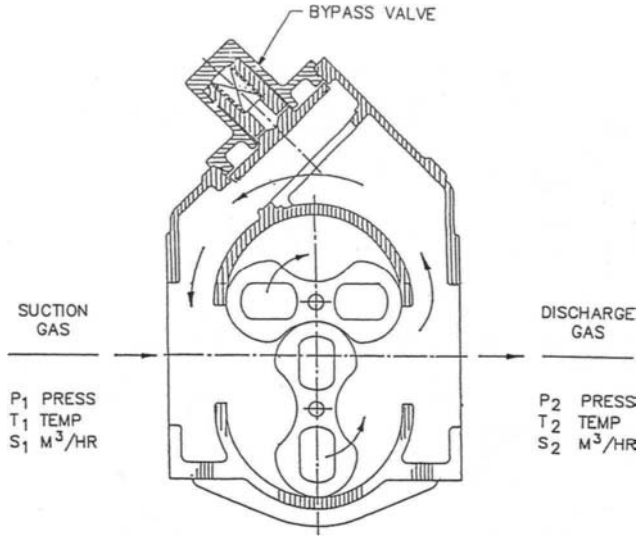
2.3.8.4 Bypass Valves

In a bypass valve type of system (Figure 10), an integral bypass valve with spring or deadweight loaded poppet or external bypass throttling valve, feeds a portion of the blower discharge gas back to the blower inlet. The valve is adjusted to provide a fixed safe differential pressure across the blower. When the throughputs of the blower and backing pump are equal, the valve closes. The valve closes at pressure $P = (P_2 - P_1)/(P_2/P_1) - 1$. Figure 10 illustrates a stock blower with integral bypass. The blower and backing pump run at full speed from atmosphere to ultimate pressure. Blower power required is a function of the pressure difference selected.

Inlet-throttling, blower rotating speed control, and bypass valve systems can produce higher pumping speeds from atmosphere to safe operating pressure than a sequenced start arrangement; see Figure 11. Assuming that the systems are arranged to provide the same safe differential pressure, the gain in pumping speed would be the same for all three systems.

Fig. 10.

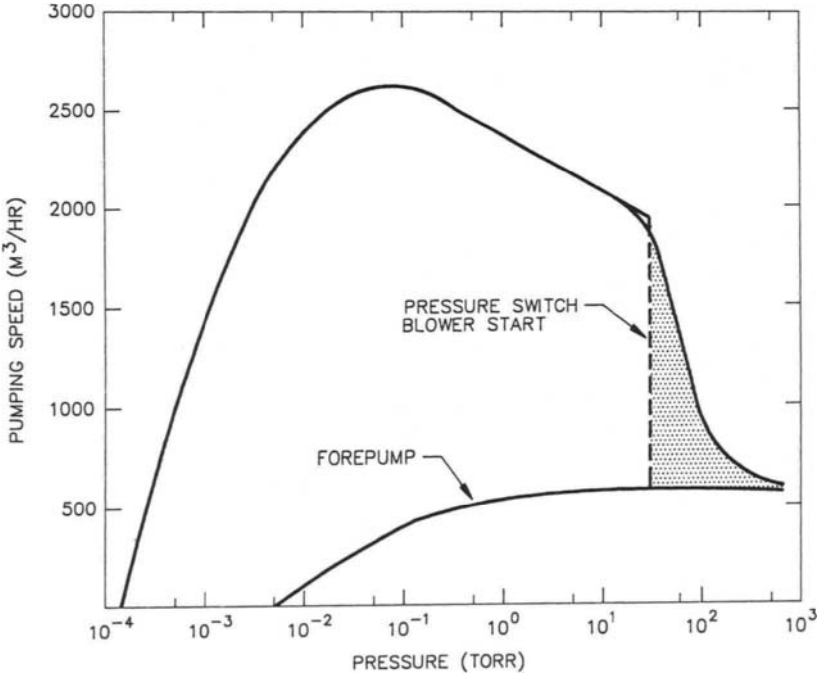
Blower with integral bypass valve.



Safe differential pressure established by spring selection.

Fig. 11.

Typical blower-backing pump speed performance with compression ratio control.



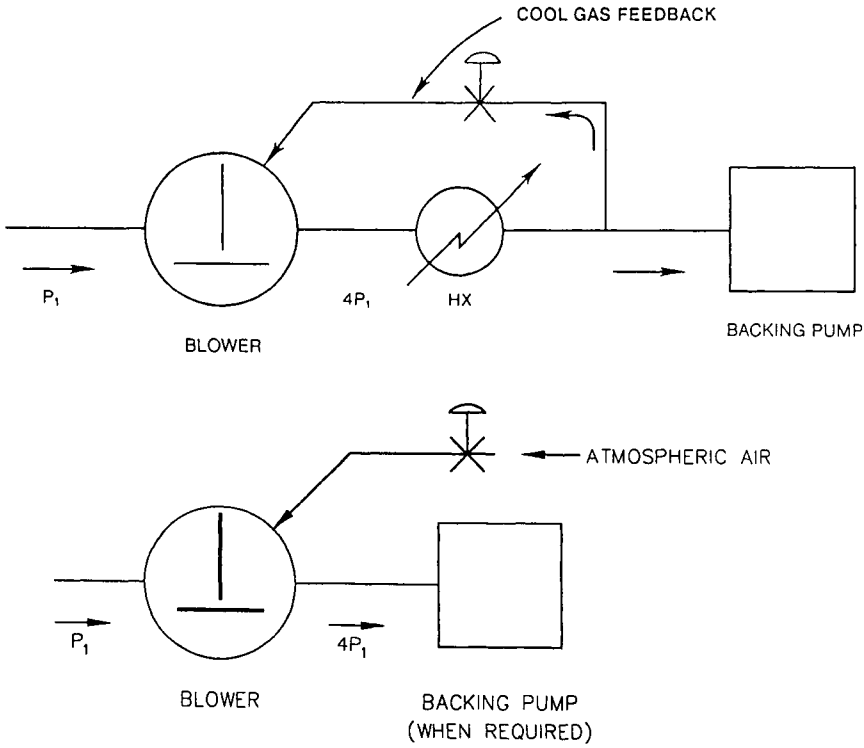
Both pumps run from atmosphere. Shaded area indicates increase of pumping speed (using ratio control).

2.3.8.5 Cool Gas Feedback Blowers

In cool gas feedback blowers, the blower is arranged to permit the admission of cool gas into the trapped volume of the blower where it mixes with and reduces the temperature of the discharge gas. These blowers are generally used at pressures of 200 torr and higher but they have also been used at lower pressures. The advantage of these blowers is that at higher pressues a compression ratio of 4:1 may be maintained, compared to 2.2:1 for uncooled blowers. This feature may make it possible to eliminate one blower stage when used alone to discharge to atmosphere or as backing pumps in chain blower systems. The penalty paid is high power consumption, as discussed in Section 2.3.6.

Atmospheric air may be used for feedback or if special gases are being pumped that must be preserved, a portion of the discharge gas may be fed back through a cooler (see Figure 12).

Fig. 12.



Cool gas feedback systems: (a) Pumped gas recirculation system, (b) Atmospheric air cooled.

2.3.9 LIQUID-SEALED BLOWERS

Blowers can be arranged to use water or other fluids (such as oil) injected into the inlet port to act as a cooling and sealing medium. With water injection, operating pressure with a single-stage blower is approximately 100 torr and with a two-stage blower, 50 torr. Pressure limit is generally based on the water vapor pressure at a given temperature. Lower pressures are attainable with lower-vapor-pressure fluids. These blowers compete with liquid ring pumps. They generally use less water and power than liquid ring pumps, but their initial cost may be higher. As with liquid ring pumps, consideration must be given to the cost of disposal of contaminated liquids and the cost, when required, of a closed loop system to cool and recover expensive liquids or to conserve water.

2.3.10 SELECTED SYSTEM ARRANGEMENTS

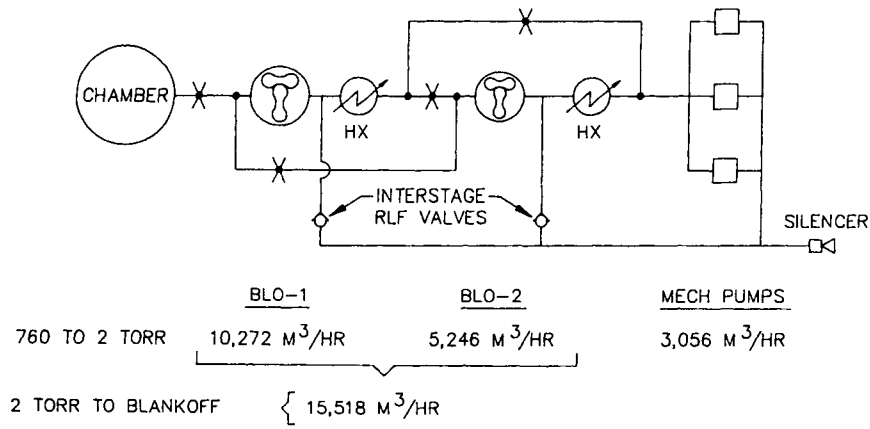
The broad range of blower and backing pump types available provides the system designer with an opportunity to optimize systems for application requirements. Following are some interesting examples of systems developed for specific applications.

2.3.10.1 Multistage (chain blower) Systems

Multistage systems are generally employed when large volumes must be pumped down in a relatively short time. The multiple stages are used to maintain the allowable compression ratio across each stage when starting from atmosphere. An example is a 3-stage chain blower system (Figure 13) built to evacuate a 1308-m^3 chamber from 760 torr to 0.01 torr in two hours with a 20 torr L/S air leak. Beginning with the last stage, each stage is sequentially started with a few seconds time delay between stages to reduce initial power surge. With an approximately 2:1 staging ratio between the three stages, the interstage relief valves maintained a maximum allowable differential pressure of 400 torr. Excess air was discharged through the valves until the throughput of each backing stage equaled that of the preceding stage. The interstage heat exchangers removed the heat of compression, maintained an approximately constant compression ratio, and limited the inlet temperatures to the downstream stages. With this type of system, the first-stage pumping speed of 10,270 m^3/hr was essentially maintained from atmosphere

Fig. 13.

Multistage blower pumping system.



Three-stage, series-parallel operation. Pumpdown of 1,308 m^3 space chamber from 760 to 0.01 torr in 100 minutes with 20 torr L/s air leak. Blowers are parallel at 2 torr to handle air leak at 0.01 torr. Courtesy: Stokes Vacuum Inc.

down to 2 torr. At 2 torr the first and second stages were valved in parallel to provide 15,518 m^3/hr to handle the air leak, which at 0.01 torr equaled 7200 m^3/hr. Actual time to 0.01 torr was 100 minutes.

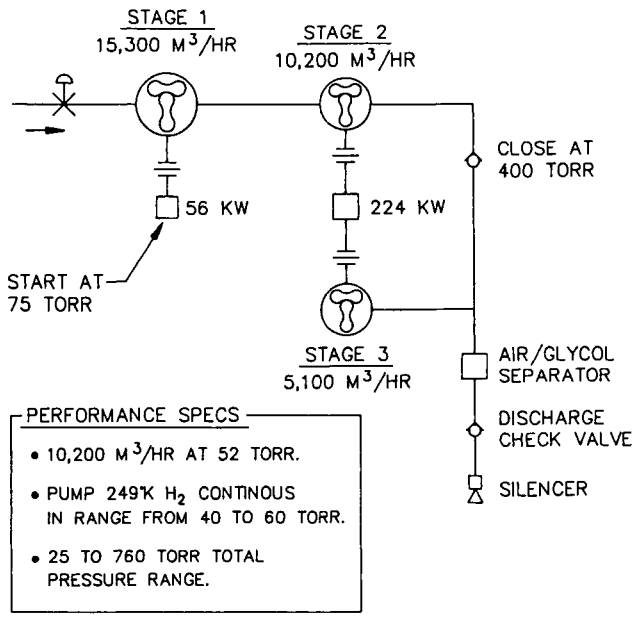
Peak power requirements for stages 1 and 2 were 159 kw and 75 kw respectively. Mean average power over a 20-minute period was approximately 60% of peak. Motors selected had ratings of 100 kw and 56 kw. Time at overload was approximately 3 minutes. High-temperature insulation was used for the windings, and special over-temperature protection was provided by sensors buried in the windings.

2.3.10.2 Slush Hydrogen Pumping System

The slush hydrogen pumping system (Figure 14) was designed to pump 10,200 m^3/hr of pure hydrogen at 52 torr from a slush hydrogen source at 14° Kelvin. Through heat gain, the temperature at the inlet was approximately 249° Kelvin. To handle the cold gas, a three-stage liquid-sealed blower system was selected, using ethylene glycol for the sealing liquid. The glycol was circulated through a closed-loop system, where it was heated to 40°C to maintain the blowers at a reasonable operating temperature. Liquid sealing increased blower efficiency, permitting the use of smaller blowers and direct discharge to atmosphere.

Fig. 14.

Slush hydrogen pumping system.



Heated glycol injected into the suction ports of the blowers for sealing and to maintain blowers at reasonable operating temperature. Courtesy: Stokes Vacuum Inc.

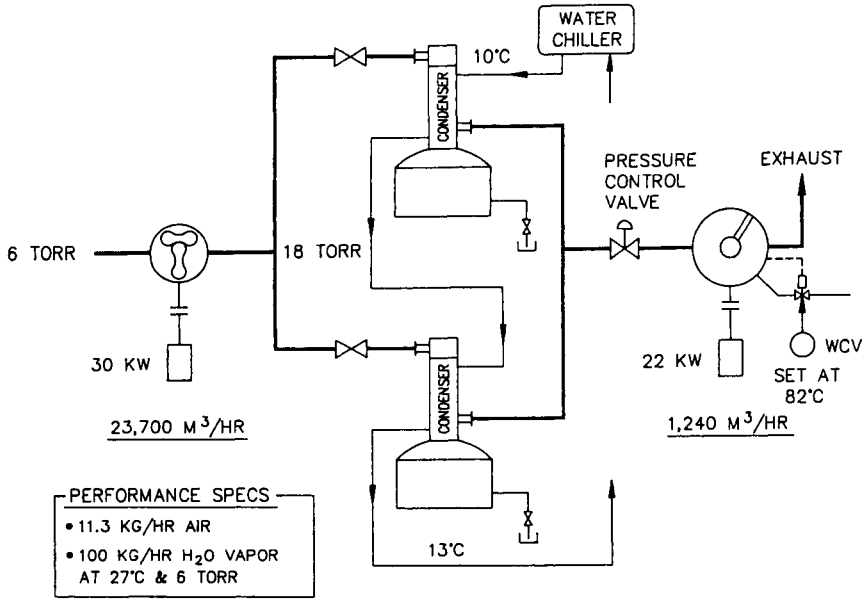
At start of pumpdown, stages 2 and 3 were in operation with a portion of the discharge from stage 2 bypassed around stage 3 through a pressure-sensitive check valve directly to the air–glycol separator and the silencer. At approximately 400 torr, the check valve closed, and the discharge of stage 2 was handled entirely by the third stage. At 75 torr, the first stage was started and hydrogen flow commenced. Pressure decreased to 52 torr as pumping progressed.

2.3.10.3 Product Dehydration System

The product dehydration system (Figure 15) was designed to hold 6 torr with 11.4 kg/hr air plus 100 kg/hr water vapor. A blower-condenser-mechanical pump system was employed to provide the net effective pumping speed of 18,500 m^3/hr. The blower was used to increase the pressure at which the condensers would operate — in this case, from 6 to 18 torr. Two condensers, each with receivers, were required to handle the entire water load, with a chiller furnishing 10°C cooling water for the condensers. An automatically operated pressure control valve lo-

Fig. 15.

Product dehydration system.



High water vapor load (100 kg/hr) plus small air load (11.3 kg/hr) to be pumped at 6 torr. The backing pump handles the air load plus partial pressure of water vapor at 10°C condenser temperature. Courtesy: Stokes Vacuum Inc.

cated between the condensers and the mechanical pump maintained a pressure of 18 torr in the condensers to prevent reflux of the condensed vapor. Most of the water was removed by the condensers, and the backing pump was small because it was only required to handle the air load plus the small water vapor load corresponding to the partial pressure of the water vapor in the condenser. The backing pump was operated at 82°C.

The condensing system was designed to operate for eight hours. With the addition of automatic valving to drain the receivers while operating, the system could have been made to operate on a continuous basis.

CHAPTER 2.4

Vapor Jet Pumps (Diffusion Pumps)

Marsbed Hablanian *(Retired)*
Varian Associates

Many high-vacuum pumping systems include a vapor jet pump and a mechanical pump.

Mechanical vacuum pumps are used to remove about 99.99% of the air from the chamber, reducing the pressure in the vacuum system to the correct operating range. This is commonly termed rough pumping or roughing. After suitable flow and pressure conditions are reached (usually 10^{-1} to 10^{-3} torr), the vapor jet pump can take over. The mechanical pump is now used to maintain proper discharge pressure conditions for the vapor jet pump at the foreline connection. This is called *backing* or *forepumping*.

Vapor jet pumps are normally used when constant high speeds for all gases are desired for long periods of time without attention. They are essentially vapor ejectors specialized for high-vacuum applications. In the past, the gas vapor diffusion aspects and the vapor condensation have been overemphasized, giving the original name for the pump.

The basic design form was stabilized approximately 70 years ago. Modern pumps direct the vapor stream at high velocity in the pumping direction. Pumped gas entrainment into this stream, in principle, is not very different from steam ejectors or other vapor pumps or compressors. The original pumping fluid was mercury. Oil-like substances were first used in 1928.

The following discussion is devoted mainly to oil vapor jet pumps.

ISBN 0-12-325065-7
$25.00

2.4.1

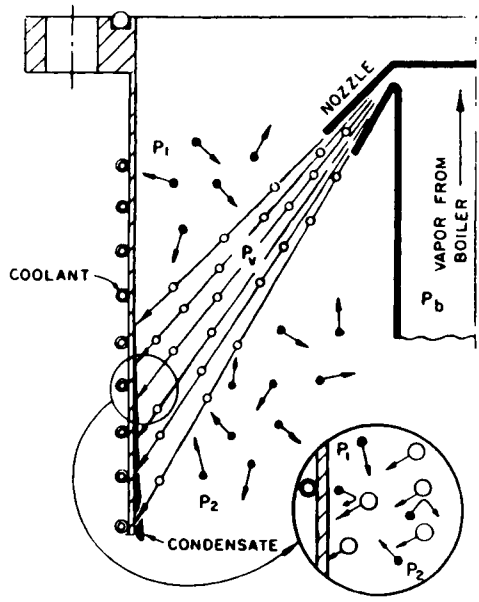
BASIC PUMPING MECHANISM

The typical vapor jet pump consists of a vertical, usually cylindrical, body fitted with a flanged inlet for attachment to the system to be evacuated. The bottom of the cylinder is closed, forming the boiler, which is fitted with a heater. The upper two-thirds of the body is surrounded by cooling coils. An outlet duct (or foreline) is provided at the side of the lower pump body for discharging the pumped gases and vapors to the mechanical forepump. The sectional view in Figure 1 illustrates a schematic arrangement of a single-stage vapor jet pump.

A jet-forming structure is located within the pump body. This consists of a concentric cylinder partially capped and fitted with flared ends to form jets through which the pump fluid vapors can emerge at high velocity. There are no moving mechanical parts.

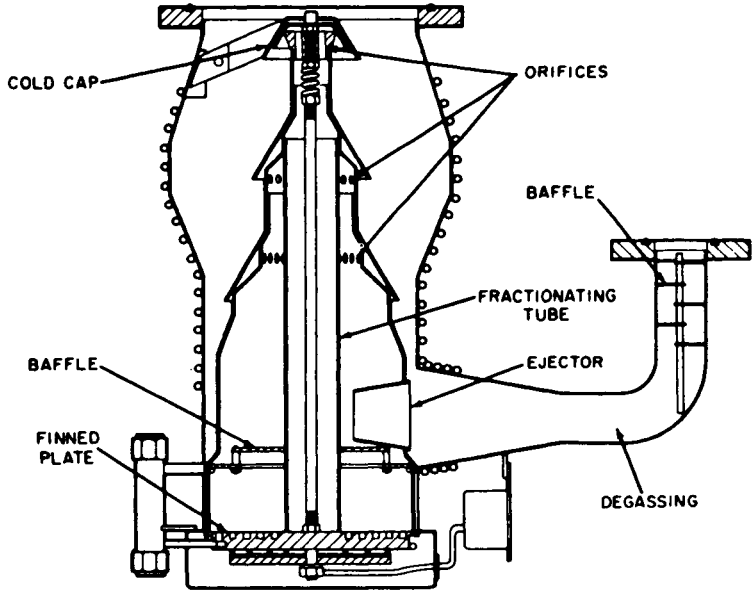
In operation, the working fluid in the boiler of the pump is heated by means of the electric element clamped to the lower body, and a vapor stream is created. This vapor rises in the jet structure and is emitted through the annular nozzle in a downward and outward direction against the water-cooled wall of the pump body.

Fig. 1.



Schematic arrangement of a single-stage vapor jet pump. Open circles = vapor jet molecules; closed circles = gas molecules.

Fig. 2.



Multistage vapor jet pump.

Gas molecules arriving at the pump inlet are entrained in the stream of working fluid vapor and are given a downward momentum. The vapor stream normally flows at supersonic velocities. The gas vapor mixture travels downward toward the foreline. When the oil vapor constituents of such a jet stream strike the water-cooled wall of the pump body, they are condensed and returned to the boiler in liquid form. The entrained gas molecules continue their path toward the exit, where they are removed by the mechanical forepump.

The condensed oil vapors return to the boiler. Heat is once more added and the oil revaporized to maintain the vapor flow to the jet assembly and the continuity of the pumping mechanism. A sectional view of a typical multistage vapor jet pump is shown in Figure 2.

The pumping action in a vapor jet pump results from collisions between vapor and gas molecules. It is more difficult for gas molecules to cross the vapor stream in the counterflow direction. Thus a pressure (or molecular density) difference is created across the vapor stream. The pressure ratio created by the vapor stream can be approximately expressed by

$$P_2/P_1 = \exp(\rho VL/D)$$

where ρ is vapor density, V its velocity, and L the width of the stream. D is related to the diffusion coefficient and depends on molecular weight (M) and diameters

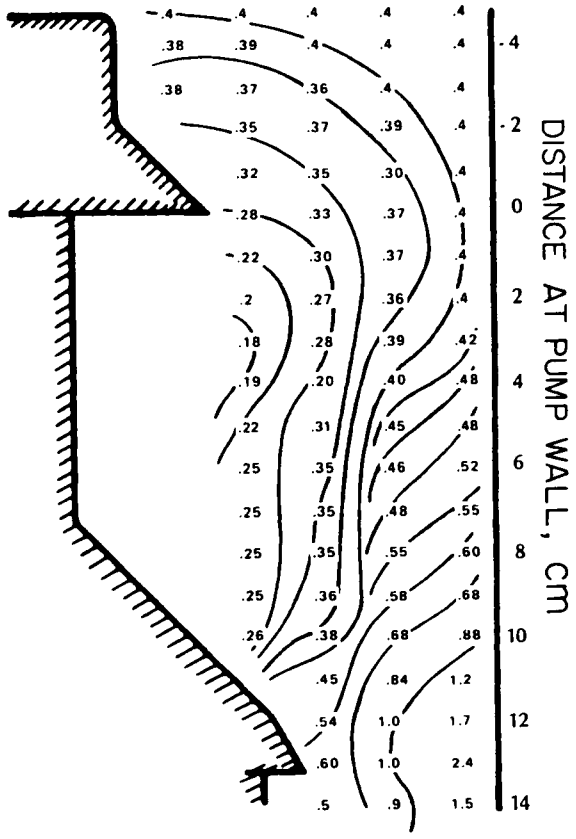
(σ) of the vapor and gas [1].

$$D = \frac{3}{8(2_\pi)^{1/2}}\left(RT\frac{M_1 + M_2}{M_1 M_2}\right)^{0.5}\left(\frac{\sigma_1 + \sigma_2}{2}\right)^{-2}$$

where subscript 1 refers to the pumped gas and 2 to the pumping fluid, R is the gas constant and T is the temperature. From this, you can see that the pressure ratio is much lower for lighter gases.

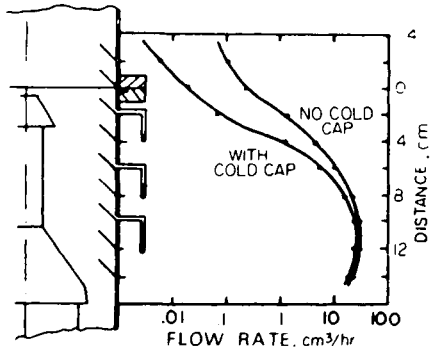
The interaction of pumping vapor and the pumped gas can be illustrated experimentally by finding the density distribution of both species in the pumping region. The gas density distribution obtained by traversing the pumping region with an ionization gauge probe is shown in Figure 3 and the distribution of vapor ar-

Fig. 3.



Gas density distribution obtained by traversing the pumping region with an ionization gauge probe.

Fig. 4.



Distribution of vapor arriving at the pump wall.

riving at the pump wall in Figure 4. The pattern of penetration of pumped air into the vapor jet, its relative absence in the core of the jet near the nozzle exit, and the gradual compression of the gas can be deduced from Figure 3. The repetition of the process in the second stage is also evident.

2.4.1.1 Number of Stages

The number of pumping stages or vapor nozzles depends on particular performance specifications. A single-stage pump would have conflicting requirements of high pumping speed and high compression ratio. Normally, the first stage at the inlet has high pumping speed and low compression ratio, and the last (discharge) stage vice versa. Small pumps often have three stages and large ones five or even six. The initial stages have annular nozzles; the discharge stage sometimes has a circular nozzle and is called an *ejector*. There are no principal differences in such variations of geometry, although certain advantages are gained by one or the other choice. Sometimes, to obtain certain performance effects two vapor jet pumps can be used connected in series. This has an effect of increasing the number of compression stages, and it allows the use of different pumping fluids in the two pumps.

2.4.1.2 Vapor and Gas Interaction

As noted before, in regard to their basic pumping action, vapor jet pumps or high-vacuum vapor pumps are related to oil or steam ejectors. The potential energy of elevated pressure inside the jet assembly (boiler pressure) is converted to the kinetic energy of high-velocity vapor stream or jet after it passes through a nozzle. The gas is pumped by the jet by momentum transfer in the direction of pumping.

Because the pumping fluids used in vapor jet pumps are easily condensable at room temperature, a multistage nozzle-condenser system can be fitted in a compact space.

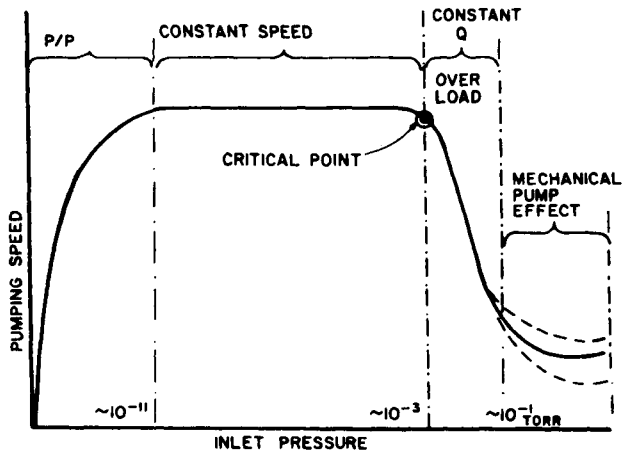
2.4.1.3 Operating Range

The pressure range for the use of vapor jet pumps is between 10^{-10} and 10^{-1} torr. At the high-pressure limit, the steady-state pressure (at the pump outlet) normally should not exceed about 1×10^{-3} torr. Without the assistance of cryogenic pumping and without baking, the lowest inlet pressures conveniently achieved are near 10^{-8} torr. With the aid of cryogenic pumping, liquid-nitrogen-cooled traps, for example, inlet pressures below 1×10^{-10} torr range can be obtained. Generally, baking and subsequent cooling of pumped chambers and inlet ducts are required in order to extend the pressure range, since they can reduce outgassing as well as produce sorption pumping effects.

2.4.1.4 Basic Performance

The pumping performance of a vapor jet pump is usually displayed in the form of a plot of pumping speed versus inlet pressure, as shown in Figure 5. The graph consists of four distinct sections. To the left, the speed is seen to decrease near the limit of obtainable vacuum. The constant speed section results from the constant

Fig. 5.



Pumping performance of a pump displayed in the form of a plot of pumping speed vs. inlet pressure.

gas arrival rate at molecular flow conditions and a constant capture efficiency of the vapor jets. The part marked "overload" is a constant throughput section that indicates that the maximum mass flow capacity of the pump has been reached. In the last section, at the right, the performance is highly influenced by the size of the mechanical backing pump. For steady-state operation, the pump should not be operated in the overload region.

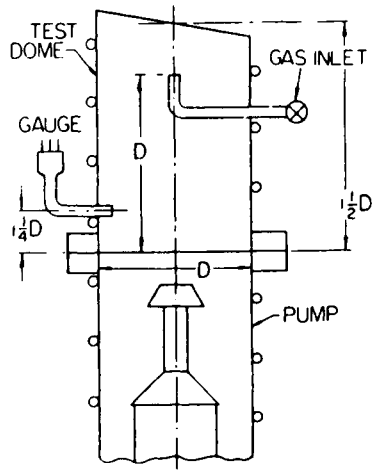
The performance of vapor jet pumps is fundamentally similar to any other pump, compressor, blower, ejector, and similar devices. The essential elements of flow and pressure relationships are analogous.

2.4.2 PUMPING SPEED

2.4.2.1 System Speed and Pump Speed

The pumping speed measured by the AVS Recommended Practice[2] refers to the inlet plane of the pump, Figure 6. Ducts connecting a pump to a chamber, baffles and traps impede flow, resulting in a pressure difference or pressure drop. Under molecular flow conditions, it is common for a baffle or a trap to have a conductance numerically equivalent to the speed of the pump. Thus, the net speed at the pumping port in the chamber can be easily half or a third of the pump speed.

Fig. 6.



Recommended standard test setup.

In steady-state flow, throughput remains constant so that the maximum throughput capacity of the pump is not affected by baffles, orifices, etc. The lower net speed at the chamber results, of course, in higher pressure for any given gas load. Because of outgassing, the gas load in high-vacuum systems is never zero, and the ultimate pressure in the vacuum chamber, therefore, is always higher than the ultimate pressure of the pump itself.

2.4.2.2 Speed Efficiency-Capture Probability

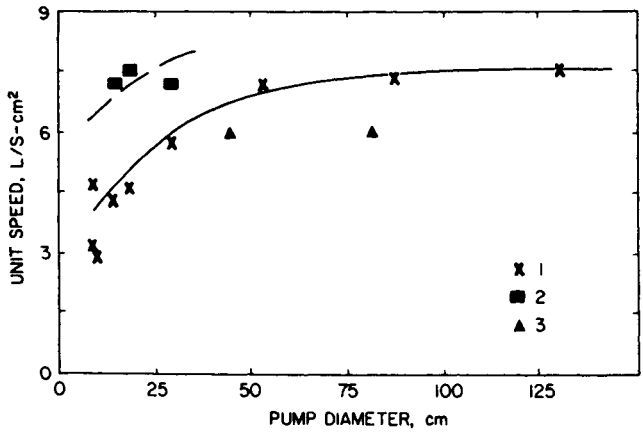
The speed is determined by the American Vacuum Society Practice as

$$S = Q/(P - P_0),$$

where Q is the flow rate (throughput) and P_0 the ultimate pressure prior to the experiment. Figure 6 shows the standard recommended test setup.

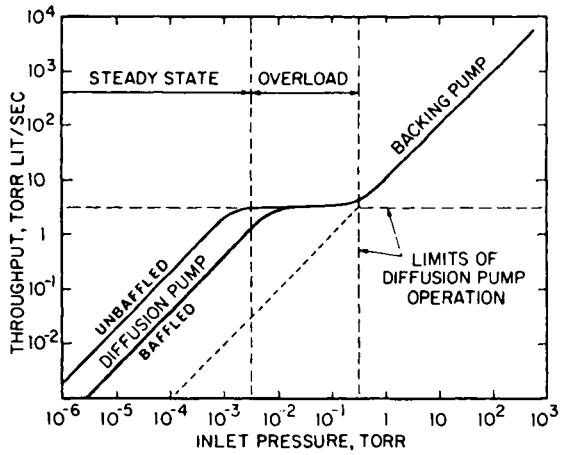
The pumping speed of vapor jet pumps is nearly proportional to the inlet area, but the larger pumps are somewhat more efficient as shown in Figure 7. The entrance geometries of small and large pumps are not strictly similar. Thus, it is possible that the largest pumps can be constructed with a speed efficiency of 50% referred to the inlet plane rather than the conical surface, where pumping action occurs. This can become an important consideration in systems where the desired pumping speed is so high that there is simply not enough wall space available for attaching additional pumps.

Fig. 7.



Efficiency of large pumps (item 2—bulged body pumps).

Fig. 8.



Throughput vs. inlet pressure graph.

2.4.2.3 Speed and Throughput

Pumping speed versus the inlet pressure relationship has caused some misconceptions about the pressure range in which vapor jet pumps can be used. Thus, it is often assumed that the pumps are unstable at system pressures above 10^{-3} torr. Such judgments should not be made without regard to the pump size and system gas load.

It is much simpler to see this relationship by looking at a throughput versus inlet pressure graph, Figure 8. Any combination of throughput and pressure included in the outlined region can be chosen for operation, provided that the pumped gas admittance is restricted (throttled) at inlet pressures above approximately 1×10^{-3} torr. The pressure stability region can be extended even to 10^{-1} torr (dashed diagonal line in Figure 8) if the pumping speed is reduced. It also depends to some extent on the size of the backing pump (as indicated in Figure 5).

2.4.2.4 Size Effects

Vapor jet pumps are made from 5 cm (2 in) to 120 cm (48 in) inlet flange sizes. An obvious difference between the smaller and larger pumps is the distance that the pumping fluid (oil vapor) must travel from the nozzle to the condensing surface or pump wall.

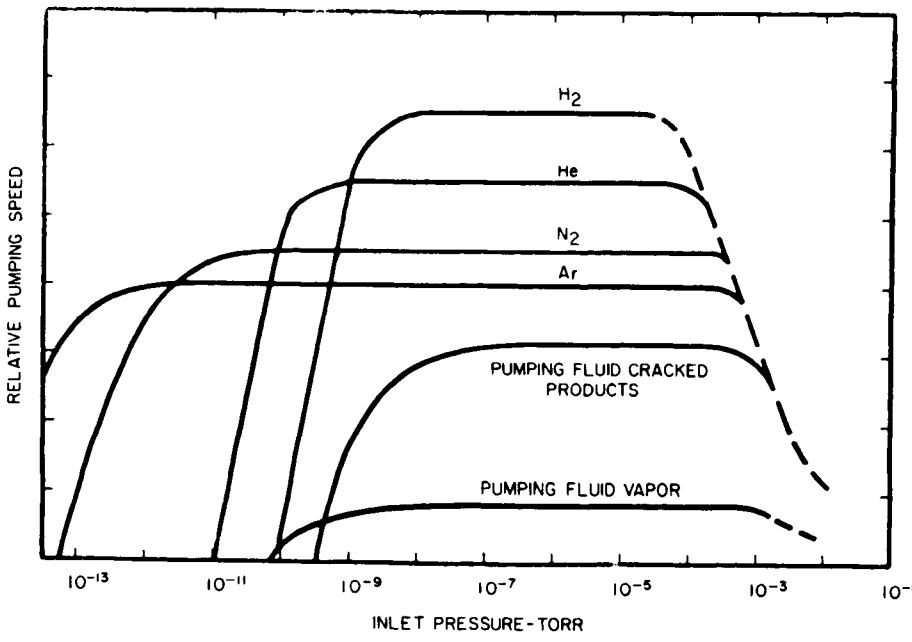
The boiler pressure in small and large pumps is approximately the same be-

cause we are limited in maximum temperature of the oil to avoid thermal breakdown. Therefore, the vapor density at the nozzle exit is nearly the same for all pumps. However, the vapor expands in both axial and radial directions and we may assume that the density is inversely proportional to the square of the distance from the nozzle. Thus, near the pump wall the jet is rare enough to be less effective in pumping gas molecules at higher pressures.

2.4.2.5 Speed for Various Gases

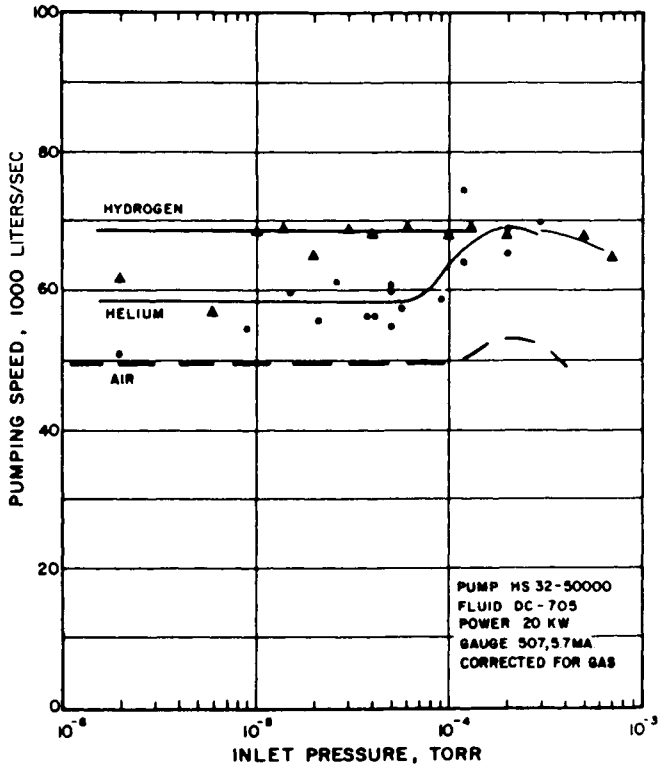
The pumping speed must be considered in relation to the partial pressure of each gas species. When pumping speed is measured, the values obtained near the ultimate pressure of the measurement system become meaningless. The total pressure values cannot be used for obtaining speed due to the uncertainties of gas composition and the condition of the gauge. The composite picture should look as shown in Figure 9. Each gas has its separate speed and, what is more important, separate ultimate pressure. The normally measured "blank-off" is due to pump fluid vapor, cracked fractions, and perhaps water vapor remaining in the system.

Fig. 9.



Pumping speed for various gases that may be present in a test chamber.

Fig. 10.



Typical set of results for helium and hydrogen in a large pump. Pump Varian = HS 32-50,000: fluid DC-705: power = 20 kw; gauge Varian = 507, 5 ma, corrected for gas.

Although the arrival rate, expressed in liters per second, should be inversely proportional to the square root of molecular weight, the lighter gases are not pumped with the same efficiency as air (or nitrogen) and pumping speed values for different gases do not peak at the same power setting. When actual values are necessary, separate measurements must be made. Usually the gases present in vacuum systems, such as water vapor, carbon monoxide and dioxide, nitrogen, and argon are pumped at approximately the same speed. Helium speed is about 20% and hydrogen speed about 30% higher than that of air. The impedance of baffles and traps is lower for lighter gases compared to air. Thus, the net system speed values for lighter gases are relatively higher than those obtained for the same baffle with air. Figure 10 shows a typical set of results for helium and hydrogen in a large pump. Note that the maximum throughput for hydrogen is nearly three times higher than for air, corresponding to the square root of molecular weight ratio. Thus, maximum throughput for argon is 15% lower compared to air.

2.4.3

THROUGHPUT

2.4.3.1 Maximum Throughput

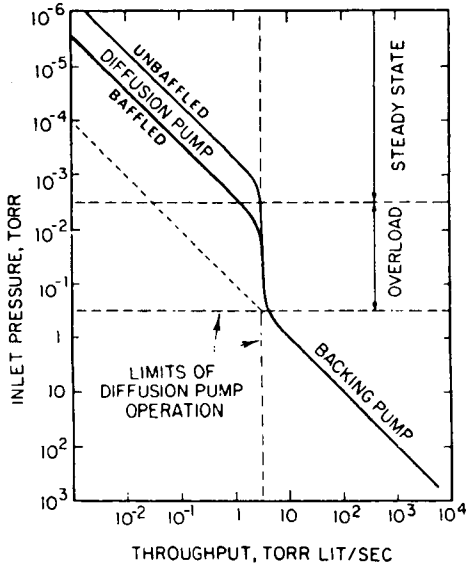
It may be observed that maximum throughput is often the important aspect of vapor jet pump performance rather than pumping speed. The value of the maximum throughput determines the amount of power required to operate a given pump. Dimensionally, throughput and power are equivalent. For pumps of current designs using modern pumping fluids, approximately 1 kw of power is required to obtain a maximum throughput of 1.2 torr L/sec.

For systems that remain under high vacuum for long periods of time, the maximum throughput is of little value. In such cases, provisions can be made to reduce the power after the initial evacuation or to use lower-power heaters. Operation of the pump at half power is possible without changes in pump design. Special designs can be made for low-power (low-throughput) operation with appropriate attention to corresponding reduction of forepressure tolerance.

For rapid frequent evacuation and for high gas load application, the value of maximum throughput determines the choice of pump size.

If the graph in Figure 8 is replotted, interchanging the axis, a more customary arrangement of dependent and independent variables is obtained, Figure 11. Then

Fig. 11.



Input pressure vs. throughput.

it is simpler to see the maximum throughput and the pressure stability limits and the concept of overload when those limits are exceeded. The more convenient way of looking at this is to keep in mind that for a given system gas load presented to the pump there is a resulting inlet pressure. This will help in selecting the required pump size and in distinguishing between the requirements of evacuating a chamber and maintaining a desired operation pressure at a given process gas load.

2.4.4 TOLERABLE FOREPRESSURE

2.4.4.1 Discharge Pressure

A vapor jet pump is designed for high-vacuum application. Thus, its boiler pressure usually is 1–1.5 torr. This implies that the maximum pressure difference the pump can be expected to produce is 1.5 torr. In addition, vapor jet pump working fluids cannot be boiled at high pressures because the resulting high temperatures will decompose the fluid at a unacceptable rate. Thus, such pumps must always be backed by another pump that produces a pressure of generally less than 0.5 torr at the discharge of the vapor jet pump.

Tolerable forepressure of a vapor jet pump is the maximum permissible pressure at the foreline.

The pumping action of the vapor jet pump collapses completely when the tolerable forepressure is exceeded [2]. Essentially, the vapor of the discharge stage of the pump does not have sufficient energy and density to provide a barrier for the air in the foreline when its pressure exceeds a certain value (usually near 0.5 torr). Then this air will flow across the pump in the wrong direction, carrying with it the pumping fluid vapor.

Approximately half of the initial pressure is recovered in the form of tolerable forepressure. When the tolerable forepressure is measured, a conventional nonultrahigh vacuum system usually does not reveal the dependence between the discharge and the inlet pressures. This is because the maximum pressure ratio for air is usually not exceeded unless the experiment is conducted under ultrahigh vacuum inlet pressures.

2.4.4.2 Backing Pump Requirements

To select an appropriate backing pump for a given vapor jet pump, several questions must be considered. First, what is the size of the initial roughing pump, and

is it to be used for both roughing and backing? Second, is the backing pump expected to perform at the maximum throughput of the vapor jet pump? Third, what is the tolerable forepressure of the diffusion pump? Also, what is the volume of the foreline ducts (or a special reservoir or ballast chamber that could be used in the line as a buffer during roughing)? The nominal pumping speed of the required backing pump for full load condition is obtained as follows:

$$S = Q_{\max}/(TFP),$$

where $Q_{\max}$ is the maximum throughput of the vapor jet pump and *TFP* stands for tolerable forepressure. Setting aside consideration of safety factors and conductances of forelines, it should be noted that mechanical pumps often have reduced speed when their inlet pressure is equal to the tolerable forepressure of the vapor jet pump. A safety factor of 2 may be required to make sure that the tolerable forepressure value is never exceeded even if mechanical pump and vapor jet pump are not operating at their best. A numerical example follows. Assume a pump with a maximum throughput of 4 torr L/sec and tolerable forepressure (maximum permissible discharge pressure) of 0.5 torr at full load (at maximum throughput). The required backing pump speed then becomes

$$s = \frac{4 \text{ torr L/sec}}{0.5 \text{ torr}} = 8 \text{ L/sec} \approx 17 \text{ CFM}$$

A good choice for the backing pump would be a nominal 14 liters/sec (30 CFM) pump, provided that the conductance between the two pumps are not severely limited.

2.4.4.3 Safety Factors

The most important rule of vapor jet pump operation is: do not exceed tolerable forepressure! In other words, in an operating pump, the maximum permissible discharge pressure should not be exceeded under any circumstance. Observance of this most basic requirement will eliminate most of the difficulties encountered with vapor jet pumps, especially problems with noticeable backstreaming of the pumping vapor into the vacuum system. High-vacuum systems should be designed with interlocks, fail-safe valve arrangements, or clearly marked instructions to preclude the possibility of exceeding the tolerable forepressure. As in most engineering considerations, a safety factor should be included in establishing the maximum permissible discharge pressure. A factor of 2 is a good general recommendation. As much as 25% reduction of tolerable forepressure can be expected near maximum throughput operation (full load), and some reduction can be expected from low heater power. The selection of the mechanical backing pump must be made with these points in mind. Also, it should be noted that the desired

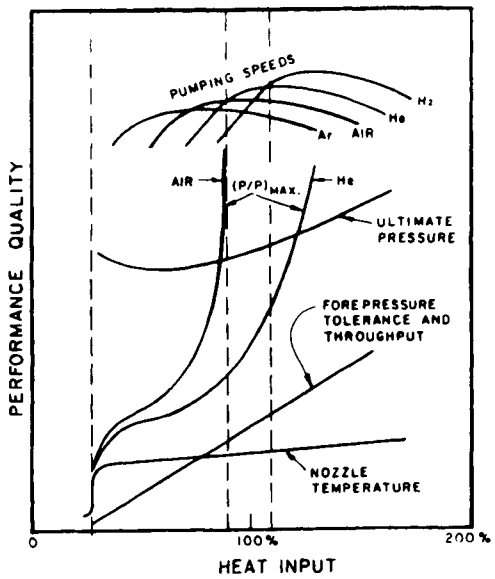
discharge pressure must be maintained near the vapor jet pump to avoid errors caused by limited conductance of forelines and the reduction of mechanical pump speed at lower pressures.

2.4.4.4 Tolerable Forepressure and Pressure Ratio

Tolerable forepressure (maximum permissible discharge pressure) and the maximum pressure ratio achievable with a given pump should not be confused. Tolerable forepressure is directly related to the boiler pressure, while the pressure ratio has a logarithmic dependence on vapor density in the jets. This distinction is indicated in Figure 12. For air the maximum pressure ratio is usually so high that it cannot be measured under normal circumstances. Only when inlet pressure is less than 1×10^{-10} torr can a dependence between discharge and inlet pressure be observed. However, for hydrogen and helium the dependence can be observed in the high-vacuum range.

The effects of boiler pressure variation on pump performance are summarized, qualitatively, in Figure 12. Due to the strong dependency indicated in Figure 12, small changes in vapor jet density may cause large variations in the maximum pressure ratio. In applications where stable pressure is required with lighter gases, this may require attention. Ordinary pumps often display inlet pressure variation

Fig. 12.



Relation of forepressure to boiler pressure and the logarithmic dependence of pressure ratio on vapor density in jets.

exceeding 5% for helium. This is unacceptable for some applications, such as highly sensitive mass spectrometer leak detectors and analytical mass spectrometers where helium is used either as a tracer or a carrier gas. Thus, special pump designs having very stable boiler pressure with power control and stable pumping action may be required.

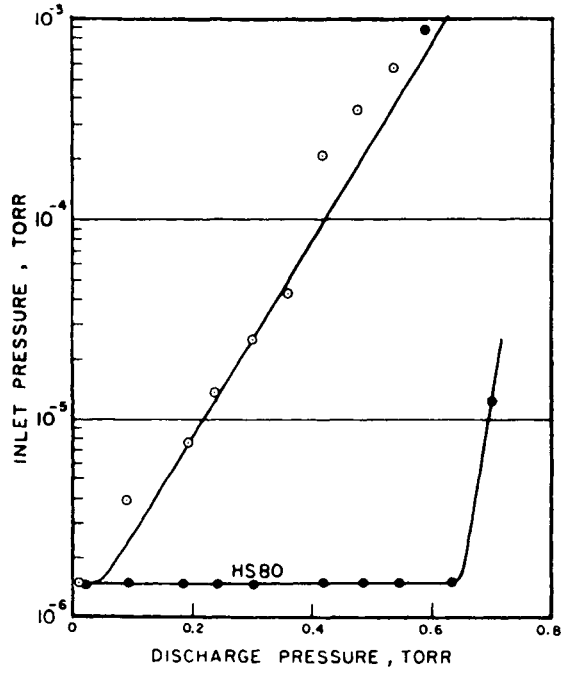
In regard to pumping speed, it can be observed that the curves for helium and hydrogen can have a rather steep slope near the nominal heat input value (100%). Therefore, if variations of vapor density occur, they may result in noticeable pumping speed variations for these gases.

2.4.4.5 Tolerable Forepressure for Various Gases

The tolerable forepressure for various gases is approximately the same as far as the complete collapse of the vapor jet (last stage) is concerned. However, in regard to the appearance of lighter gases at the inlet of the pump, when they are introduced in the foreline the maximum pressure ratio effect becomes noticeable.

The results of pressure ratio measurements are shown in Figure 13. The data were obtained by introducing helium into the foreline of the pump and observing

Fig. 13.



Results of pressure ratio measurements.

the resulting effect on the inlet pressure. The upper curve represents an older 5-cm pump. The comparison between the two pumps indicates a thousandfold improvement in maximum pressure ratio, which can be sustained across the pump.

For the older pump in Figure 13, it is difficult to establish a concept of forepressure tolerance for helium, at inlet pressures below 10^{-6} torr. The newer pump shows a behavior that can be interpreted as if a forepressure tolerance of 0.6 torr for helium is present. This pump had a forepressure tolerance for air of about the same magnitude.

2.4.5 ULTIMATE PRESSURE

2.4.5.1 Pressure Ratio Concept

Two distinct observations can be made regarding the ultimate pressure of a pump. Ultimate pressure may be considered to be a gas load limit or a pressure ratio limit. Both are of significance in practice, the latter usually only with light gases. The pumping action of vapor jets does not cease at any pressure, however low. The ultimate pressure of the pump depends on the ratio of pumped versus backdiffused molecules, plus the ratio of gas load to the pumping speed. In addition, the pump itself can contribute a gas load through backstreaming of pump fluid vapor and its cracked fractions and the outgassing from its parts. Thus, in practice, the observed total ultimate pressure is a composite consisting of several elements. In practice, the commonly observed first limit is due to the pumping fluid, although with the best fluid some system degassing (baking) may be necessary to observe this limit when it is below 1×10^{-8} torr.

When liquid nitrogen traps are employed, the limiting pressure is usually given by various gas loads in the system, even if the system consists only of the measuring dome and a pressure gauge. Below 5×10^{-9} torr thorough baking is usually required.

For helium and hydrogen, a true pressure ratio limitation can be observed in most pumps, although special designs can be made to improve the limit beyond observable level.

2.4.5.2 Baffle and Trap Effects

Room temperature baffles at the inlet to the pump do not affect the ultimate pressure except by shielding a gauge from a direct entry of pumping fluid vapor into the gauge. Water-cooled baffles suppress the rate of reevaporation of condensed

or intercepted fluid, thereby reducing the density of vapor in the space between the baffle and the trap. For substances such as vapor jet pump fluids, each 20°C temperature change near room temperature will result in about an order of magnitude change in vapor pressure and hence the rate of evaporation. The lowered vapor density will reduce the possibility of intercollisions and consequent bypass through the trap without touching a refrigerated surface.

Cryogenic or refrigerated traps have basically two effects. They act as barriers for the flow of condensable vapors from pump to system, but they also act as cryopumps for condensable vapors emanating from the system. The latter may be the primary effect on the ultimate pressure in many cases. In the high-vacuum region and for unbaked systems using modern low-vapor-pressure pumping fluids, the reduction of pressure (when traps are cooled) is primarily due to water vapor pumping. In unbaked systems after the initial evacuation, water may constitute 90% of the remaining gases and cooling of the trap simply increases the pumping speed for water vapor (usually by a factor of 2 or 3).

2.4.5.3 Pressure Ratio for Lighter Gases

As noted previously, the pressure ratio can be sufficiently small for the light gases to reveal the dependence of inlet pressure on the discharge pressure. Measurements of pressure ratio for various gases have been reported as follows: hydrogen 3×10^2–2×10^6, helium 10^3–2×10^6, neon 1 or 2×10^8, CO and argon 10^7, oxygen and krypton 3–5×10^7, and hydrocarbons (C_nH_{2n+2}) 7×10^8. In modern pumps the helium pressure ratio is closer to 10^7, and it can be increased even as high as 10^{10} by doubling the heat input. In practice, even an ion gauge operated in the foreline can produce sufficient hydrogen to cause an increase of the inlet pressure (the same occurrence has been observed in a turbomolecular pump system).

As far as ultimate pressure is concerned, hydrogen can be a substantial part of the residual gas composition due to its presence in metals, in the pumping fluid, and in water vapor. This can be an important consideration for ultrahigh vacuum work where some diffusion pumps may need a second pump in series. The same consideration may apply to helium in leak detectors, mass spectrometers, molecular beam experiments, etc.

2.4.5.4 Pumping Fluid Selection

A variety of "organic" liquids have been used as motive fluids in vapor jet pumps. Criteria for the selection of the fluid are low vapor pressure at room temperature, good thermal stability, chemical inertness, nontoxicity, high surface tension to

Table I

Properties of Diffusion Pump Fluids

Trade name	Chemical name	Molecular weight	Vapor pressure at 20°C (torr)	Flash point (°C)	Viscosity at 20°C (cS)	Surface tension at 20°C (dyn/cm)
Octoil	Diethyl hexyl phthalate	391	10^{-7}	196	75	<30
DC-704	Tetraphenyl tetramethyl trisiloxane	484	10^{-8}	216	47	30.5
Apiezon C	Paraffinic hydrocarbon	574	4×10^{-9}	265	295	30.5
DC-705	Pentaphenyl trimethyl trisiloxane	546	5×10^{-10} (25°C)	243	170 (25°C)	<30.5
Santovac-5	Mixed five-ring polyphenyl ether	447	1.3×10^{-9} (25°C)	288	2500 (25°C)	49.9
Neovac SY	Alkyldiphenyl ether	405	$<1 \times 10^{-8}$	230	250 (25°C)	<30
Fomblin 25/9	Perfluorinated polyether	~3300	2×10^{-9}	None	270	20 (25°C)
Krytox 1625	Perfluoro polyether	~4600	2×10^{-9}	None	250	19

minimize creep, high flash and fire points, reasonable viscosity at ambient temperature, low heat of vaporization, and, of course, low cost. A list of the presently popular fluids and their properties is given in Table 1. The selection of the fluid should be made giving due consideration to its operational stability in the pump boiler. A breakthrough in the selection of fluids was made with the use of five-ring polyphenyl ethers consisting of chains of carbon atoms interbonded by oxygen [3]. This fluid offered exceptional thermal and chemical stability and enabled reaching ultimate pressures of 10^{-9} torr (approaching its vapor pressure at ambient temperatures) with only water-cooled baffles.

Operational characteristics of another low-vapor-pressure silicone fluid (DC-705) were discussed by Crawley et al. [4]. Ultimate pressures of 10^{-9} torr with water-cooled baffles and 10^{-10} torr with a baffle at $-20°C$ were reported [5].

Whatever fluid is used, its vapor may pervade the pumped system, dependent on its vapor pressure, the pump design, and the type of trapping used. The vapor can be broken down by the presence of hot filaments and bombardment by charged particles. The polymerization of silicone fluids resulting from bombardment by charged particles may cause an insulating film to be produced on electrode surfaces, changing the characteristics of the electronic instrumentation. Octoil or polyphenyl ethers are usually recommended to eliminate this problem in applications where mass spectrometers and other electron optical devices are used.

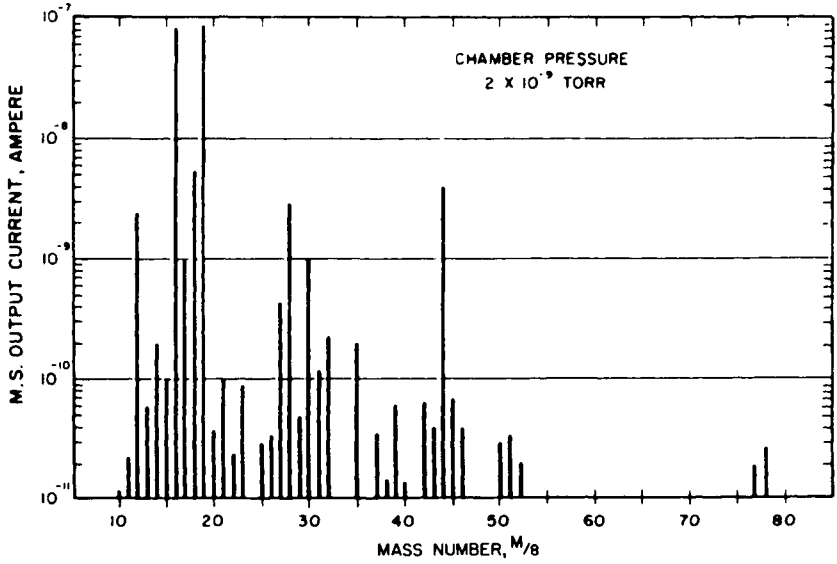
2.4.5.5 Residual Gas Analysis

With condensable gases of high molecular weight, it is extremely difficult to correlate the ion currents indicated by the mass spectrometer with the rate of backstreaming through the baffle. Generally, the residual gas analysis of a trapped diffusion pump system cannot be reliably performed with spectrometers having poorer detectability than 10^{-9} torr. When the spectrometer tube and other parts of the system are baked, the results can be very misleading. It may take weeks before equilibrium conditions are established.

Qualitative measurement can be made, however, with no more difficulty than the ultra-high pressure measurements made with total pressure gauges such as an ionization gauge.

A typical mass spectrum from a baffled vapor jet pump system, unbaked, operating in the ultra-high-vacuum region, with DC-705 motive fluid, is shown in Figure 14. The mass numbers 16, 19, and 35 are unusually high owing to the characteristic properties of the particular spectrometer tube and the hydrogen peak is not shown. The peaks 50, 51, 52, 77, and 78 are characteristic for the pumping fluid. Crawley [4] also gives the residual gas analysis of a vapor jet pump system using silicone DC-705 fluid under both unbaked and baked conditions. Wood and

Fig. 14.



Typical mass spectrum from a baffled pump. Chamber pressure is 2×10^{-9} torr.

Roenigk [6] and Cleaver and Fiveash [7] provide the mass spectra from which the significant identifying peaks can be chosen for the detection of the pumping fluid.

2.4.5.6 Fluid Breakdown and Purification

The ultimate pressure in a vacuum system consists of the partial pressures of various gases. Components arising from gas evolution from sources other than the vapor jet pump (including the inlet gasket) are not discussed here. From the components arising in the pump, the major constituents are the vapor of the motive fluid, the products of its decomposition in the boiler, the contaminant vapors from other system parts, and back-diffused gases. In a system without the use of cryogenic trapping, the ultimate pressure due to the vapor pressure of the fluid at the ambient temperature is the minimum achievable. In pumps that have a horizontal ejector stage in the foreline, in addition to increasing the forepressure tolerance, the ejector stage maintains low pressures in the region under the lower annular stage. With this arrangement, the volatile components of the returning condensate are more easily removed to the forevacuum region. Pumps with a strong side ejector stage are very little influenced by the addition of a booster vapor jet pump for the improvement of ultimate pressures even in the 10^{-10} torr range [8].

With the use of cryogenic traps and modern fluids, such as DC-705 or Santovac-5, ultimate pressures of 10^{-10} torr and below can be achieved where contaminants from sources other than the vapor jet pump are controlled.

An increase of heat input will usually increase the pressure ratio obtainable with a pump. However, depending on pump design, higher power can either improve or worsen the ultimate pressure. This depends on the ratio between fluid thermal breakdown rate and the purification ability of the particular pump.

2.4.6 BACKSTREAMING

Any transport of the pumping fluid into the vacuum system may be called backstreaming. The possibility of this backflow is perhaps the most undesirable characteristic of vapor jet pumps. It is in this area that most misunderstanding and misinformation exists. Pump manufacturers usually report the backstreaming rate at the inlet plane of an unbaffled pump. System designers are often concerned about the steady-state backstreaming above liquid nitrogen traps. The users are often troubled by inadvertent high-pressure air inrush into the discharge end of the pump. The quantities of backstreaming fluid involved in these cases can range over many orders of magnitude.

As far as the pump itself is concerned, there can be several sources of backstreaming:

1. The overdivergent flow of vapor from the rim of the upper nozzle
2. Poorly sealed penetrations at the top nozzle cap
3. Intercollision of vapor molecules in the upper layer of the vapor stream from the top nozzle
4. Collisions between gas and vapor molecules, particularly at high gas loads (10^{-3} to 10^{-4} torr region)
5. Boiling of the returning condensate just before the entry into the boiler (between the jet assembly and the pump wall), which sends fluid droplets upward through the vapor jet
6. Evaporation of condensed fluid from the pump wall

2.4.6.1 Primary and Secondary Backstreaming

All the items just listed that can be stopped or intercepted by the room temperature baffle may be called primary backstreaming. The reevaporation of the pump-

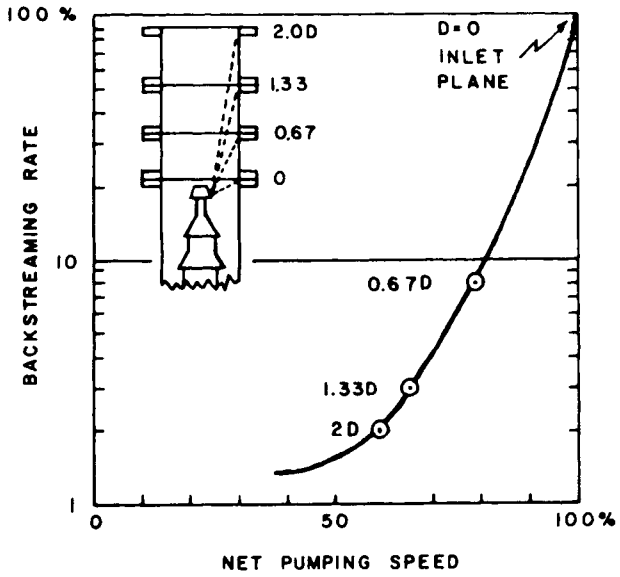
ing fluid from the baffle and the passage through the baffle may be called secondary backstreaming. The primary backstreaming can be effectively controlled by the use of cold caps surrounding the top nozzle. Cold caps are water-cooled baffles located above the top jets of a pump.

In systems with liquid nitrogen traps (barring accidents and high gas load operation), the backstreaming level can be controlled at such a low level that contaminants from sources other than the vapor jet pump will predominate [9]. Properly operated and protected vapor jet pump systems can be considered to be free of contamination from the pumping fluid for most applications.

2.4.6.2 Speed and Backstreaming

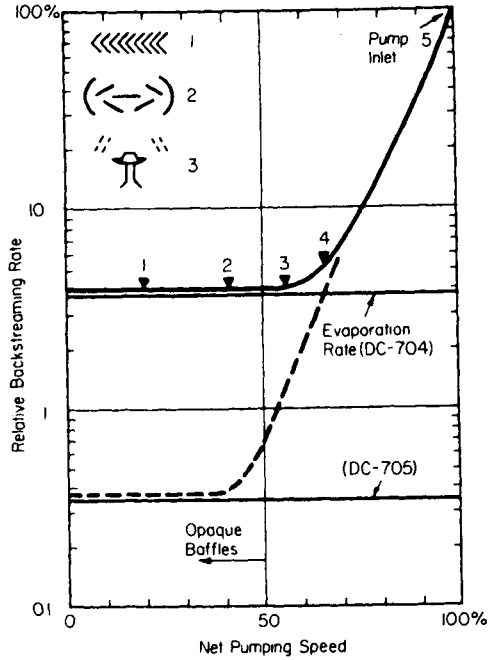
The backstreaming rates measured by the American Vacuum Society Recommended Practice and Manufacturer's Specifications refer to the inlet plane of the pump. It is important to understand that this rate quickly diminishes at some distance from the pump inlet. At a distance from inlet equivalent to two pump diameters, the rate is reduced typically 50 times (Figure 15). A 90-degree bend in the inlet duct will act as a baffle and reduce backstreaming to the level of the natural

Fig. 15.



Reduction of backstreaming rate.

Fig. 16.



Qualitative relationship between conflicting requirements of minimum backstreaming and maximum retaining pumping speed with baffles at ambient temperature.

evaporation rate of the fluid at the ambiance temperature. Without cryogenic traps or other similar devices, this rate cannot be reduced much further.

Figure 16 shows the qualitative relationship between conflicting requirements of minimum backstreaming and maximum retained pumping speed with baffles at ambient temperature. If a series of elements that reduce backstreaming is introduced above the pump, the backstreaming rate approaches the evaporation rate of the fluid. It may be observed that in many applications, complete opaque baffles are redundant. With efficient cold caps and low-vapor-pressure pumping fluids, a system can be operated for a year or more with continuous liquid nitrogen cooling of the trap. The condensed pumping fluid will be returned into the pump, and the buildup on the trap will not be detrimental.

In normal practice, an optimum design for a vapor jet pump and trap combination can have nearly 40% net pumping speed and reduce backstreaming to less than 1×10^{-10} g/cm^2 min (at the inlet plane of the trap). Values of this magnitude have been measured [10].

2.4.6.3 Surface Migration

Some pumping fluids may have a tendency to spread on metal surfaces as a thin film. This spreading is discussed in the literature dealing with lubrication of small instruments [11]. The spreading may be noticed sometimes as wetness on rough (sand-blasted) surfaces. Modern low-vapor-pressure pumping fluids having surface tensions above 30 dyne/cm do not spread on ordinary metal surfaces. They are so-called autophobic liquids because they do not spread on their own monolayers covering a metal surface.

Surface diffusion is not likely to be a significant backstreaming mechanism on surfaces at room temperature or below for pumping fluids with vapor pressures in the 10^{-9} torr range and high surface tension.

2.4.6.4 Accidental Backstreaming

It is important to remember that malfunction and misoperation can destroy the intentions of most intelligent designs. The most common causes of gross backstreaming are accidental exposure of discharge side of the vapor jet pump to pressure higher than the tolerable forepressure, high inlet pressure exceeding maximum throughput capacity over long periods of time, incorrect startup, and incorrect bakeout procedures.

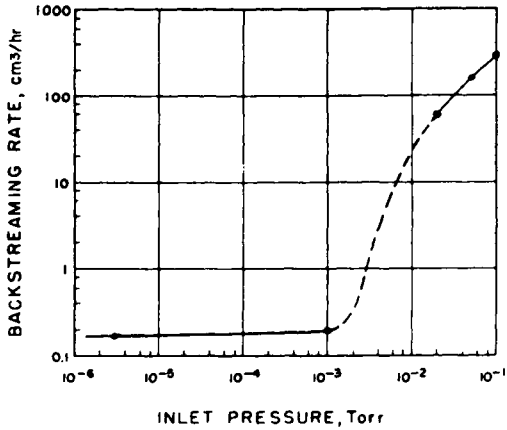
BACKSTREAMING RATE VS. PRESSURE

From lowest pressures to about 1×10^{-4} torr, the backstreaming rate appears to be independent of pressure indicating that oil–gas collisions are not significant in this region. Between 10^{-4} and 10^{-3} torr, a slight increase may be noticed [12]. Above the critical pressure point when the maximum throughput is reached (pumping speed begins to decline), the backstreaming rate may rise markedly.

Measurements of the backstreaming rate at various inlet pressures made under undesirable conditions are plotted in Figure 17. The tests shown were obtained with a collecting surface in the immediate vicinity of the top nozzle. If the measurements were made at the inlet plane as specified by the American Vacuum Society Recommended Practice, the rate would have been 10 or 100 times lower (see Figure 15). It can be seen that the rate does not change significantly as long as the inlet pressure is below about 10^{-3} torr. This is the point where most vapor jet pumps have an abrupt reduction in speed indicating that the top jet essentially stops pumping.

The sudden increase of backstreaming at inlet pressures above 10^{-3} torr points

Fig. 17.



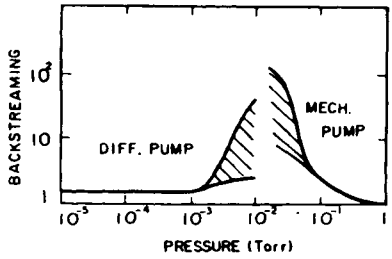
Measurements of backstreaming rate at various inlet pressures.

out that a vapor jet pump should not be operated in this range for more than a brief time. It is common practice in vacuum system operation to open the high-vacuum valve after the chamber has been rough pumped to about 10^{-1} torr exposing the pump to this pressure. Extended operation at high pressure can direct unacceptably high amounts of oil into the vacuum system. (See Section 2.1.9 for more information on crossover procedures.)

The higher-pressure conditions exist at the inlet to the pump briefly after the high-vacuum valve is opened. Unless the vapor jet pump is severely overloaded, the associated backstreaming peak normally lasts only a few seconds.

The same condition will exist when the pump is initially heated. In this case, the duration is a few minutes and the amplitude a few times higher than the steady-state backstreaming at lower pressures. The generalized picture is shown in Figure 18. At pressures higher than 0.5 torr, depending on diameter and length

Fig. 18.



Higher-pressure conditions that exist at the inlet to the pump.

of ducts, the pumping fluid vapor may be swept back into the pump. This can be used to reduce initial backstreaming and prevent mechanical pump backstreaming by arranging a flow of air (or other gas) through a leak valve.

2.4.6.6 Mechanical Pump Effects

Another possibility for hydrocarbon contamination arises from the roughing pump. A roughing pump should not be left connected to the system for long periods of time. After initial evacuation, the roughing valve should be closed and pumping switched to the vapor jet pump.

In systems where only a mechanical pump is used for maintaining vacuum conditions in the 10^{-2}–10^{-1} torr range, sometimes it is advisable to pump periodically rather than continuously. A valve and a gauge are needed to start and stop pumping whenever necessary. This prevents continuous back-diffusion of lubricating oil and, in addition, keeps the oil cold, reducing its vapor pressure.

2.4.6.7 Pumping Fluid Loss

The pumping fluid level in a well-designed pump need not be precisely controlled. Generally, 30% above and below normal level should be tolerable without noticeable effects. When the level is too low, the boiling process may pass from nucleated boiling to partial film boiling, which leads to overheating of the boiler heating surface. If this condition is continued for an extended period of time, particularly for large pumps, it can cause distortions of the boiler plate. This, in turn, may expose the center of the boiler plate above the liquid level and lead to further overheating. The resulting poor contact between heaters and the boiler plate may also overheat the heating elements and cause their failure.

If the liquid level is too high, the boiling process may foam the fluid and raise its level as high as the foreline opening.

Excluding normal backstreaming, the fluid may be lost out of the pump in several ways: prolonged operation above or near maximum throughput; accidental high pressure and high-velocity air flow through the pump in either direction; and evaporation of higher-vapor-pressure fluids due to incorrect temperature distribution.

With relatively low gas loads, modern pump fluids, and correct system design and operation, vapor jet pumps can be operated for many years without adding or changing pumping fluid. Operation exceeding 10 years has been reported in the

case of a particle accelerator [13]. Large pumps usually have means of monitoring the fluid level. To reduce fluid loss, some pumps have built-in foreline baffles.

2.4.6.8 Fluid Films on LN Traps

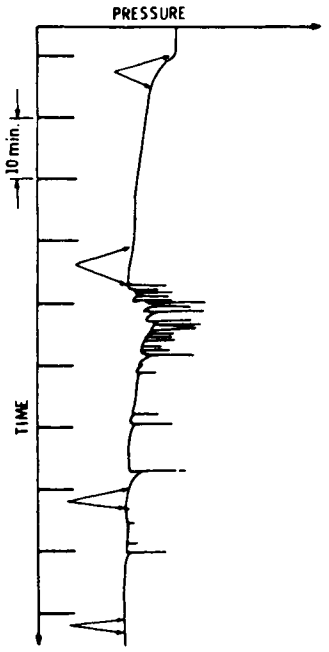
An unusual mechanism of gross backstreaming has been observed in vapor jet pump systems with liquid nitrogen traps. It occurs sometimes during or following the charging of the trap with liquid nitrogen and can be recognized by the appearance of small droplets upstream from the trap and by accompanying pressure fluctuations. The phenomenon is produced by fracture of the frozen pumping fluid film due to the unequal temperature expansion coefficients between it and the metal surfaces on which it has been previously deposited. The elastic energy stored in the film is sufficient to impart high velocity to the fragments resulting from the fracture. The design can be improved by avoiding certain geometric configurations that tend to accumulate heavy fluid films and produce highly stressed films that are likely to fracture.

Liquid nitrogen traps generally should not be placed too close to unprotected vapor jet pumps. Ambient or water-cooled baffles, partial baffles, or efficient cold caps essentially remove primary backstreaming. Traps that have removable internal structures should be cleaned periodically. The frequency will depend on the vapor pressure of the pumping fluid and possibly the degree of accumulation of other condensables such as water vapor, assuming that the trap is continuously kept refrigerated. A thin film (less than 10^{-3} mm thickness) is not likely to fracture. In this connection, traps that have long liquid nitrogen holding time are particularly recommended.

If the trap has to be desorbed occasionally, such as during weekends, the restarting procedure needs some attention. It is good practice to keep the high-vacuum valve closed during and after the trap-filling periods. In this manner, the fractured frozen film particles and products of desorption can be prevented from reaching the vacuum chamber. The valve opening should be delayed after recooling the trap, to allow time for repumping of condensable matter.

With the combination of the design and procedural remedies mentioned, the difficulties associated with the described phenomenon can be avoided. Figure 19 shows a record of a typical pressure variation due to this effect.

Fig. 19.



Record of typical pressure variation when trap is desorbed. Arrows show trap filling periods.

2.4.7

OTHER PERFORMANCE ASPECTS

2.4.7.1 Design Features

In their specifications, manufacturers of vapor jet pumps traditionally report only pumping speed, ultimate pressure, forepressure tolerance (measured at the foreline), throughput, and sometimes the backstreaming rate (referred to the inlet plane of the pump). However, other performance aspects, listed next, can be used as guides in selecting pumps for particular applications:

- *Pumping speed per unit inlet area.* This can be called speed efficiency, and it is not likely to exceed 50% compared to a hypothetical pump with 100% capture probability for molecules that cross the inlet plane into the pump.
- *Power required to obtain a given maximum steady-state throughput without overloading and without oversized forepumps.*

- *Ratio of maximum throughput and forepressure tolerance.* This determines the minimum required forepump speed that must be provided at the discharge of the vapor jet pump (at its full load).
- *Maximum pressure ratio for light gases.* Of particular interest are helium and hydrogen (pumping speed for helium and hydrogen, compared to the speed for air, should be 1.2 times higher).
- *Ratio of evaporation rate of pumping fluid at ambient temperature and the actual pump backstreaming without traps.* This ratio is approaching unity in some modern pumps that have efficient cold caps surrounding the top nozzle.
- *Ratio of vapor pressure of working fluid at ambient temperature and ultimate pressure obtained by pump without cryogenic traps.* The target for this ratio should also be near unity. To achieve this, the pump must have low pumping fluid breakdown level and a high degree of fluid purification.
- *Ratio of forepressure tolerance and boiler pressure.* This ratio is usually about 0.5. The significance of this figure is in keeping the fluid temperature as low as possible to reduce thermal breakdown, while keeping the forepressure tolerance as high as possible.
- *Ratio of pump diameter and height.* The height is normally minimized for the sake of compactness, but some performance improvements could be realized if pumps were allowed to be taller.
- *Pressure stability in constant speed region.* This can be expressed as a percent variation referred to an average value. Pressure instability is more common in smaller pumps and with lighter gases.
- *Sensitivity to heat input fluctuations.* In addition to variations of heater power, cooling water flow rate and temperature may have some significance.

The low degree of fluid breakdown and self-purification qualities of the pump can be judged by the degree with which the ultimate pressure of the pump follows the expected vapor pressure of the pumping fluid at room temperature. Figure 20 illustrates the performance of various fluids in a pump which must have a given boiler pressure (for example, 0.5 torr). In modern pumps, the pressure of the oil vapor inside the jet assembly may be about 1.5 torr. From Figure 20 it can be seen that above the temperature rate of 250°C, the performance of the pump deteriorates, as far as ultimate pressure is concerned. Generally, it is useless to employ fluids of very low vapor pressure because of resulting excessive boiler temperature.

When using extremely low-vapor-pressure fluids (for example, Santovac-5), it is sometimes necessary to adjust water-cooling rate or power to obtain optimum performance. Some pumps may have to operate with more effective heat input, some with less, depending whether they are below or above the valley of the curve in Figure 20.

Fig. 20.

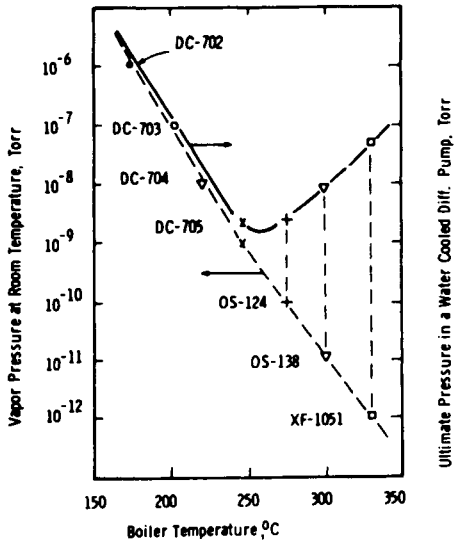


Illustration of performance of various fluids in a pump with a given boiler pressure.

2.4.7.2 Cold Caps

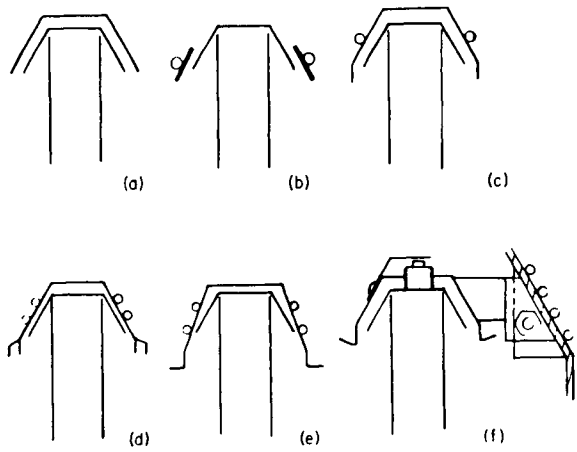
Cold caps surrounding the top nozzle of the pump reduce backstreaming rates 50 or more times [14–18]. In medium-sized pumps, they are usually made of copper and cooled by radiation or conduction through supports contacting the water-cooled pump wall. This allows the fluid drops from the cold cap to fall into the pump, preventing evaporation into the regions outside the pump. Figure 21 shows the evolution of cold cap design.

2.4.7.3 Pressure Stability

Pressure instabilities occasionally seen in vapor jet pump systems can originate in the pump as well as from sources outside the pump. The sources outside the pump are

1. Gas bubbles from the elastomer seals
2. Liquid dripping from the baffle
3. High foreline pressure, light gases
4. Throughput overload

Fig. 21.



Evolution of cold cap design.

5. Trap defrosting
6. Explosive breakup of a frozen layer in the trap

The sources inside the pump are

1. Eruptive or unstable boiling
2. Boiling outside the jet assembly
3. Low pressure ratio, light gases
4. Liquid droplets in the nozzles
5. Cold top nozzle
6. Leaks near the boiler

Vigorous boiling in vacuum causes droplets to splash upward and be blown out through the nozzle opening, causing momentary blockage as well as wet running in the nozzle. A splash baffle located inside the boiler is used to eliminate this effect. Low pressure ratio for light gases can cause pressure instabilities. An increase in the heater power, as well as an increase of the number of stages, is beneficial.

Outside the pump, gas bubbles from the elastomer seals are the major and frequent source of pressure bursts. Gas evolved from and permeating through the elastomer seals is sometimes trapped by oil films. As the pressure builds up in the bubbles, they break through the oil film and release the gas into the system. The inlet flange of the pump is the most vulnerable location. Careful design of the O-ring groove can minimize this instability. Fluid dripping from the baffles on the hot surface of the top jet cap can evaporate producing vapor and temporarily

affect the pumping speeds. Pumps having pressure fluctuations for light gases of $\pm 1\%$ at the pump inlet have been reported [19, 20].

REFERENCES

1. R. Jaeckel, *KleinsteDrücke* (Springer-Verlag, Berlin and New York, 1950).
2. H. G. Nöller, Physical events in diffusion pumps, *Vacuum*, **5** (10) (1955) 59.
3. K. C. D. Hickman, *Trans. Am. Vac. Soc.* (Pergamon, New York, 1961), p. 307.
4. D. J. Crawley, E. D. Tolmie, and A. R. Huntress, *Trans. Am. Vac. Soc.* (AIP for AVS, New York, 1967), p. 67.
5. H. Okamoto and Y. Murakami, *Abs. Am. Vac. Soc. Symp.* (AIP for AVS, New York, 1967), p. 67.
6. G. M. Wood, Jr. and R. J. Roenigk, Jr., *J. Vac. Sci. Technol.*, **6** (1969) 871.
7. J. S. Cleaver and P. N. Fiveash, *Vacuum*, **20** (1970) 49.
8. B. D. Power, *High vacuum pumping equipment* (Van Nostrand-Reinhold, New York, 1966).
9. M. H. Hablanian, *J. Vac. Sci. Technol.*, **6** (1969), p. 265.
10. G. Rettinghaus and W. K. Huber, *Vacuum*, **24** (1974) 249.
11. F. M. Fowkes, Contact angle, *Adv. Chem. Ser. 43*, Am. Chem. Soc. (1964).
12. D. H. Holkeboer, D. W. Jones, F. Pagano, and D. J. Santeler, *Vacuum Engineering* (Technical Publishers, Boston, 1967).
13. J. Moenich and R. Trcka, *Trans. Am. Vac. Soc. and Int. Congr. Vac. Sci. Technol.* (Pergamon, New York, 1961), p. 1133.
14. B. D. Power and D. J. Crawley, *Vacuum*, **4** (1954) (published 1957) 415.
15. S. A. Vekshinsky, M. I. Menshikov, and I. S. Rabinovich, *Proc. Int. Cong. Vac. Tech. 1st.* (Pergamon, London, 1958), p. 63.
16. N. Milleron and L. Levenson, *Trans. Am. Vac. Soc.* (Pergamon, New York, 1960), p. 213.
17. M. H. Hablanian and H. A. Steinherz, *Trans. Am. Vac. Soc.* (Pergamon, New York, 1961), p. 333.
18. M. H. Hablanian and A. A. Landfors, *Trans. Am. Vac. Soc.* (Pergamon, New York, 1960), p. 555.
19. M. H. Hablanian and A. A. Landfors, *J. Vac. Sci. Technol.*, **13** (1976) 494.
20. J. D. Buckingham and N. Dennis, *Vacuum*, **21**, **5** (1975).

CHAPTER 2.5

Cryogenic Pumps

Gary S. Ash
CTI-Cryogenics Division
Helix Technology Corporation

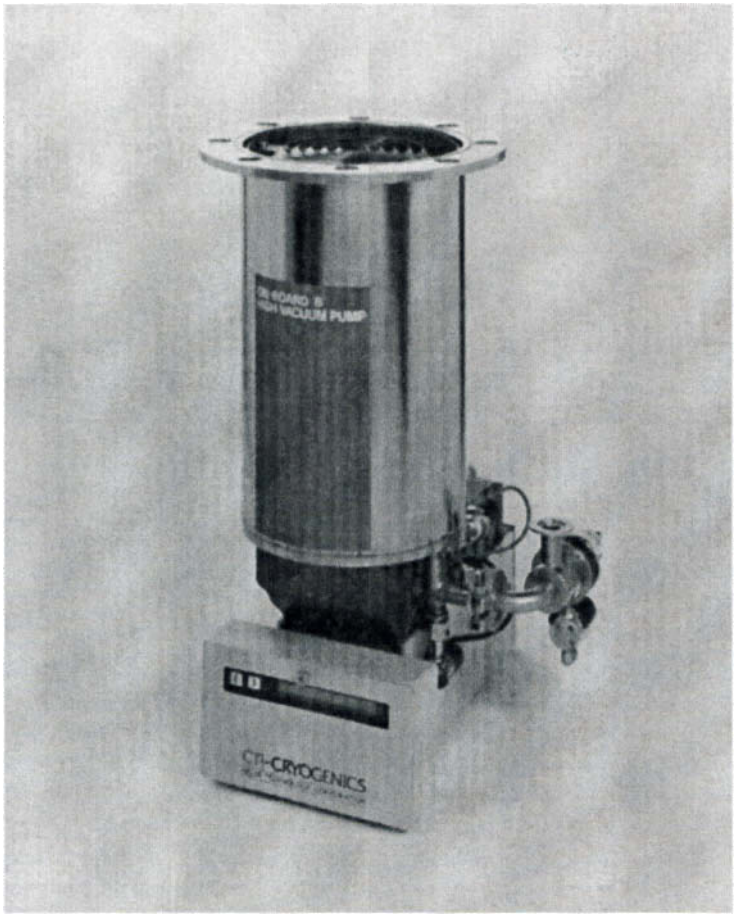
2.5.1 INTRODUCTION

Cryogenic high-vacuum pumps provide clean, high-speed pumping of all gases. They function through a combination of condensation and adsorption of gases and vapors on surfaces held at very low temperatures. Most cryopumps employ sets of surfaces held at two specific temperature ranges by a reciprocating mechanical refrigerator using helium gas as a refrigerant. The helium is supplied at high pressure and room temperature by a remote compressor, and returns to the compressor at low pressure but only slightly warmer than room temperature. All the cold parts of the refrigerator lie within the vacuum environment of the pump, although there are no moving parts within the vacuum. This approach produces a rugged, cost-efficient vacuum pump that is suitable for many applications.

The housing of a cryopump, such as the one in Figure 1, is typically a stainless steel or aluminum vacuum vessel with a high-vacuum flange for mounting to the high-vacuum valve on the system. A two-stage cryogenic refrigerator is bolted or welded into the vessel. The refrigerator surfaces exposed to the vacuum volume are stainless steel. Massive copper heat stations, sometimes plated with nickel for corrosion protection and appearance, are brazed to the cylinder of the refrigerator for efficient heat transfer to the condensing arrays. A copper thermal radiation shield, like a large bucket painted black on the inside, is attached to the first-stage

ISBN 0-12-325065-7
$25.00

Fig. 1.

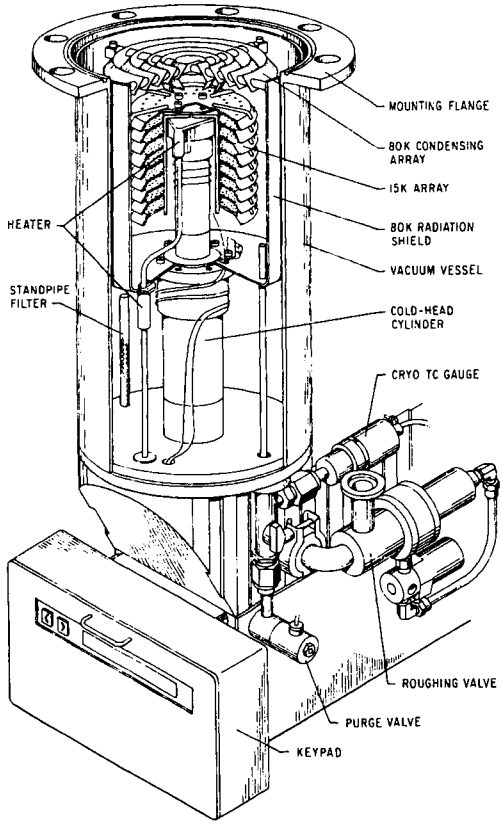


An 8-inch (200-mm) diameter cryogenic high-vacuum pump. (Photo of On-Board 8 cryopump courtesy of CTI-Cryogenics.)

heat station. This thermal radiation shield, kept at 65 K or so by the refrigerator, absorbs incoming heat from the system on the black paint. The outside of the shield is nickel plated to reduce the thermal radiation transfer from the room-temperature vacuum vessel. A nickel-plated copper inlet array, commonly a circular chevron, attaches to the opening of the thermal radiation shield and is also maintained near 65 K. The inlet array acts as a pump for water vapor (see Figure 2).

Within the volume enclosed by the inlet array and radiation shield, a second-stage array set is located. Attached to the lower temperature (10–20 K) heat station, a number of copper disks, plates, or cones are used as a second pump for ar-

Fig. 2.



Components of a cryopump. (Line drawing of cut-away view of On-Board 8 cryopump courtesy of CTI-Cryogenics.)

gon, oxygen, nitrogen, and other gases by condensation. These surfaces are also usually nickel plated. Over much of the second-stage array surface is a layer of activated carbon or charcoal. These provide the third type of pumping within the cryopump: a sorption pump for the light gases hydrogen, helium, and neon. The size, shape, and number of these plates determine the speed and capacity for all the gases to be pumped on the second stage.

2.5.1.1 History

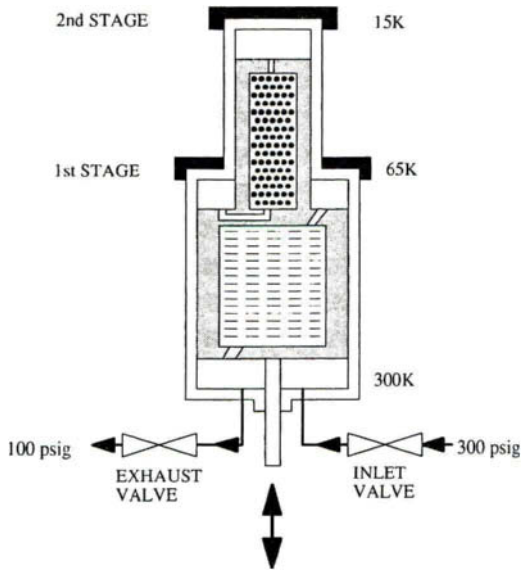
Early cryogenic pumps from the 1930s onward used liquid cryogens, such as liquid nitrogen at 77 K and liquid helium at 4.2 K to produce refrigeration. Although

simple to construct, they were limited to a single orientation and required constant replenishment of cryogens [1]. The most significant breakthroughs in cryopumping technology were the development of the Gifford-McMahon refrigeration cycle and the use of activated charcoal for the sorption of hydrogen, helium, and neon in the 1960s. Originally developed at Arthur D. Little, Inc. for condensing hydrogen at 20 K and for cooling microwave receivers for satellite communication ground stations, the Gifford-McMahon refrigerators proved their capability of running for many years with little or no maintenance in remote sites around the world. Other refrigerator versions were initially developed for cooling infrared detectors and solid-state lasers for military programs.

2.5.1.2 Refrigerator Operation

The Gifford-McMahon cycle is a "no work" thermodynamic expansion cycle using regenerative heat exchangers. In cryogenic vacuum pumps, the refrigerator configuration is usually of two stages of different diameters, one above the other. (See Figure 3.) A pistonlike displacer, with stepped diameters, fits within the cylinder of the refrigerator [2]. The larger diameter portion is packed with hundreds of disks made of metal screen. The smaller-diameter section of the displacer is packed with lead balls or similar materials. The screens and balls are the heat exchanger matrix. The displacer body itself is made of a rugged plastic with low

Fig. 3.



A Gifford-McMahon refrigerator.

thermal conductivity. A number of holes in the displacer body direct the flow of helium gas through the displacer. The screens and balls are packed so as to encourage the relatively free flow of gas through the matrix while exposing the surfaces of the balls and screens to the gas for good heat exchange.

A slow-speed motor or pneumatic drive moves the piston up and down within the cylinder bore at a nominal rate of 60–72 strokes per minute. When the refrigerator is operating, the end of the large-diameter, first-stage cylinder near the drive mechanism is essentially at room temperature. The other end of the first-stage cylinder, to which the first-stage heat station is brazed, is at approximately 65 K. The smaller-diameter, second-stage cylinder is welded to the 65 K end of the first stage. The opposite end of the second-stage cylinder achieves a temperature of about 10–15 K. The exact stage temperatures are determined by the thermal loads placed on them, the design of the displacers in the refrigerator, the supply and return pressures of the helium gas from the compressor, and a number of other variables.

Gas supplied from the compressor at about 20 atmospheres pressure (300 psig) is admitted to the warm end of the cylinder by a mechanically timed inlet valve. The helium flows through the displacer matrices and fills the entire cylinder volume to the supply pressure. As the gas passes through the screens and balls, it is cooled as it goes. In turn, each screen or ball becomes slightly warmer. The moving displacer passes through the gas, eventually reducing the volume of gas at the warm end to nearly zero. This also produces small volumes of gas at 65 K and 10 K at the other end of the cylinders as the displacer reaches the end of its stroke. At this point, the inlet valve closes and the exhaust valve opens to the return line to the compressor. Because the return line is at about 6–7 atmospheres (100 psig), gas within the cylinders expands, cools, and flows back through the displacers. On each stroke, a little heat is removed from each heat station (and attached array) to maintain the operating temperatures. The cold gas flowing back over the balls and screens cools them back to their previous temperatures. As the displacer continues its reciprocating motion, the small cold volumes are reduced, forcing the remaining cold gas back through the displacer matrix. The gas going out the exhaust valve of the refrigerator is slightly warmer than the gas that came into the cylinder. Because the refrigerator cylinders are welded tubing, none of the helium gets into the vacuum space. Similarly, none of the gases condensed in the cryopump ever enter the helium stream.

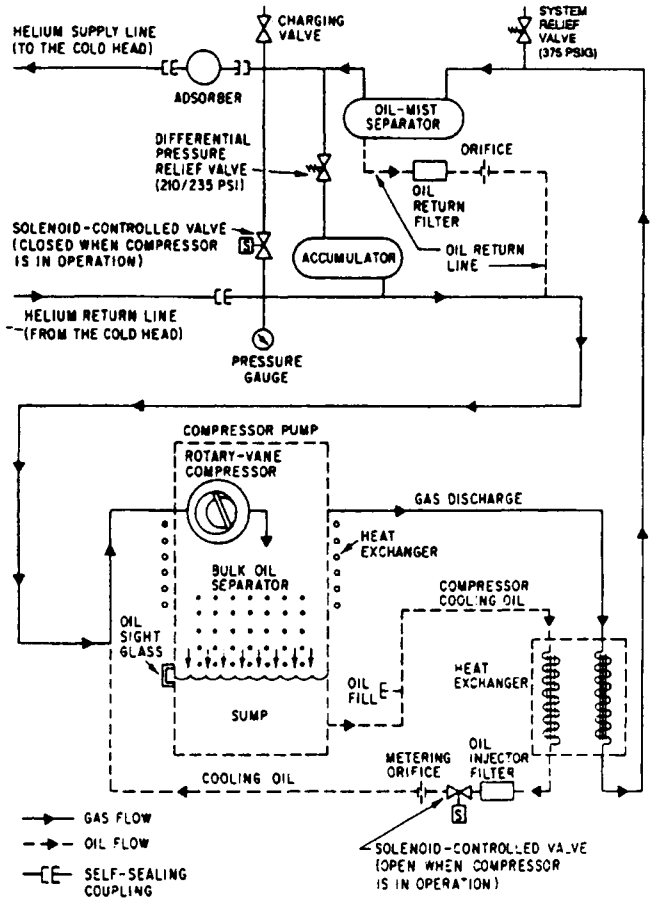
2.5.1.3 Compressors

The compressor performs one major function: It takes the return gas from the refrigerator at 100 psig and boosts the pressure to 300 psig before returning it to the refrigerator. To accomplish this, it must also perform some related functions. Compressing helium generates lots of heat, so the heat of compression must be

removed from the helium and transferred to either cooling water or air supplied from the outside. Since the compressor is a mechanical device requiring lubrication, oil is injected into the returning helium and carried to the compressor pump. The oil also serves as an efficient cooling and heat transfer medium. After compression and cooling through a heat exchanger, the helium stream is stripped of the oil by a mist separator. An additional adsorber cartridge provides a final cleanup of the helium before it goes out to the refrigerator. A typical piping diagram of a compressor is shown in Figure 4. Gas purity is held at the level of tens to hundreds of ppm.

As the compressor only takes in and sends out a relatively constant volume of helium, there is no limit to the separation distance of the compressor and attached

Fig. 4.



Piping circuit of compressor for cryopump showing flow of helium, oil, and cooling water. (Piping diagram of helium compressor courtesy of CTI-Cryogenics.)

cryopumps. Helium line lengths of 10–60 feet are common, but a compressor may be placed hundreds or thousands of feet from the pumps when lines of adequate diameter are used. No loss of refrigerator performance is created by separation distance.

The major factors in cryopump performance controlled by the compressor are the absolute and differential pressures of the supply and return lines. Other things being equal, higher refrigeration capacity is produced by higher differential pressure. Performance is slightly better at lower absolute pressures. Higher refrigeration capacity provides shorter cooldown and regeneration times, as well as greater ability to handle heat loads at constant temperature. Since compressors are made with positive displacement mechanical pumps, increased helium output flow reduces the differential pressure. This means that when more small pumps are connected to one large compressor, the refrigeration capacity of each pump decreases, increasing cooldown time. To limit the maximum pressure differential and absolute pressure to a reasonable level, a spring-loaded bypass valve in the compressor limits the differential pressure when the compressor is on but has no output (cryopumps in regeneration) or only a low gas output. Depending on the application, the performance of the cryopumps may be increased by providing as much compressor capacity as is economically practical. On multipump systems with automated controls, the compressors may account for only 15–20% of the total pumping cost, so that there is great leverage on small investments in additional compressor capacity.

Normally, compressors require little maintenance. At periodic intervals of one to three years, the adsorber should be changed to continue high gas purity. If some helium charge has been lost and the static charge pressure (compressor off) has dropped, 99.999% pure helium gas should be added. In the event of helium gas contamination, as indicated by the onset of knocking noises on each stroke of the refrigerator, the compressor and each of the attached cryopumps should be purged and recharged. In addition, the adsorber should be changed at that time. Compressors will run for many years without any other maintenance. Operating lifetimes of 50,000 to 100,000 hours are not unusual.

2.5.1.4 Applications

High-vacuum evaporators were the first commercial application of cryopumps, where they replaced diffusion pumps. Freedom from oil vapor backstreaming and the ability to reach pressures in the 10^{-7} torr range and below, make cryopumps appealing for use in the deposition of a variety of electronic and optical thin films by evaporation or sputtering. Ultra-high-vacuum versions of cryopumps operate in the 10^{-8} to 10^{-11} torr range for surface science and molecular beam epitaxy systems. In the semiconductor industry, cryopumps serve on cluster tool sputtering systems at 5×10^{-3} torr and on ion implanter end stations and beam lines.

Ion milling processes using inert gases such as argon, as well as reactive gases, are pumped with cryopumps. Industrial equipment such as electron beam welders and vacuum brazing and heat-treating furnaces rely on the oil-free environment of cryopumped chambers for high-quality metal processing. NASA and the aerospace companies use large cryopumps of 36- to 48-inch diameter for evacuating 100-foot-long space simulation chambers to 10^{-8} torr, although these larger pumps may use multiple refrigerators and LN_2 for part of the refrigeration. Small cryopumps of 100 mm diameter are used on load locks of vacuum systems and on analytical instruments such as secondary ion mass spectrometer (SIMS) systems.

The most popular size of cryogenic vacuum pump is of 200-mm (8-inch) diameter, and most of these are used on commercial semiconductor production systems. Typical characteristics for a 200-mm cryopump are given here:

8-inch (200-mm) Cryopump Characteristics

Water vapor speed	4000–4500 L/s
Nitrogen speed	1500 L/s
Argon speed	1200 L/s
Hydrogen speed	2200–2500 L/s
Argon capacity	1000 standard liters
Hydrogen capacity	12–18 standard liters
Argon throughput	700 std. cm^3/minute
Crossover rating	150 torr-liters
Cooldown time	90–100 minutes
Cold-to-cold full regeneration time	2.5–3 hours

2.5.2 CRYOPUMP BASICS

2.5.2.1 Flow Regimes

Cryopumps can operate over a wide range of pressures, typically from about 5×10^{-3} torr on the high end to 1×10^{-11} torr on the low end. Laboratory pumps have reached even lower pressures. This range of pressures is generally in the molecular flow regime for systems of reasonable dimensions. In molecular flow, the trajectories of the molecules are straight lines with a mean free path very long compared to system dimensions. The molecular density is low enough to ignore all molecule–molecule interactions and scattering. This simplifies the design of the elements of the cryopump. In addition, heat transfer by gas conduction can be ignored below 10^{-3} torr. At sputtering pressures in the mid-10^{-3} torr range, however, both gas scattering and conduction may be significant [3].

An aspect of molecular flow that is important for all high-vacuum pumps is that the pump does not draw or "suck" molecules toward the entrance. The pump must wait for the arrival of a molecule in its quasi-random motion, and then trap it. Because the cryopump traps the molecule by adsorption and condensation onto a cold surface, it can be considered a piece of cryogenic "flypaper." A molecule lands on the cold surface and never takes off again. Other molecules land on the surface, to form progressively thicker layers of condensate.

2.5.2.2 Vapor Pressure

Cryopumps are omnivorous: They condense all gases at very high speeds. The ability of a surface to trap a particular gas depends on the vapor pressure characteristic of the gas. Nearly all gases follow the basic Clausius-Clapeyron relation [4] of the vapor pressure P and the temperature T in Kelvins.

$$\log P = -A(1/T) + \text{constant}$$

The slope of the line, $-A$, is the heat of vaporization of the gas. An example of the Clausius-Clapeyron plot is shown in Figure 5. Many readers are more familiar

Fig. 5.

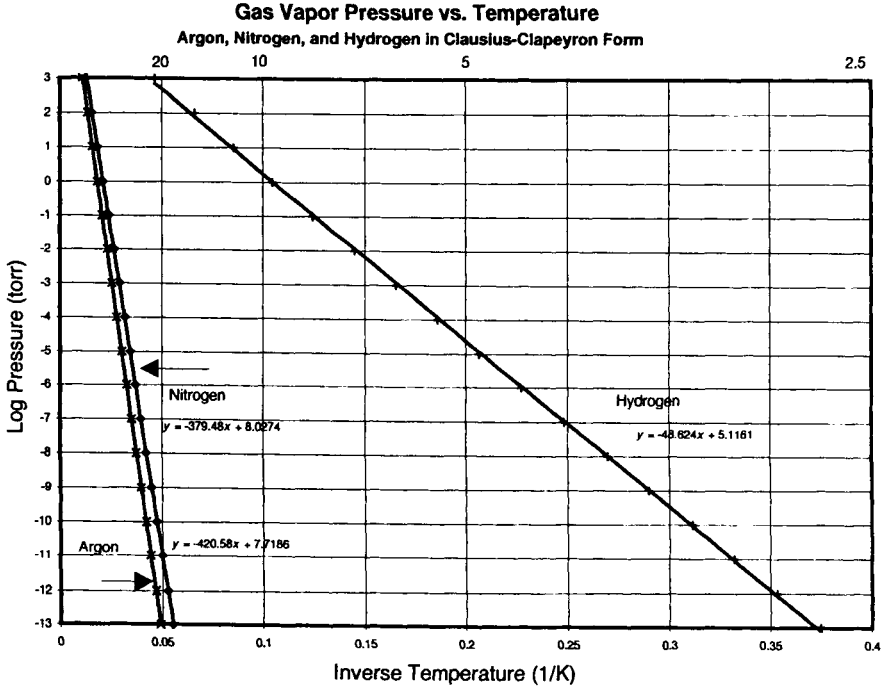
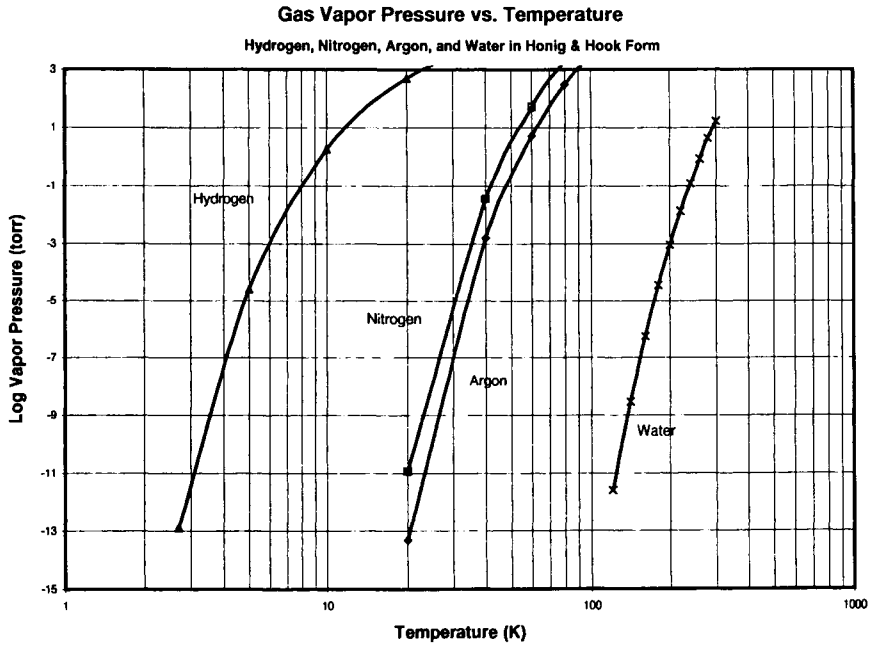


Fig. 6.



with an alternative presentation of vapor pressure data by Honig and Hook [5]. The data from Figure 5 is replotted in Honig and Hook form in Figure 6. For a surface at a given temperature in a closed volume, equilibrium will be attained at the equilibrium vapor pressures for the various gases in the volume. The resulting total pressure will be the sum of the equilibrium partial pressures of the individual gas components. For system pressures more than 100 times the equilibrium vapor pressure of the cryopump array, full speed is obtained. As the system pressure drops below 10 times the equilibrium pressure, the speed begins to drop off. Zero speed occurs when the base pressure gradually approaches the equilibrium vapor pressure of the array surface, or the ultimate pressure for the gas in the cryopump. In most cases, these ultimate pressures are well below 10^{-11} torr for condensable gases.

Normal dry room air is about 78% nitrogen, 21% oxygen, 1% argon, and other gases down to about 7 ppm of helium. Depending on the relative humidity, water vapor may make up 1 or 2% of the air. In a vacuum system, nitrogen, oxygen, and other gases are rapidly removed by the initial pumping process, leaving water vapor as the major constituent of the vacuum atmosphere. The water vapor is polar-bonded to the surfaces in 50–100 monolayers, while other gases are easily stripped

away. Water may make up 98% of the gas load of a vacuum system at 10^{-7} torr. Consequently, high pumping speeds for water vapor are crucial in all vacuum systems.

2.5.2.3 Condensation

Water is relatively easy to pump onto cryogenic surfaces. At 130 K, an equilibrium vapor pressure of 10^{-10} torr is obtained, ensuring nearly 100% pumping efficiency to pressures below 10^{-8} torr. A surface below 113 K will provide greater than 99% efficiency to below 10^{-11} torr. Consequently, operation of a cryopump surface in the 65–90 K range is 100% efficient for pumping water and high-molecular-weight gases and vapors in any practical vacuum system. Temperatures below 20 K are necessary for condensing nitrogen, oxygen, argon, and most other condensable gases. At these temperatures, the condensates are dense solids, like ice. It should be noted that it is not practical to condense hydrogen, helium, and neon, even at temperatures in the 6–20 K range. These three gases must be pumped by special adsorption surfaces within the cryopump.

For all other gases, however, the condensation surfaces in the 10–20 K range provide 100% efficiency pumping at constant speed over the entire pressure range from below 10^{-10} torr to above 10^{-4} torr. As the pressure rises into the low 10^{-3} torr range, the pumping speeds increase by about 20–40% due to the onset of transitional flow. In most applications, the speed of the cryopump can be treated as essentially constant at all pressures.

The process of condensation requires removal of heat from the gas molecules through contact with the cold surfaces. Remarkably, the quantity of heat removed from gases near room temperature is quite low. For nitrogen, the conversion of gas at 300 K to a solid at 20 K is a complex though instantaneous process.

Nitrogen Condensation Budget

Cooling Process	Enthalpy (calories/gram-mole)
Gas cooled 300 K to 77 K	1600
Liquefaction at 77 K	1300
Cool liquid 77 to 63 K	189
Transition liquid to solid I	172
Cool solid I 63 K to 36 K	267
Transition solid I to solid II	54
Cool solid II to 20 K	123
Total	3700

This translates to about 1 watt of cooling required for each 100 std. cm^3/minute of nitrogen gas flow into the system, about the flow in a sputtering system at 2×10^{-3} torr. For the same pump operating in the 10^{-6} range or below, the total condensation load is less than 1 milliwatt. It is for this reason that even very large cryopumps can be constructed with a relatively small refrigeration capacity at 20 K and below. For comparison, the mechanical work involved in the transport of 100 std. cm^3/minute (0.2 torr-L/sec) of gas in a vacuum system requires only about 30 milliwatts of power.

It is frequently advisable to control the operating temperature of the inlet array and thermal radiation shield at a preset value. Generally, the temperature is set in the range of 65 to 90 K to prevent the partial adsorption of Ar, N_2, CO, and other gases on the first-stage array parts. For example, the first-stage temperature can drop to 40–45 K in a cryopump with very low thermal loads. If the pump is then used for sputtering with Ar at 10^{-3} torr, some argon is adsorbed in a very thin layer on inlet array surfaces. When the gas flow is shut off at the end of the process, the chamber pressure may take a long time to recover to the 10^{-6} to 10^{-7} range as the lightly adsorbed gas is evolved from the inlet array. The presence of a thick water ice deposit on the inlet array can make the problem worse due to the larger effective surface area of the microcrystalline ice deposits. Raising the first-stage array temperature to the 65–90 K range reduces the adsorption of Ar and speeds the recovery to base pressures.

In principle, a first-stage array temperature in the range of 90–105 K might be desirable for some processes requiring repetitive cycling over a large dynamic pressure range. However, since the first and second stages of the refrigerator are thermodynamically linked, it is not normally possible to drive up the first-stage temperature to these high values without also dragging the second-stage temperatures well above the optimum 10–14 K range.

2.5.2.4 Adsorption

The three light gases H_2, He, and Ne cannot be pumped by condensation and must be adsorbed. Activated carbon or charcoal from coconut shells, wood, or other organic materials is used for this purpose. Treatment of the charcoal with steam at high temperatures produces huge effective areas of 1000 m^2/gram or more. The carbon then has an interconnected network of minute channels or fissures with tightly controlled pore size in the range of 10–30 angstroms. A gas molecule landing on the outer surface of the carbon loses kinetic energy as it cools to 10–12 K. Such loosely held molecules then diffuse rapidly into the inner volume of the charcoal. At these temperatures, hydrogen molecules can be retained at very low equilibrium vapor pressures as up to a monolayer of gas is ad-

sorbed onto the carbon. Although this may not seem impressive, at approximately 10^{15} atoms/cm^2 a gram of charcoal (about a teaspoon) can hold 200 std. cm^3 of hydrogen. A handful of charcoal can hold 10–20 liters. A layer of charcoal granules or pellets is attached to copper sheets with vacuum-compatible epoxy or other adhesives.

The remarkable ability of cold, activated carbon to adsorb gas is demonstrated by the reduction in vapor pressure. A flat copper surface at 12 K would hold only a small fraction of a monolayer of hydrogen while coming to an equilibrium pressure of about 20 torr. By adding 50–100 grams of charcoal to the copper surfaces, many liters of gas can be adsorbed before reaching an equilibrium pressure of about 10^{-10} to 10^{-8} torr. This is a reduction in pressure by 12 orders of magnitude with a simple technique.

Carbons are used in preference to zeolites because carbons are hydrophobic. That is, when the cryopump is eventually regenerated, any adsorbed water vapor can be removed from the charcoal by heating at 30–40°C, instead of the 200–250°C required of zeolites. The activated carbons used in cryopumps have been processed at very high temperatures to obtain nearly pure carbon with few impurities. Since both the carbon and the adhesive are at about 12 K whenever the high-vacuum valve is open to the rest of the vacuum system, they are always "sinks" for gases and never sources of outgassing.

The operating temperature of the charcoal affects how much gas can be retained at a given equilibrium pressure. At 10–15 K, a gram of charcoal will hold about 200 std. cm^3 of hydrogen at 5×10^{-6} torr or about 100 std. cm^3 at 5×10^{-8} torr. These are the normal conditions of use of a cryopump at maximum rated capacity. Temperatures below 10 K do not significantly improve effective capacity, although lower ultimate pressures may be obtained in clean, baked UHV systems with light hydrogen loads. At high hydrogen gas loads that produce partial pressures of 10^{-5} torr and above, very low carbon temperatures are harmful. The hydrogen may not be able to diffuse to the inside of the carbon granules rapidly enough to maintain low pressures. Instead, operation of the charcoal in the 12–14 K range is optimum to balance good capacity with maintainable, constant pumping speed [6]. Above 25 K, though, the capacity of carbon to adsorb hydrogen falls rapidly. It has the ability to adsorb (pump) many other gases, such as nitrogen or argon, at temperatures well above 90 K.

Helium is more difficult to adsorb than hydrogen. A pump that can adsorb up to 20 liters of hydrogen while holding pressures below 5×10^{-6} torr, can adsorb only 5 std. cm^3 of helium at the same maximum pressure level. It is fortunate that helium is only a minor constituent in our atmosphere. The very low capacity for helium precludes the use of cryopumps for any process that bleeds helium into the vacuum chamber. It also requires that helium remain as only a trace impurity in any other process gas, a condition not always obtained. In particular, commer-

cial bottled gas suppliers and on-site air separation facilities may not count helium or neon as impurities when they certify argon and nitrogen gas purity. The presence of 10–100 ppm of helium in a process gas stream can quickly cause base pressure problems in cryopumped systems.

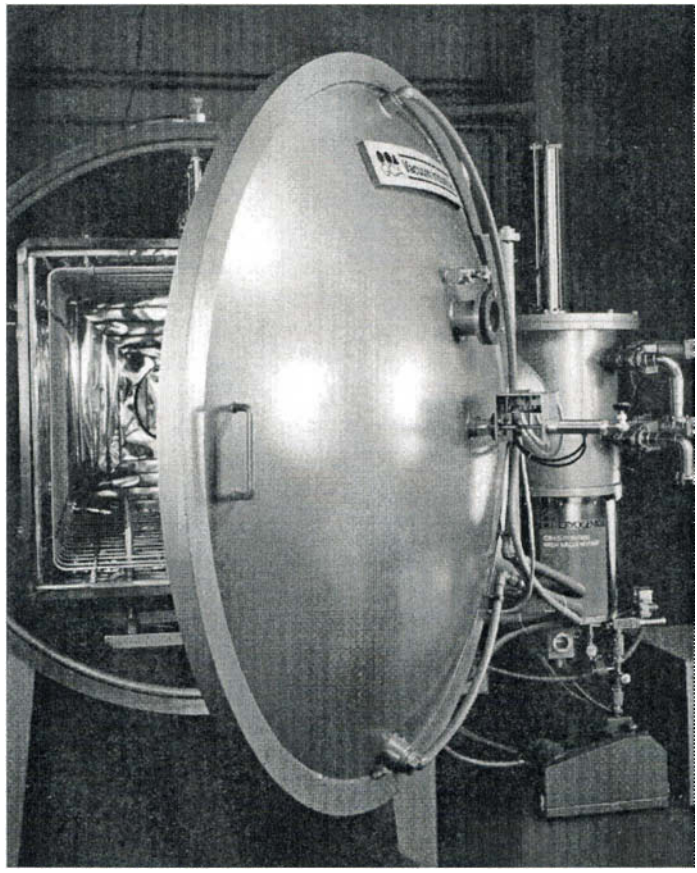
2.5.2.5 Thermal Considerations

Thermal radiation loads also play a role in the design and operation of cryogenic high-vacuum pumps. Pumps are usually mounted on a port of the vacuum chamber with a thin gate valve at the opening. The vacuum chamber acts as a large optical integrating cavity or "black body" at the wall temperature if there are no other heat loads in the chamber. With a room temperature (295 K) wall, the thermal load from infrared radiation on the inlet of the pump is approximately 0.05 watt/cm^2, or about 16 watts across a standard 8-inch diameter pump. With hotter chamber surfaces, the thermal load goes up as T^4 (in Kelvins). This limits most cryopumps to operating on chambers with wall temperatures in the range of 50–100°C.

Most of the thermal load is absorbed by the first stage of the refrigerator as the radiation strikes the inlet array of the pump. The inlet array of the pump is typically nickel plated to reduce the absorbed thermal load on the pump during the cooldown process when the pump is empty of all gases. Once the pump has operated for a few hours or days, a thin layer of infrared-absorbing water ice forms on the inlet array. This effect raises the thermal load on the inlet array. Some thermal radiation passes through or around the inlet array to fall on a black thermal radiation shield, which is also cooled to about 65–80 K by the first stage of the refrigerator. Because the refrigerator typically has three times the cooling capacity on the 65 K first stage as on the 12 K second stage, the inlet array also serves as a shield for the second-stage array to prevent thermal overloading of the low-temperature surfaces.

Thermal sources in the vacuum chamber, such as heat lamps and deposition sources, add other heat loads to the cryopump. Stainless steel vacuum chamber walls are 90% reflective in the infrared, whereas aluminum surfaces may approach 98% reflectivity at long optical wavelengths. This allows many bounces of the radiation around the chamber with little attenuation. Infrared radiation, unseen by the human eye, winds up in the "black hole" of the cryopump. Consequently, attention to thermal shielding of the heat sources in the chamber or at the pumping port may be necessary. That this shielding process can be effective is demonstrated by the use of cryopumps on vacuum furnaces, such as the one shown in Figure 7, having hot zones operating at 2000°C and above.

Fig. 7.



A 12-inch (320-mm) diameter cryopump with 2000 liters of argon ice and some water vapor. (Photo of vacuum furnace with cryopump courtesy of CTI-Cryogenics.)

2.5.2.6 Crossover

Another heat load on the cryopump is the transient load from the initial opening of the gate valve to the vacuum chamber. Typically, a vacuum chamber is rough-pumped to 0.1 to 1 torr before opening the valve to the high-vacuum pump. The sudden inrush of air and water vapor, as measured by the product of chamber pressure and volume, produces an impulsive heat load on the cooled surfaces of the pump as the gas is condensed. The pressure in the chamber drops to 10^{-5} torr

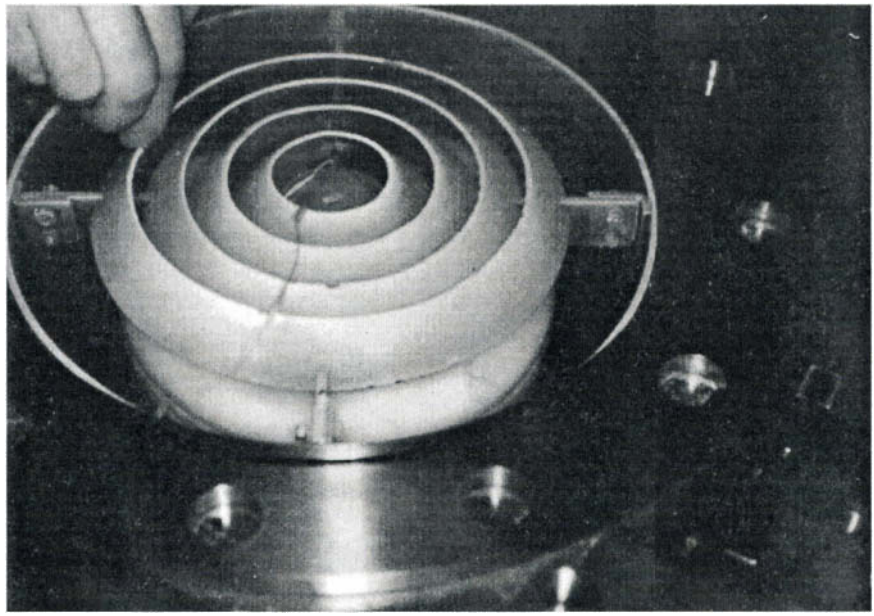
or less in a few seconds. The joules of heat extracted from the gas during cooling momentarily exceed the ability of the refrigerator to remove heat and maintain temperature. Now, the thermal mass of the cooled parts determines how high the temperature will rise. The *crossover rating* of the pump, usually given in torr-liters, is that product of chamber pressure and volume that can be accommodated without raising the second-stage array parts above 20 K. This ensures that any adsorbed hydrogen will not be released back into the system, causing the pressure to go up instead of down. Interestingly, continued use of the cryopump improves the crossover rating because the condensed gas loads have significant thermal capacity of their own. For example, 300 standard liters of argon gas condensed into about 375 cm^3 of argon ice weighing about 535 grams. This ice at 12 K has a heat capacity approaching that of the mass of copper and stainless steel of the cryopump parts at 12 K, effectively doubling the crossover rating.

2.5.2.7 Capacity

The capacity of a cryopump to store condensed gases depends on the dimensions of the pump. Water vapor — and other Type I gases that have low vapor pressures at 65 K — condense on the top surfaces of the inlet array and on the inside of the thermal radiation shield. For a 200-mm (8-inch) cryopump, nearly a liter of water can be stored as ice, spread out over these surfaces in layers up to a few centimeters thick. The ultimate limit on water capacity is that the ice will eventually choke off the ability of other gases to reach the interior of the pump, thus reducing their pumping speed. In most applications, a few tens of cubic centimeters of water will be collected over a few weeks of operation. Under UHV conditions, the pump will not collect a thimble full of water or other gases in a year! A cryopump with only a little condensed water is shown in Figure 8.

Argon, nitrogen, oxygen, and other Type II gases condensable at 20 K, collect on the second-stage surfaces at 10–20 K. As they form ice layers of several centimeters thickness, several things begin to happen. The ice grows upward until it reaches the 65 K inlet array. Since the Type II gases will not condense at 65 K, the ice continues to grow up through the inlet array without actually touching the warmer surfaces. The Type II ice ball also begins to block access of the Type III gases, which do not condense at 20 K (hydrogen, helium, neon) to the carbon-coated surfaces. Furthermore, the Type II ice is no longer thermally shielded by the inlet, so the thermal loads increase on the second stage of the refrigerator. Between the increased thermal load and the increasing thermal gradient through the ice, the equilibrium pressure obtainable begins to rise. This sets the practical limit on capacity for Type II gases at about 1000 standard liters of gas for a 200-mm cryopump before the equilibrium pressure rises into the 10^{-5} torr range. For systems that must cycle rapidly and frequently from above 10^{-3} for sputtering to be-

Fig. 8.



A 12-inch cryopump showing argon ice thickness of 2–3 cm as regeneration continues and ice begins to melt. (Photo of cryopump inlet array with ice courtesy of CTI-Cryogenics.)

low 10^{-7} torr for chamber cleanup, reduced effective capacities may be observed as the time to recover to base pressure increases.

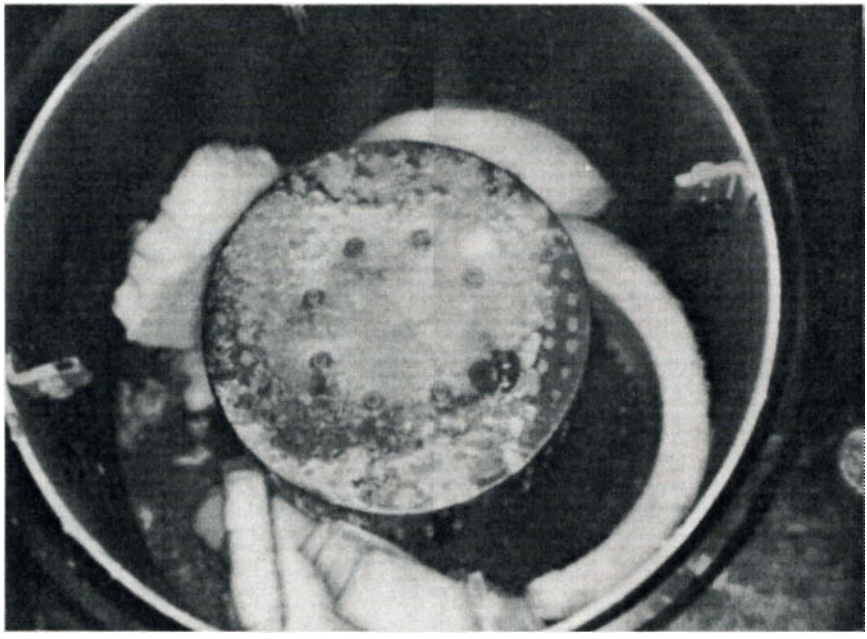
In practical terms, a cryopump can be used for a long time before reaching capacity. For static, high-vacuum pumping at 10^{-6} torr, an 8-inch pump could run for one to three years before requiring regeneration. On load-lock or batch applications without additional process gas loads, 2000–5000 pumpdown cycles may be obtained between regenerations. Sputtering systems have typical argon gas loads of 50–150 std. cm^3/minute coupled with requirements for rapid recovery, limiting time between regenerations to 1–2 weeks. Similarly, the heavy hydrogen gas loads of ion implanters require regeneration at intervals of 5–15 days in full-production situations.

2.5.2.8 Regeneration

As the effective capacity is reached, it is necessary to regenerate the cryopump to restore full pumping speed and the ability to recover rapidly to base pressure. Because three types of adsorption processes have been used to collect the gases, a

number of effects take place during regeneration. The simplest way to regenerate a pump is simply to close the high-vacuum valve and turn off the refrigerator. Since the pump will initially be under high vacuum, with all the surfaces cold, the arrays and condensed gases will warm slowly as heat is conducted through the cylinder of the refrigerator to the second-stage array. As the second-stage array temperature (and the temperature of the charcoal attached to the array) climbs above 25 K, hydrogen and other Type III gases are released into the pump vessel, raising the pressure. As the pump pressure rises above 10^{-3} torr, the thermal conductivity of the gas increases, providing more heat load on the arrays. Above 30 K, the equilibrium vapor pressures of the Type II gases condensed on the second stage begin to rise, and nitrogen and argon add to the rising pressure in the pump. For a few minutes, the still-cold charcoal adsorbs the Type II gases, suppressing the pressure rise. Eventually, the second stage warms sufficiently to liquefy the Type II gases, which then flow to the warmed thermal radiation shield and outer vacuum vessel surfaces. A cryopump in the middle of the regeneration process is shown in Figure 9. In some pump orientations, the liquid cryogens may flow to the gate valve or other parts of the chamber. The outer vessel of the pump becomes very cold, condensing water vapor from the room air onto the outside surface to liquid droplets or frost.

Fig. 9.



Vapor pressure of water versus temperature. (Photo of interior of cryopump during regeneration with melting ice courtesy of CTI-Cryogenics.)

Now the increased vaporization rate overwhelms the ability of the charcoal to adsorb any more gas, and the pressure rises to above 1 atmosphere. A spring-loaded pressure relief valve opens to vent the cold gases and maintain internal pump pressure at a few pounds/in^2 above atmosphere. If the pressure relief valve were not present, 1000 liters of condensed argon could raise the pressure in a 10-liter vacuum vessel volume to 100 atmospheres, causing damage to the pump, high-vacuum valve, or the vacuum system itself. For this reason, it is important to allow free flow from the pressure relief valve for gases released during regeneration.

Many processes use or generate flammable, toxic, or otherwise hazardous gases. Exhausted gases from the cryopump should generally be vented to approved stacks or treatment systems. Since 1000 liters of gas may be vaporized in a few minutes during pump regeneration, the exhaust system must be sized for peak flow rates. Where the gases of several processes may be combined in a vent system, attention must be paid to possible adverse chemical reactions and flammability. For example, during regeneration pumps used on ion implanters may release tens of liters of hydrogen as well as toxic by-products of boron, arsenic, or phosphorous. These gases should not be released into the room where pumps are located.

Eventually the inlet array and thermal radiation shield of the pump rise above 0°C (273 K), and water frozen on the arrays begins to melt. The water flows to the low spots and begins to evaporate. The charcoal on the second-stage array adsorbs a great deal of water vapor and even liquid droplets that reach it, although now it may be approaching room temperature. The inside surfaces of the chilled vacuum vessel also condense water vapor to droplets or ice. The pressure in the pump drops to 1 atmosphere as the final Type II gases escape and the pressure relief valve closes automatically. Without any purge gas added, all the pump parts eventually return to room temperature. If there is considerable water in the pump, the charcoal may take on half its own weight in water vapor at 100% relative humidity. Before the pump can be restarted, the residual gases and any water in the pump must be completely removed. This can be done by warming the pump, purging with dry nitrogen gas, and rough-pumping the vacuum vessel.

2.5.3 ADVANCED CONTROL SYSTEMS

2.5.3.1 Temperature Indicators

Although stable operation of cryopumps is obtained over a very wide range of array operating temperatures, knowledge of the actual array temperatures is useful. Older pumps used a hydrogen vapor pressure (H_2VP) gauge system with a gas-

filled bulb mounted on the low-temperature stage. Since the normal boiling point of hydrogen is 20.2 K, an indication of 14.7 psig or 760 torr (1 atmosphere) on a mechanical pressure gauge showed the array had reached 20 K. As the temperature dropped below 12 K, the gauge reading went to less than 0.2 psig or 10 torr. Although simple and reliable, the H_2VP gauge is less useful when other temperature ranges and electrical outputs are desirable.

Silicon diodes are used as temperature sensors in most cryopumps. When excited by a few microamps of direct current, the junction voltage drop provides an accurate, repeatable indication of temperature. High sensitivity, stability over time, and low cost make them particularly attractive for control circuits. Although their response is nearly linear over the 25–330 K range, the lower-temperature response is very nonlinear. Digital electronic circuits make it relatively simple to convert diode voltage to very accurate temperatures.

2.5.3.2 Pressure Sensors

A thermocouple vacuum gauge can be attached to the vessel of the cryopump to measure pressure over the range of about 1 torr to 10^{-3} torr. Although pressure sensors tend to have absolute accuracy no better than ± 20–30 percent at the end of their ranges, this is sufficient for good operation of the pump. Since they actually measure the thermal conductivity of the gas that is present, they do measure the property that is most important to the regeneration of the cryopump. Given their mechanical ruggedness and low excitation power, thermocouple gauge tubes have proved very stable and reliable. It is normally sufficient to re-zero the gauge occasionally when the pump is at operating temperature and the high-vacuum valve is closed, a condition certain to produce a vacuum well below the sensor's lowest limit of detectability.

Normally, a gauge tube design is selected to have a filament temperature below the ignition point of hydrogen to prevent explosion in case the pump is started into regeneration with the gauge turned on. For similar safety reasons, it is strongly recommended that hot-filament ion gauges not be placed on the pump side of the high-vacuum valve.

2.5.3.3 Electrical Heaters

Attached to the two heat stations of the refrigerator, electrical heaters may be used to obtain two important functions: maintaining exact temperatures during operation and quickly warming the interior parts of the pump during regeneration. Since in most cases it is desirable to hold the first-stage temperatures near 65 K, a closed-loop control circuit linked to the diode temperature sensor is used to add a few watts of power to the first-stage heater. Argon hangup problems are

thus prevented. It is not normally necessary to drive the second-stage temperature up, since other thermal radiation loads will provide the watt or so needed to put the refrigerator into the 10–14 K range.

During regeneration, the heaters can be driven at full power to provide 100–200 watts or more to bring the arrays to well above room temperature in a few minutes. Again, closed-loop control is important so that the refrigerator and array parts are not overheated. Damage to the refrigerator can occur if the pump is heated above 80–100°C for prolonged periods. The heaters provide the joules necessary to vaporize the liquid argon or nitrogen frozen on the second stage and melt the water ice on the first stage. The heated arrays also warm the flow of nitrogen purge gas coming into the pump during regeneration. Heat is transferred to the outer vacuum vessel by conduction and convection. The control circuit slowly reduces the power delivered to the heaters as the various parts of the pump come up to temperature, typically about 40–60°C.

Because the arrays, temperature sensors, and heaters are thermally well coupled to the refrigerator heat stations, the heater sheaths do not reach very high temperatures. That is, they remain very close to the 40–60°C maximum temperature setting of the refrigerator process. The sheaths are hermetically sealed to prevent the entry of water or gases, and also to avoid exposure of the heater filament to the pump vacuum space. These heater characteristics are important to ensure that the heater itself cannot become an ignition source for flammable gases during regeneration.

2.5.3.4 Automated Regeneration

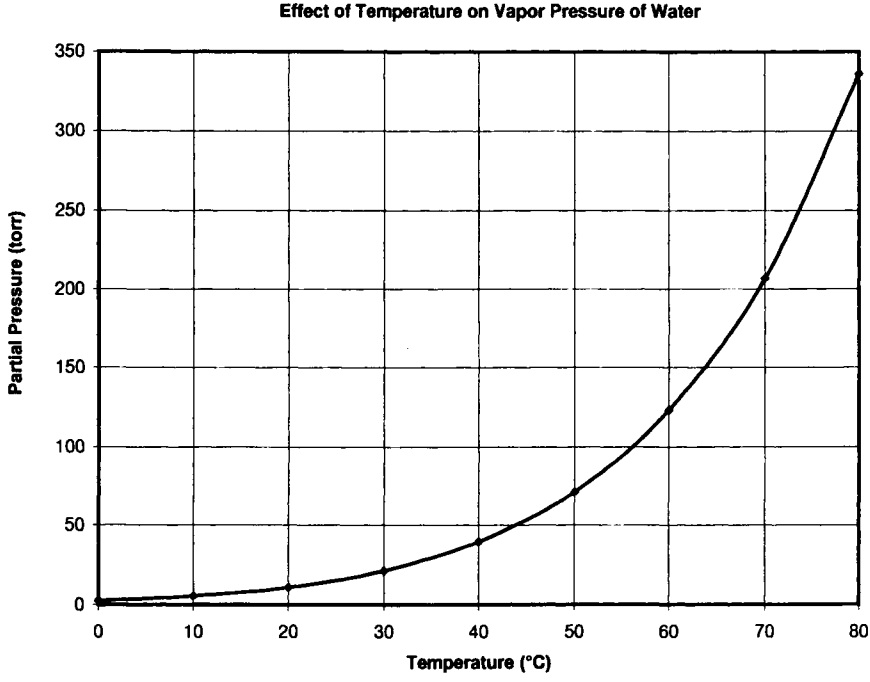
The latest generation of cryopumps use automated regeneration processes incorporating internal heaters for the arrays, pressure and temperature monitoring and control, purge gas, and a controlled roughing valve. The purpose of automated regeneration is to complete the process in a minimum amount of time while ensuring that all the residual water and gases are removed from pump surfaces and activated carbon. By roughing to a predetermined base pressure and rate of rise, all water can be removed and a repeatable level of cleanliness of the activated carbon can be obtained.

At the beginning of the automated regeneration cycle, a flow of nitrogen purge gas is started to quickly bring the vacuum vessel to atmospheric pressure for enhanced heat transfer from the room. In addition, the 20- to 60-liter/minute purge gas flow dilutes hydrogen during the initial stages of regeneration and carries it from the pump. The electrical heaters attached to the refrigerator heat stations, or to the arrays themselves, quickly melt the frozen gases and assist in rapid vaporization. The heaters bring the arrays — and the charcoal — to a temperature near 40°C. The heated arrays warm the incoming purge gas to speed the vaporization of water and the warming of the vacuum vessel walls. Purge gas flow must con-

tinue until all the liquid water is evaporated, the array and pump vessel surfaces are dry and warm, and most of the water vapor has been dried from the activated carbon. When only a little water is present, as for pumps on vacuum process chambers of load-locked systems, 20 minutes may provide sufficient purging. However, for cryopumps on chambers regularly exposed to atmosphere and humidity, such as the load locks themselves, many tens of cubic centimeters of water may be present in the pump, requiring an hour or more of purge time. For best performance, purge gas lines should be made of metal tubing and in-line dryers should be used to avoid water vapor contamination of the purge gas.

Rough-pumping of the vessel begins when all surfaces are warm and dry. A base pressure of .05 torr is typically selected as a requirement. Roughing removes gases and water adsorbed into the charcoal at atmospheric pressure and strips most of the water vapor from the metal surfaces. The heaters for the arrays are kept on during the rough process to prevent evaporative cooling of the activated carbon. Water will be evolved at a much faster rate at higher temperatures. For example, the vapor pressure of water is nearly four times higher at 40°C than it is at 20°C [7]. Figure 10 shows the vapor pressure of water over the 0 to 80°C range.

Fig. 10.



A cryopumped large-vacuum furnace with thermally shielded, high-temperature hot zone.

In some regeneration routines, it is possible to perform simultaneous heating, purging, and roughing to enhance removal of water and heavy organic deposits in the pump, such as photoresist by-products in ion implanters. By keeping the pressure above 10 torr, rapid volatilization of the water is possible without the possibility of freezing the water to ice by evaporative cooling. The simultaneous flow of purge gas carries the water through the rough pump, and prevents recondensation in dry rough pumps. If the pressure in the cryopump is reduced below about 10 torr when liquid water is present, ice quickly forms and does not sublime. A false indication of good base pressure might then be obtained with a large-enough rough pump, leading to an incomplete regeneration.

On reaching base pressure, the rough valve closes and the rate of rise of pressure is monitored. If all liquid water is gone and the charcoal is sufficiently clean, the rate of rise will be less than about 10 microns/minute. If the target rate is not met, the rough valve opens again. If the number of rough cycles exceeds a preset count, an additional purge cycle is performed to evaporate any remaining water. The rough cycle then begins again.

Once both base pressure and rate-of-rise values are met, the heaters are turned off, the refrigerator is turned on, and the cooling-down process of the cryopump starts. At the internal pump pressure of 50–100 microns, most of the residual gas is water vapor. Once the refrigerator reaches about 150 K, the metal surfaces become very efficient for pumping water vapor and the pressure starts to drop. In addition, the charcoal also begins to efficiently adsorb gases, removing any residual air or nitrogen purge gas. As the pressure again falls below 1 micron (10^{-3} torr), an insulating vacuum is achieved, removing thermal conduction loads from the pump arrays and speeding the cooldown process. When the second-stage array reaches about 17 K and the first stage is below about 130 K, the pump is ready to use again. In most cases, the vacuum process chamber has been vented to air during the cryopump regeneration process, so high-vacuum pumping can begin even as the pump itself continues to cool to its ultimate low temperatures.

For ultra-high-vacuum chambers that have not been vented to atmosphere during cryopump regeneration, it may be desirable to wait to reopen the high-vacuum valve until some time after the pump reaches its lowest temperatures. This is to be sure that any residual water vapor has been released from the cryopump vessel and high-vacuum valve surfaces and is condensed on the cold surfaces of the pump.

2.5.3.5 Partial Regeneration

In many applications, such as sputter chambers on semiconductor production systems, rapid regeneration of a cryopump is economically appealing. Fortunately these systems have load locks that prevent the ingress of water into the process chambers and cryopumps located on them. In that case, regeneration of the 12 K

surfaces and condensed loads of argon or nitrogen process gases and removal of adsorbed Type III gases in the charcoal is all that is required. Since there is little water present, it is not necessary to raise the first-stage arrays from 65 K back to room temperature to melt ice.

For partial regeneration, the refrigerator is kept running to maintain the first stage near 65 K. The heater on the second stage is turned on to raise the second-stage array temperature to about 180 K. At the same time, a short burst of nitrogen purge gas is added to bring up the pressure in the pump. As the second-stage arrays rise above 60–80 K, frozen nitrogen and argon melt and fall to the thermal radiation shield (still near 65 K) or the vacuum vessel wall. In a few minutes, the liquid nitrogen or argon vaporizes. The temperature of the inlet array and thermal radiation shield is kept low enough to prevent sublimation of water vapor from these surfaces into the activated carbon. The vapors help to keep the thermal radiation shield cold in spite of heat conduction through the vapors from the warmer vacuum vessel. Vapor exits the pressure relief valve. Activated carbon quickly fills with large amounts of nitrogen and argon. After a few minutes, the rough valve is opened. If liquids remain, the pressure remains high. If all the liquids are gone, the pressure will quickly drop below 1 torr. If a rapid pressure drop is not detected when the rough valve opens, the valve is closed again and a short burst of purge gas is injected to raise the pressure again and encourage the vaporization of the liquid.

Once the pressure drop verifies the absence of liquids, rough pumping is continued to base pressure, while the activated carbon array is maintained at about 180 K. Release of nitrogen and argon from the activated carbon at this temperature and a base pressure of 50 microns is sufficient to completely restore the hydrogen capacity to its initial value. The second-stage heater is turned off, and the refrigerator returns the second-stage array to 10–20 K for operation. In most cases, the fast regeneration process can be completed in about one hour or less. When properly completed, 10 or more sequential fast regenerations may be done without any measurable loss in hydrogen speed or capacity. If higher water loads are present in the system, then perhaps only two to four fast regenerations may be done between full regenerations, since the increasing water burden adds to the thermal load on the pump. A consequence of using fast regeneration is that purge times must be extended during full regeneration, because 4–10 times as much water may be present in the pump as compared to when full regenerations are performed at every cycle.

Variations of the fast regeneration process may be appropriate for specific applications, such as ion implantation. The high amounts of hydrogen evolved by the implant process [8] compared to relatively low amounts of nitrogen or other Type II gases, may require additional purge–rough cycles during fast regeneration to remove all the hydrogen. Full regeneration may be required more frequently than in sputtering to eliminate the larger amounts of water and photoresist by-products.

2.5.4

CRYOPUMP PROCESS APPLICATIONS

2.5.4.1 Chamber Evacuation

Even the most basic vacuum chambers require attention to the high-vacuum pump if they are to quickly reach their intended operating pressures. When starting from atmospheric pressure, a mechanical pump is used to produce initial rough vacuum. The point of crossover to the high-vacuum pump depends on the type of rough pump used, the speed of the pump, and the crossover rating of the cryopump. Consider a 20-inch-diameter, 30-inch-high bell jar (5.5 ft^3 or 155 liters) with an 8-inch (200-mm) cryopump. The 150 torr-liter crossover rating of the cryopump indicates that the high-vacuum valve can be opened when the 155-liter chamber gets below 1 torr. Many chambers of this size use rough pumps of about 17–35 ft^3/minute (~30–60 m^3/hour) displacement, which equates to a speed of only 8–16 liters/second. Clearly, it is advantageous to switch to the cryopump's 1200–4000 liter/second speeds as soon as possible. In addition, oil-sealed rough pumps tend to backstream oil at pressures below 200 microns (0.2 torr). Dry pumps avoid the backstreaming problem, but still suffer from relatively low speeds. At crossover, the pressure will quickly drop to the point where the outgassing of the chamber determines the pressure. For chambers that have been exposed to room air, this means that the chamber will probably reach a pressure in the 10^{-4} to 10^{-5} torr range in the first minute.

After crossover, the pressure will slowly decrease in the absence of leaks. Most outgassing phenomena, especially water vapor, follow a $\frac{1}{t}$ decrease in outgassing rate [9]. That means that it takes ten times as long to get to one-tenth the pressure. For example, if it takes 1 minute from crossover to reach 10^{-4} torr and 10 minutes to reach 10^{-5} torr, it will take 100 minutes to reach 10^{-6} torr and 1000 minutes (17 hours) to get to 10^{-7} torr. If there are system leaks or large amounts of O-rings to allow permeation, low base pressures may never be obtained.

The value of a mass spectrometer-type residual gas analyzer on all vacuum systems cannot be overemphasized. It is the only way to determine what type of gas is responsible for determining the pressure [10]. The good news for cryopump users is this: If the pump is cold, it is probably pumping at its full rated speed. The causes of poor pressure are generally outgassing and leaks in the chamber, not a decrease in the speed of the cryopump. The chamber pressure P is determined by the relationship of outgassing rate Q and pumping speed for a particular gas S. That is,

$$P = Q/S$$

Sometimes Q can be estimated by assuming the pumping speed is constant, accounting for all the conductance losses between the pump and the chamber, and

plotting the decrease in pressure over time. Unfortunately, when water vapor outgassing predominates, as it usually does, the outgassing rate *cannot* be determined by closing the high-vacuum valve and noting the pressure rise with time. This is because the water outgassing rate will immediately decrease as the chamber pressure rises, resulting in a gross underestimate of the outgassing rate.

Leak checking of cryopumped systems can sometimes be difficult. When a helium leak detector is connected to the chamber with the gate valve open to the cryopump, only a very low sensitivity is obtained. That is because the net pumping speed of the leak detector may be only 1 to 2 liter/second, whereas the helium pumping speed of the cryopump is 1500–1800 liter/second. This results in a 1000:1 reduction in leak sensitivity. Closing the high-vacuum valve between the cryopump and the chamber improves the sensitivity if the leak detector itself can maintain a low-enough pressure in the vacuum system. Because the helium capacity of the cryopump is only a few standard cm^3, the cryopump should be regenerated after any helium leak checking.

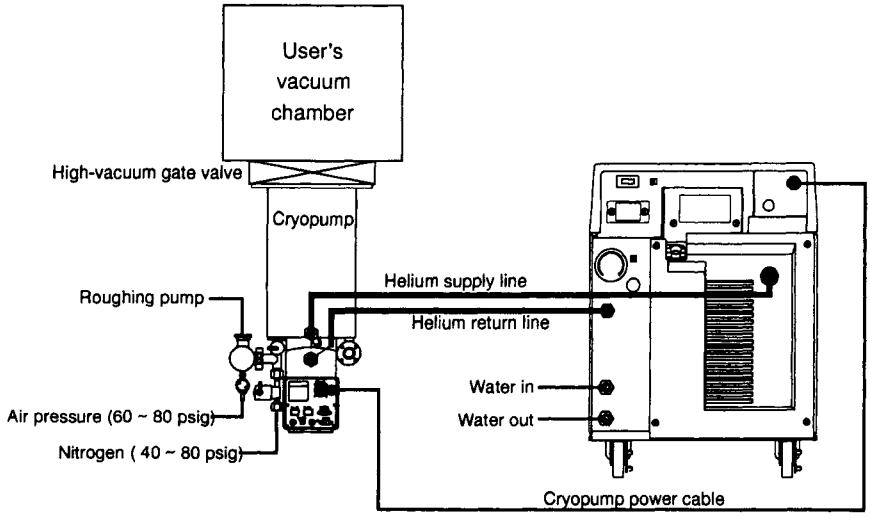
On ultra-high-vacuum systems, base pressures in the 10^{-9} to less than 10^{-11} torr range may be obtained with baking of the chamber for 24 hours or more at temperatures of 100–200°C while the pump is running. Peak temperatures may be limited by the ability of the pump to maintain its own temperature below 20–25 K during chamber bakeout. Although high pressures may be observed during bakeout, the pressure will drop as the chamber (and cryopump) cools. It may be beneficial to then close the high-vacuum valve and regenerate the cryopump to rid the pump of any hydrogen outgassed by the chamber. When base pressures below 2×10^{-10} torr are required, the O-ring sealed pressure relief valve of standard cryopumps is sometimes replaced by an all-metal burst disk. As long as provisions are made to prevent the pump vessel from rising above atmospheric pressure during regeneration, the burst disk may permit slightly lower ultimate pressures. It is not practical to use gas purge or automated regeneration when a burst disk is used, however.

The connections of the purge gas supply, rough pump line, and other utilities are shown in Figure 11. The compressor supplies helium to the cryopump, and may also provide electrical power.

2.5.4.2 Evaporation Systems

Thin-film deposition by evaporation is performed for optical and electronic coatings. Evaporation can be used for a wide variety of coating materials and on a greater range of substrate shapes than sputtering. Cryopumps work well on evaporation systems because they are usually not bothered by the vapors of the deposited materials, and they produce clean vacuum conditions. The evaporation of reactive metals, such as aluminum, produces a conversion of residual water vapor

Fig. II.



Connections of the cryopump to the vacuum chamber and utilities. (Connection diagram for cryopump and compressor courtesy of CTI-Cryogenics.)

into metal oxide and hydrogen at a prodigious rate. It is therefore beneficial to pump the chamber to a base pressure below 10^{-5} torr before beginning deposition to avoid filling the hydrogen capacity of the cryopump prematurely. In addition, thermal and optical baffling of the pumping port may be required to limit the thermal radiation from electron beam guns and other sources from overloading the cryopump.

If ion-assisted deposition (IAD) is used with oxygen at partial pressures in the 10^{-5} to low 10^{-4} torr range, traces of ozone may be detected in the exhaust gases during pump regeneration. Although only parts per million of ozone may be generated during IAD processes, the cryopump regeneration exhaust should be vented to a separate stack to make sure that the ozone is not vented into the room air. This will also prevent reaction with oil vapor or other potential fuels in a common facility stack. Ozone destruction units may be required to meet local codes.

Vapors of metals, oxides, and other evaporated compounds may reach the inlet array of the cryopump. There they will deposit on the ice layer growing on the inlet array and thermal radiation shield. Some may reach the outer surfaces of the second-stage (charcoal) array. In general, these deposits remain after regeneration. Although they may slowly build up into an ugly deposit, they will have little or no effect on pumping speed or capacity. Ordinarily, these deposits may be removed from the inlet array at annual (or more frequent) system maintenance intervals by removing the array from the pump and scrubbing it with a brush in detergent and water.

Because many evaporation chambers have quite large volumes, it may be necessary to rough-pump them to lower base pressures before crossover to the cryopump. This invites oil backstreaming from the mechanical pump into the rough lines and process chamber. Over time, any oil condensed in the chamber will eventually migrate into the cryopump. A possible result is the contamination of the activated carbon of the second-stage array, reducing the hydrogen capacity in extreme cases. The use of Roots-type blowers between the oil-sealed pump and chamber reduces backstreaming when crossover pressures of 10–100 microns are needed. Otherwise, promptly crossing over at the appropriate pressure will at least reduce the time that the rough pump runs in transitional flow.

2.5.4.3 Sputtering Systems

For semiconductor wafer metallization, sputter deposition is the process of choice for aluminum and titanium. Oxide and nitride layers are also sputtered. Cryopumps perform well in these applications because of their inherent cleanliness, reliability, and high speeds for both water vapor and process gases. Since sputtering is normally done in the range of $1-5 \times 10^{-3}$ torr, gas flow is increased to raise the chamber pressure to the desired level. For a closely coupled pump, the 1200 liter/second argon speed would require a flow of 285 std. cm^3/minute (3.6 torr-liter/second) to maintain 3×10^{-3} torr. To obtain long times between regeneration and reduce the expenses for process gas, conductance limiters are used to reduce gas flow to the 50–150 std. cm^3/minute range [11].

A warm (room temperature) conductance limiter, such as a baffle in the pumping port in front of the gate valve, reduces water vapor speed as well as process gas speed by about the same ratio. An alternative is to use a *sputter plate*. The sputter plate replaces the standard 65 K chevron array with a flat copper plate having a number of small holes. At sputtering pressures, a $\frac{1}{2}$-inch-diameter hole has an argon conductance of about 20 liters/second. By selecting a plate with an appropriate number, size, and location of holes, a lower net argon speed will be obtained while maintaining full water vapor pumping speed. A reduction in argon speed to 20–40% of the base speed is common. Although this will also reduce the speeds of air (oxygen and nitrogen), little negative effect is noticed on chamber pumpdown, since water vapor is the principal gas load at that time. Hydrogen speeds are slightly reduced by the sputter plate, but not to the same proportion as argon speeds.

Sputtering with RF magnetron sources of high power in oxygen or oxygen-rich gas mixtures should be avoided with cryopumps. The creation of atomic oxygen and excited O_2 states leads to the formation of ozone in the oxygen ice condensed on the second stage. At low ozone concentrations, the gases exhausted from the

pump during regeneration can be irritating or toxic. Under certain process conditions that are not well defined or limited, more efficient ozone generation takes place. The likelihood for a spontaneous ignition or explosion during regeneration increases due to the catalytic reaction of fine dust from the sputtered reactive metal and liquid ozone. Cryopumps that accumulate oxygen or other reactive and flammable gases should be regenerated as often as possible.

2.5.4.4 Ion Implantation

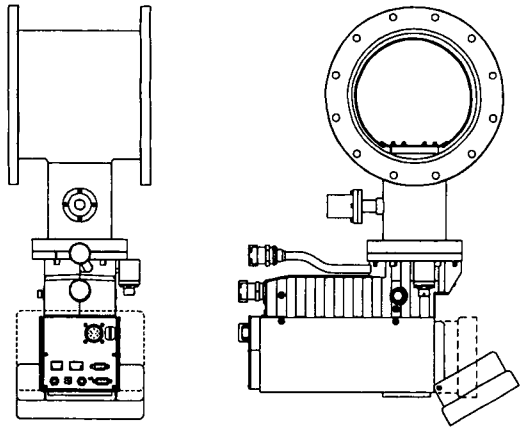
As previously discussed, ion implanters produce heavy hydrogen loads through the decomposition of photoresist on the silicon wafers by the energetic ion beam. Some additional hydrogen load may be added from the ion source itself, although this is usually small by comparison. On an implanter with a high-vacuum pump on the load lock, the regeneration schedule of the cryopump will be determined solely by the hydrogen load. Other things being equal, the hydrogen generated is proportional to the beam current or total dose of the implant. The hydrogen load also goes up as the beam energy (voltage) increases. The impinging ion beam also dissociates the photoresist into heavy hydrocarbons and other gummy by-products that gradually migrate into the cryopump. Implanters without high-vacuum pumps on the load locks are subject to higher water vapor loads in the end station.

For these reasons, special cryopumps have been developed for implanters. First- and second-stage arrays have been optimized for higher hydrogen speeds and capacities. Unique regeneration routines for purging out all the hydrogen and driving out water and photoresist by-products are used. Implanter pumps present special challenges during overhaul and maintenance procedures because the accumulated deposits in the pumps may contain hazardous amounts of arsenic, boron, or phosphorous. Pumps returned for service should be carefully packaged and clearly marked to indicate the presence of hazardous materials. Generally, array parts from these pumps should be discarded under strict hazardous material controls, and cleaning fluids (water and solvents) must also be treated.

2.5.5 CRYOGENIC PUMPS SPECIFICALLY FOR WATER VAPOR

A cold surface held in the range of 105–135 K has very high efficiency for condensing water vapor. At these temperatures, it will condense no other normal gases, except heavy hydrocarbon vapors. Using only a single-stage refrigerator operating at about 107 K, several alternate designs are available to provide eco-

Fig. 12.



An in-line water pump configuration with annular condensing ring. (Line drawing of On-Board waterpump courtesy of CTI-Cryogenics.)

nomic, high-speed auxiliary pumping of water vapor in vacuum applications. In some semiconductor processes, such as etching, reduction in water vapor has resulted in lower particle counts on wafer surfaces.

For process chambers on load-locked systems, the refrigerator may be mounted on the outside wall of the chamber with the cold tip of the cylinder passing through a small-diameter hole in the chamber wall. A copper or aluminum plate within the chamber is attached to the refrigerator heat station. With a net water pumping speed of 10–20 liters/second per square centimeter of plate area, the size of the plate is matched to the heat load it will generate on the refrigerator. A thick plate has lower temperature gradients and can therefore pump effectively at higher average temperatures, although a thinner plate will cool down faster.

The in-line water pump configuration, as shown in Figure 12, uses a cold annular ring in front of a turbomolecular pump. Since turbopumps are relatively expensive for the pumping speed they provide, it becomes economical to add a single-stage cryopump for water. Then high water vapor speeds are obtained for process cleanliness, and adequate speeds for process gas still result. The water pump array configuration produces little loss in oxygen, helium, or hydrogen speed of the turbopump while collecting 95% of the incident water vapor. A turbopump/waterpump combination is useful for applications where the process gases may not be suitable for a standard cryopump.

2.5.6

COMPARISON OF CRYOPUMPS TO OTHER TYPES OF PUMPS

2.5.6.1 Benefits of Cryopumps

There are a number of benefits from using cryopumps.

SPEED

Cryopumps have the highest speeds for any given inlet diameter. For water, cryopumps have essentially 100% of the maximum theoretical pumping speed. For argon, oxygen, nitrogen, and other Type II gases, they are about 40% efficient. Hydrogen efficiencies of cryopumps are about 18–25% of theoretical maximum. This compares well to the 8–13% efficiencies of turbopumps for all gases and the 20–25% efficiencies of LN_2 trapped diffusion pumps for all gases except water (~100% water efficiency of the trap). Ion pumps achieve about 20% for active gases but only 5% for argon and other neutral gases.

COST

They have the lowest cost in liters/second per dollar invested. Although a basic diffusion pump is cheap, the addition of a trap and control systems raises the cost substantially, and operating costs are high for power and cooling water. Turbopumps are expensive to buy for the equivalent pumping speed, and the required inlet diameter is larger. To achieve complete freedom from oil in a turbopump, a magnetic levitation system is required, doubling the cost of the turbopump. The turbopump also requires a backing pump. Ion pumps are not practical in size or cost for high pumping speeds; they cannot be used in high-throughput situations.

CLEANLINESS

In the event of worst-case power failure, cryopumps can only put back into the vacuum system what they took out initially. Cryopumps do not backstream oil, and they cannot provide a path for oil from a backing pump.

RELIABILITY

Cryopumps are highly reliable. The relatively simple mechanical parts of a pump run at very slow speeds, so they do not wear out rapidly. Most pumps will run two

to five years without maintenance. Overhaul costs are a small fraction of the initial purchase cost. In the event of a mechanical failure, none of the parts are injected into the vacuum system.

ORIENTATION

Cryopumps can run in any orientation. Cryopumps operate the same whether they are flange side up, flange side down, or sideways. There is no problem with orientation as in oil-lubricated or magnetically levitated turbopumps, or as in diffusion pumps. You may, however, want to think about where liquids go when the cryopump regenerates.

2.5.6.2 Limitations of Cryopumps

As good as cryopumps are, they do have some limits.

REGENERATION

Cryopumps do need to be regenerated. The capacities for argon and nitrogen are high enough to get through a week or so of heavy process gas flows, but then one to three hours must be devoted to regeneration. The same is true for ion implanters. Some applications have such high gas flows that daily regeneration is required.

HEAT CAPACITY

They have limited heat capacity. Processes with elevated temperatures in the process chamber may overheat the cryopump. Thermal shielding may reduce the net pumping speed to where another type of pump is more practical.

PROCESS GASES

Cryopumps cannot pump all processes. Systems that produce high gas loads of hydrogen or helium are poor candidates for cryopumps. Processes that use highly toxic, flammable, or other dangerous gases in large quantities should not use cryopumps because of concerns for operator safety. This includes arsine, phosphine, silane, and methane, and ozone, among others. Vacuum ovens and furnaces that load oily parts should consider another type of pump.

2.5.7 FUTURE DEVELOPMENTS

The last 5 or 10 years have brought significant improvements in cryopump speeds, capacities, and reductions in regeneration time. Automatic control and communication functions have been added to improve integration of the cryopump with the rest of the vacuum system. Continued advances in all these areas can be expected.

Control of the speed of the drive system of the cryopump promises to bring a number of benefits. At more rapid stroking speeds, the refrigerator gains efficiency. Significantly shorter cooldown and refrigeration times result as the stroke rate is doubled or tripled, if somewhat higher vibration levels can be tolerated during this period. Because this speedup occurs only for a short interval, little reduction in overall life occurs. At slower than standard speeds, gas utilization from the compressor is improved and vibration levels are reduced.

Development of microprocessor controls incorporated into the cryopump, and also distributed in cryopump network controllers, has already enabled pumps to perform regeneration processes that are not practical with manual controls. Further integration of these capabilities can allow the efficient brokering of system resources, such as rough pumps, between load-lock rough-pumping and cryopump regeneration. The ability to sense and control more parameters of pump operation allows analytic and diagnostic functions to be built into the pump. Needs for maintenance can be predicted and communicated to the system operator, either locally or remotely.

It is also becoming more common for cryopumps to be purpose-built for specific processes and even for specific vacuum systems. Various performance parameters can be traded off to obtain the best pump characteristics for a particular process. For these reasons, system designers and pump specifiers need to understand the operating characteristics of cryopump systems to obtain the best level of performance at the right cost.

REFERENCES

1. J. Dewar, *Proc. Roy. Inst.*, **18** (1907) 747–756.
2. W. E. Gifford and H. O. McMahon, *Prog. Refrig. Sci. Technol.*, **1**, eds. M. Jul and A. Jul (Pergamon, Oxford, 1960), p. 105.
3. J. F. O'Hanlon, "A User's Guide to Vacuum Technology" (2nd ed.) (Wiley-Interscience, New York, 1989), p. 410.
4. M. A. Paul, *Physical Chemistry* (D.C. Heath, Boston, 1962), p. 205.
5. R. E. Honig and H. O. Hook, *RCA Review* **21**, (1960) 360.

6. P. A. Lessard, *J. Vac. Sci. Technol.*, **A7 (3)**, (1989) 2373.
7. R. C. Weast, ed., "CRC Handbook of Chemistry and Physics — 66th edition" (CRC Press, Boca Raton, Florida, 1985).
8. R. B. Liebert, L. Ficara, R. Eddy, W. Ghen, and P. A. Lessard, *Ion Implant Technology — 94*, ed. S. Coffa, G. Ferla, F. Priolo, and E. Rimini (Elsevier Science B.V., Amsterdam, 1995), p. 548.
9. D. H. Holkeboer, D. W. Jones, F. Pagano, and D. J. Santeler, "Vacuum Technology and Space Simulation" (American Institute of Physics, New York, 1993), p. 184.
10. M. J. Drinkwine and D. Lichtman, *Partial Pressure Analyzers and Analysis* (monograph) (American Vacuum Society, New York, 1979).
11. G. S. Ash, *37th Ann. Technical Proceedings, Soc. of Vac. Coaters* (Albuquerque, 1994), p. 100.

CHAPTER 2.6

Turbomolecular Pumps

Hinrich Henning
Leybold Vakuum GmbH

2.6.1 TURBOMOLECULAR PUMPS (TMP)

The turbomolecular pump (TMP) invented by Becker [1] in 1957 became commercially available in 1958. Since then it has become very popular in every field of high- and ultra-high-vacuum technique, due to the clean, consistent, and predictable vacuum created, the easy operation, and the advanced degree of operating reliability. The TMP is the only mechanical vacuum pump that together with a roughing pump can attain ultimate pressures in the range below 10^{-10} torr.

2.6.1.1 General Information

PRINCIPLE OF OPERATION

The turbomolecular pump is a bladed turbine that compresses gases by momentum transfer from the rapidly rotating blades of the rotor disks to the gas molecules. The rotor impulse is transmitted to the particles by the superposition of the thermal velocity of colliding particles with the velocity component of the moving rotor surface. The nondirected motion of the particles is changed to a directed motion, creating the pumping process. When the mean free path of the particles is larger than the spacing between rotor and stator blades (molecular flow range,

ISBN 0-12-325065-7
$25.00

typically $<10^{-3}$ torr), particles collide primarily with the rotor, resulting in an efficient pumping process, and there is no interacting influence of the different gases.

In the laminar flow range (typically $>10^{-3}$ torr), the action of the rotor is diminished by the frequent collisions between the particles. Therefore, normally, a TMP is not capable of pumping gases against atmospheric pressure and must be backed by an adequate roughing pump.

HISTORY OF TMP

Historically, the development of the TMP goes back to the year 1913, when Gaede introduced his "molecular drag pump". The "early" molecular pumps (Gaede [2] "molecular drag pump", 1913; Holweck [3] "dual-flow molecular pump", 1923; Siegbahn [4] "disk-type molecular pump", 1940) were never really successful because of their relatively low pumping speed and their questionable reliability, and furthermore, in those days there was no real demand for these pumps. To attain low ultimate pressures with these pumps, the clearances between rotating and stationary parts were made a few hundredths of a millimeter. Therefore, any change in temperature or intruding solid particles could result in a failure of the pump, caused by a seized rotor.

Recently, however, the basic ideas of Gaede (disk) and Holweck (drum) were used successfully in the design of modern pumps (molecular drag pumps; combination pumps), in order to attain extremely low pressures and/or to make use of simple dry roughing pumps.

The Becker design ("turbomolecular pump") avoided these disadvantages: The pump is composed of a series of disks with a row of blades, alternately fixed and moving. In each disk the blades are inclined with respect to the disk plane, in one direction for the blades of the moving disks and in the other for the blades of the fixed disk. The moving disks have a high rotational speed, so that the peripheral speed of the blades (up to 500 m/s) is of the same order of magnitude as the speed of the molecules of the pumped gas.

The distances between these disks are in the range from several tenths to a few millimeters. The channels between the inclined blades of the disks act like elementary molecular pumps, similar to the Gaede molecular drag pump. All channels on one disk, $\approx$20 to 50, are connected in parallel and together can yield a high pumping speed up to several thousand L/s.

2.6.1.2 Theoretical Considerations: Performance Data

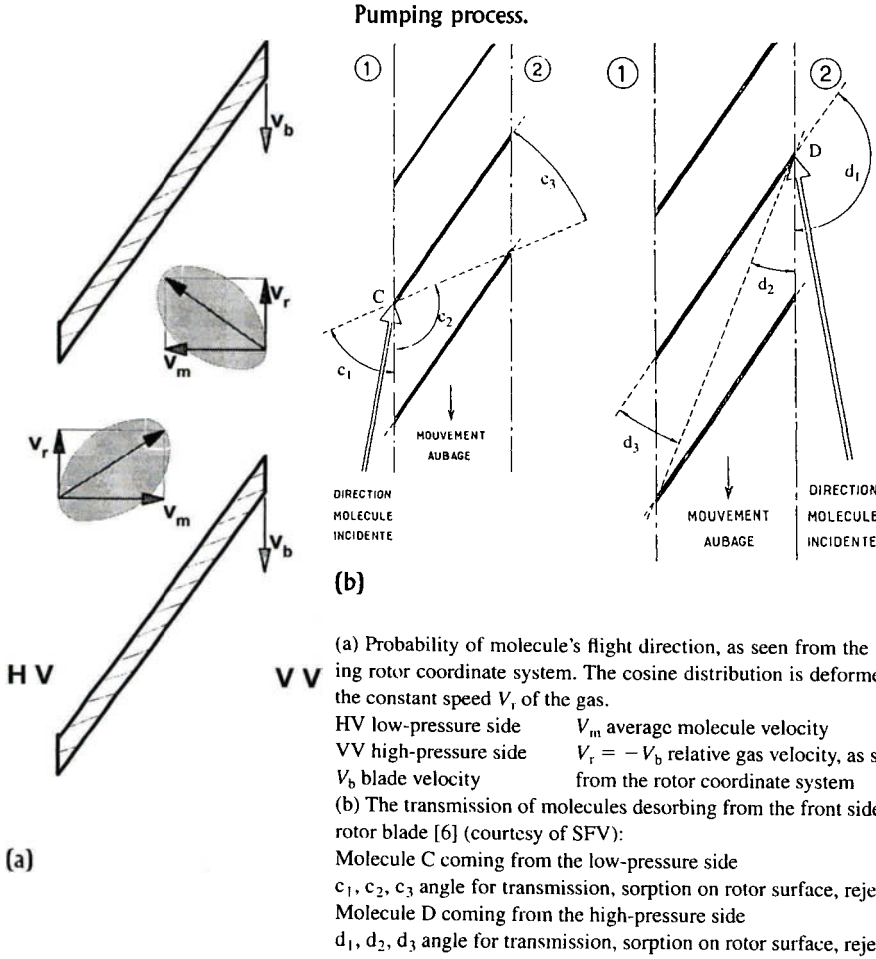
Fundamental theoretical work was performed by A. Shapiro [5].

The main performance data of a TMP, the compression and the pumping speed, can be calculated from the data of single rotor and stator disks. A rotating disk

with blades pumps molecules from one side to the other, and some particles move opposite to the primary flow direction through the blades.

The structure of most common TMPs is shown in Figure 1 [6]. Distribution for the probability of particle direction from the rotor coordinate system is governed by the cosine law. As seen from the moving rotor this distribution will be deformed to a nearly elliptic shape by subtracting the constant rotor speed $v_b = -v_r$. Figure 1(a) shows clearly that particles coming from the high-vacuum side will preferentially pass the rotor structure without collision with the blades. Particles coming from the discharge side will preferentially strike the "lower" side of the blades, where they are desorbed again but now with an undeformed cosine distribution, i.e., less than 50% of them to the high-vacuum side. The ratio

Fig. 1.



(a) Probability of molecule's flight direction, as seen from the moving rotor coordinate system. The cosine distribution is deformed by the constant speed V_r of the gas.

HV low-pressure side — V_m average molecule velocity

VV high-pressure side — $V_r = -V_b$ relative gas velocity, as seen from the rotor coordinate system

V_b blade velocity

(b) The transmission of molecules desorbing from the front side of a rotor blade [6] (courtesy of SFV):

Molecule C coming from the low-pressure side

c_1, c_2, c_3 angle for transmission, sorption on rotor surface, rejection

Molecule D coming from the high-pressure side

d_1, d_2, d_3 angle for transmission, sorption on rotor surface, rejection

of particles coming from the high-vacuum side to the discharge side will be clearly higher than in the opposite direction. As for particles striking the blade surfaces, Figure 1(b) shows

1. A molecule coming from space 1 collides with the blade of the pump rotor; the probability of returning to space 1 is proportional to the angle c_1 and is smaller than the probability of rebounding on another blade c_2 or to be transmitted into space 2 (c_3 probability).
2. The situation is quite different for a molecule coming from space 2, which has a lower transition probability of crossing the blade structure than a molecule coming from 1. The pumping speed of the pump is the net difference between the molecules $1 \rightarrow 2$ and the molecules $2 \rightarrow 1$.

To calculate the performance of a single disk, we must know two extreme values, the maximum compression K_{max} (at zero flow) and the maximum pumping speed S_{max} (for equal pressure on both sides of the disk). These values depend on the transmission probabilities of the particles moving in and opposite the pumping direction. For the molecular flow, these probabilities depend on the ratio v/u of the blade speed v and the mean thermal velocity u of the gas particle, the blade angle α, but not on the pressure. K_{max} of a TMP is exponentially dependent on the blade speed v, a pump-specific factor g (rotor and stator geometry), and the square root of the molecular weight M of the particles pumped:

$$K_{max} \sim \exp\left(M^{1/2} v g\right) \tag{1}$$

S_{max} is proportional to the product of a pump-specific factor G (rotor and stator geometry) and the blade speed v. It does not depend on the pressure and the species of gas pumped:

$$S_{max} \sim vG \tag{2}$$

A disk in operation works between these two extremes. For real pumping conditions (gas throughput $Q = S \cdot p_{inlet}$) and different pressures on both sides of the disk ($K = p_{outlet}/p_{inlet}$), the following relation exists between compression K and real pumping speed S of a single disk:

$$S = S_{max} \frac{K_{max} - K}{K_{max} - 1} = S_{max} \frac{K_{max} - S/S_v}{K_{max} - 1} S \tag{3}$$

More exact calculation takes into account the gap between the rotating disk's outer diameter and the inner diameter of the stator housing and the variation of the blade geometry with the radius. However, the basic relation shown in Equation (3) remains valid.

A TMP is assembled from several disks in series with different blade geometries. Each disk can be regarded as a separate pump. If the pumping speed S_v of the next disk downstream is known and the throughput is constant, K in Equation (3) can be replaced by S/S_v, and beginning with the pumping speed S_v of

the backing pump and using the values S_{max} and K_{max} for the first disk at the fore-vacuum side, the real pumping speed of this disk can be calculated. The formula can be used as a recurrence formula to calculate step by step the pumping speed of the whole pump in the molecular flow. Then the pumping speed of the whole TMP depends on the compression and the pumping speed S_v of the backing pump, as given in Equation (3).

From there it can be derived that if K_{max} for a gas is small, the pumping speed for this gas is a function of the ratio S/S_v. For practical purposes, a TMP with low K_{max} for H_2 needs a larger backing pump to pump H_2 effectively (low S/S_v ratio).

The inlet conductance of the blade area of the disk limits the pumping speed, and the actual pumping speed S of a TMP becomes a function of the ratio v/u [7], and therefore S depends on the molecular weight of the gas pumped.

The pumping speed and the compression of a TMP decrease at pressures of above 10^{-3} torr, caused by the interaction of the particles with one another, as the mean free path becomes smaller than the blade distance and the blades of the rotor disks no longer are in the molecular flow range. This value corresponds to a foreline pressure of 10^{-2} torr to 10^{-1} torr.

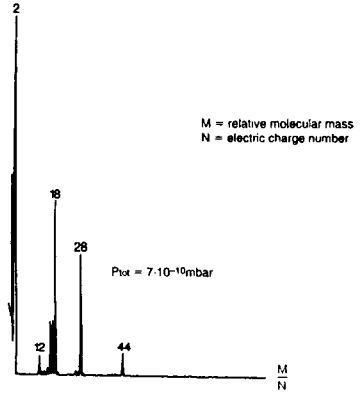
As for any vacuum pump, the ultimate pressure p_0 a TMP can attain can be calculated from the pressure p_v at the outlet side of the pump by dividing it by the maximum compression K_{max}:

$$p_0 = p_v/K_{max} \tag{4}$$

The increase of K_{max} with the molecular weight M means that heavy gases with heavy molecules are highly compressed and have a low backflow probability; this is the reason for the "clean" vacuum without contamination by oil vapors and hydrocarbons (Figure 2). The smaller compression for light gases is the reason

Fig. 2.

Residual gas spectrum of a clean TMP.



• Ordinate in arbitrary units [12]

that the residual gas atmosphere of a TMP consists mainly of H_2. This, however, holds true only for a "clean" system with metallic flange seals. In the case of viton or rubber seals, the ultimate pressure and the residual gas composition look different [8].

2.6.1.3 Typical Vacuum Data of Commercial TMPs

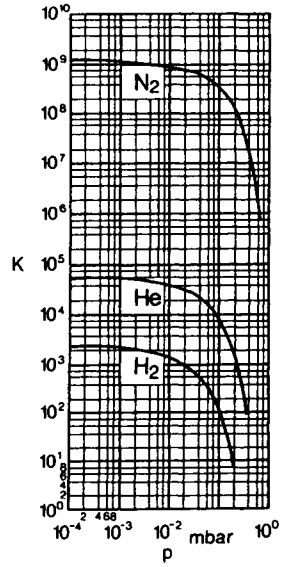
The differences in the data of the different manufacturers depend on their actual design (e.g., blade design, rotor staging) of the TMP.

COMPRESSION

Figure 3 shows the maximum compression K_{max} of a TMP for different gases as a function of the pressure. Table 1 shows typical compression values for TMPs from the catalogues of different manufacturers. The differences in the indicated values depend on the overall design of the TMP (disks, number of disks, staging of disks).

Fig. 3.

K_{max} for different gases as a function of the backing pressure.



[12]
N_2 nitrogen
He helium
H_2 hydrogen

Table 1

Typical Compression Values for TMPs from the Catalogues of Different Manufacturers

K_{max} / Gas Pumped	Manufacturer [9]	[10]	[11]	[12]	[13]
H_2	$3\times10^2\ldots2\times10^3$	$1\times10^3\ldots2\times10^4$	$8\times10^2\ldots1\times10^5$	$1\times10^2\ldots1\times10^4$	$4\times10^2\ldots8\times10^3$
He	$3\times10^3\ldots2\times10^4$	$2\times10^4\ldots1\times10^6$	$6\times10^2\ldots1\times10^7$	$5\times10^2\ldots6\times10^5$	$4\times10^3\ldots3\times10^5$
N_2	$8\times10^7\ldots1\times10^9$	$1\times10^8\ldots1\times10^9$	$5\times10^6\ldots1\times10^9$	$2\times10^7\ldots1\times10^{10}$	$5\times10^7\ldots1\times10^{10}$

PUMPING SPEED

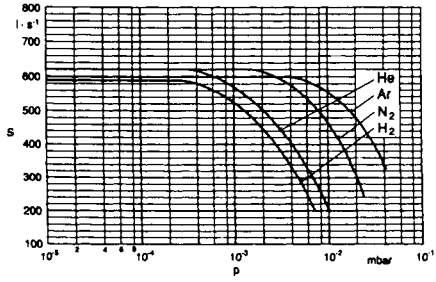
Figure 4 demonstrates the pumping speed of a TMP for different gases.

The product name of a TMP often includes the pumping speed for nitrogen, which is used as the 100% value. This value is compared in Table 2 with the pumping speed data for H_2 and He from the catalogues of different manufacturers.

In addition to the dependence of the indicated values on the internal design parameters already given, other factors also influence these pumping speed ratios. Most commercial TMPs for a certain rotor/stator dimension come with inlet flanges of different diameters, to economize on the accessories (e.g., valves) connected to the inlet flange. The conductance of this flange lowers the effective pumping speed for the heavier gas molecules, for which conductance is low. The maximum compression of the TMP also has some influence on pumping speed for different gases — see Equation (3).

Fig. 4.

Pumping speed for different gases as a function of the inlet pressure.



[12]
N_2 nitrogen
He helium
H_2 hydrogen
Ar argon

Table 2

Pumping Speed Ratios from the Catalogues of Different Manufacturers

Pumping Speed Ratios	Manufacturer				
Gas Pumped	[9]	[10]	[11]	[12]	[13]
H_2 (%)	34 . . . 63	96 . . . 133	57 . . . 98	79 . . . 126	66 . . . 115
He (%)	60 . . . 80	100 . . . 133	73 . . . 110	87 . . . 137	88 . . . 124
N_2 (%)	100	100	100	100	100

ULTIMATE PRESSURE

The ultimate pressure of a commercial TMP is generally between 10^{-10} torr and 10^{-9} torr, using metal flange seals and a two-stage rotary backing pump. The partial pressure of H_2 above an oil-filled two-stage backing pump is approximately 5×10^{-7} torr. Assuming a K_{max} value for the TMP of 1.000, and using Equation (4), an ultimate pressure for H_2 above the TMP of 5×10^{-10} torr could be expected, which corresponds to the typical value of the ultimate pressure of this pump combination.

The bakeout status of a TMP and the kind of sealing material used influence the ultimate pressure and the composition of the residual gas atmosphere. To attain the just-mentioned typical ultimate pressures, the TMP must be baked out, and metal must be used as the sealing material at the high-vacuum side to reduce the gas desorption from the internal surfaces and the sealing material.

2.6.1.4 Actual Design of Turbomolecular Pumps

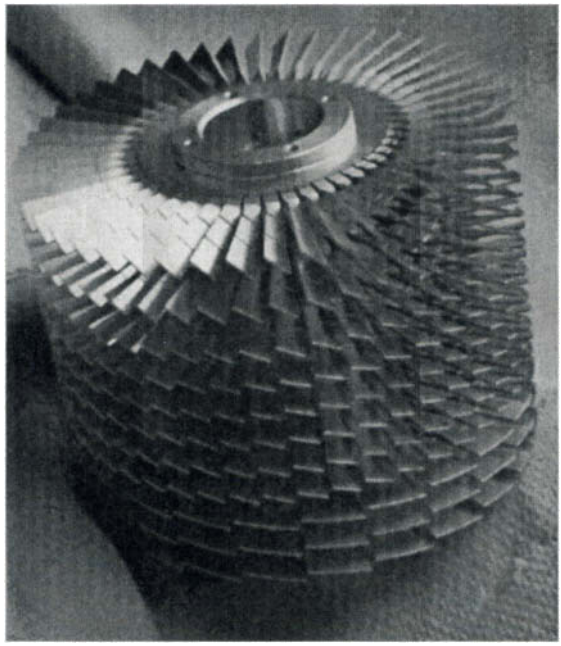
ROTOR AND STATOR GEOMETRY

The pumping speed and the compression of a TMP depend strongly on the rotor geometry and rotational speed. Starting from the 1958 original geometry of the rotor and stator disks, new disk geometries have been developed. These geometries, together with increased rotor speeds, allow much smaller and lighter rotors for the higher-speed range.

Because the gas throughput is constant (the product of pressure and pumping speed) in each stage, the blades nearest to the inlet of the TMP are designed to have a high pumping speed and a low compression, whereas the blades nearest to

Fig. 5.

Rotor with different disk design.



	Pumping Speed	Compression
Strong inclined blades on the HV side (40° . . . 45°)	High	Low
Less inclined blades in the transition region (35° . . . 22°)	Medium	Medium
Low inclined blades on the discharge side (20° . . . 10°)	Low	High

Courtesy of Leybold Vacuum

the foreline port are designed for high compression and low pumping speed (Figure 5). For economic reasons, it would be impractical to make each stage different from its neighbor. A compromise results in groups of two to four different types of blades, in which each is designed for a particular speed and compression ratio.

The methods of manufacturing the rotors and stators influence the pumping speed and compression. Rotors can be made of individually machined disks that are heat-shrunk to the rotor shaft, by machining complete groups of disks from a single block of material or by manufacturing the rotors using spark erosion. The individually machined disks offer the advantage of making them optically opaque, maximizing them for the compression. The other methods of rotor production yield disks with less opaqueness and lower compression, maximizing them for pumping speed.

The stators are either manufactured from individually machined disks or from stampings.

The first commercial TMPs were dual-flow ("horizontal") pumps having a double-ended rotor, pumping gas from a central inlet toward both sides and conducting the gas flow in a common foreline. Single-flow ("vertical") TMPs, using single-ended rotors, became available in 1969 [14]. The double-ended rotor design allows a more stable bearing design, which is advantageous for easy balancing and lower vibration levels. The single-flow design has little conductance losses between the inlet flange and the rotor, whereas the dual-flow design suffers losses from the inlet to both sides. Today only a few models of commercial TMPs still use the dual-flow design.

ROTOR SUSPENSION

The change in the size of the TMP has been quite spectacular. It was made possible by an increase of the circumferential (tip) speed from 150 m/s in 1958 to approximately 400 to 500 m/s today, and by changes of the rotor geometry as well. These high tip speeds relate to high rotational speeds of the rotors, which exert high loads on the rotor's suspension.

Mechanical Bearings

Today most commercial TMPs are equipped with lubricated mechanical rotor bearings, or by a combination of permanent magnet bearing at the high-vacuum side with a lubricated mechanical bearing at the forevacuum side. Depending on the disk diameter, the rotational speed of the rotor goes up to 90,000 rpm. These high rpm have become possible, due to the advances in bearing and balancing technology, without sacrificing the high reliability of the TMP. Today, high-precision ball bearings are available that, when specially tuned to a certain TMP rotor, at comparable radial and axial loads, have a longer lifetime, even at much higher rpm, than bearings of older design.

"Ceramic" bearings (ceramic balls) are widely used today. The ceramic balls exert lower centrifugal forces and lower stress on the races than metal balls, they are harder and more temperature resistant and, therefore, have a stable spherical shape and minimal wear on balls and races. Their surface is smoother, leading to less friction and the pairing of different materials (ceramic balls/steel races) avoids micropitting. Therefore these bearings are more reliable even under lubrication-starved conditions.

Lubrication of Mechanical Rotor Bearings

Three main requirements have to be fulfilled by the lubricant used for mechanical rotor bearings: (1) the lubricant has to cool the rotor bearings, because

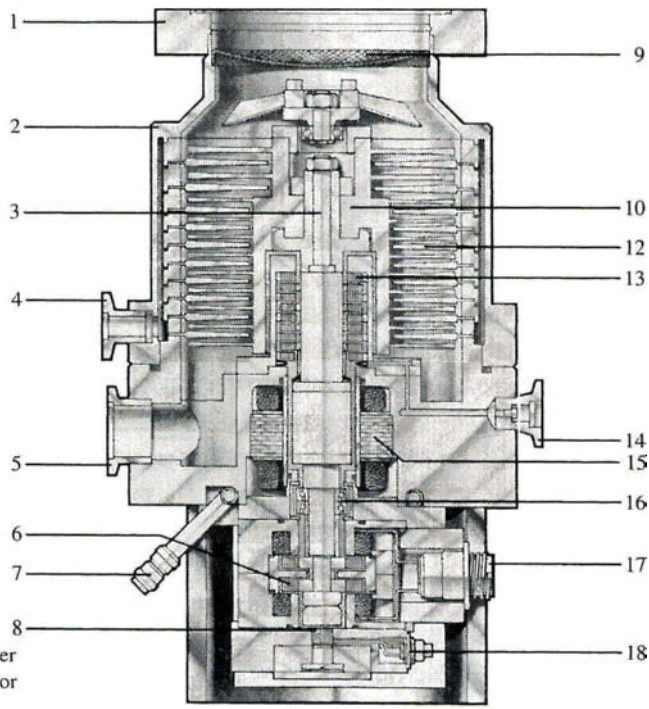
these are running under forevacuum conditions and a heat transfer from the inner to the outer race is possible only by the lubricant, (2) a low vapor pressure, and (3) good lubrication properties at high speed. Today most lubricants used have a synthetic base.

Even at low rotor weights a minimum fluid flow through the bearings for heat transfer is necessary. Many smaller TMPs use a wick lubrication system, while most larger TMPs have pump systems circulating the oil. "Ceramic" bearings with grease lubrication are also used.

Dry Bearings: Magnetic Suspension

After some unsuccessful experiences with commercial "gas-bearing" TMPs [15], today commercial magnetic bearing TMPs are available in which one, two, three, or all the possible five degrees of freedom of the rotor are actively controlled by an electronic system (Figure 6). The position of the rotor spindle is monitored by sensors, and the electronic circuit corrects the spindle position from the data supplied by the sensors. There is no mechanical friction, and hence no wear. The

Fig. 6. Magnetic suspension.



1 HV flange
2 pump case
3 driver shaft
4 venting connection
5 VV flange
6 magnetic bearing stabilizer
7 water-cooling entry
8 axial sensor
9 splinter guard
10 touchdown bearing
11 rotor
12 stator disk
13 permanent magnet bearing
14 purge gas entry
15 electric driver (DC motor)
16 safety ball bearing
17 connector to stabilizer
18 axial sensor connector
[12]

price of these magnetic pumps is still considerably higher than of standard ball-bearing TMPs, limiting their general use.

BALANCING AND VIBRATION

The dynamic balancing of the rotor of a TMP is very important for minimizing vibration and noise levels, which are related to the mechanical bearing lifetime. Due to the high rotational speed of TMP rotors, the centrifugal forces associated with the residual imbalance (material inhomogeneities, radial bearing clearance, geometric imperfections) attain considerable values and transmit vibrations to the body of the pump.

The balancing process modifies the mass distribution of the rotor by adding or taking off material to bring the rotational axis as close as possible to the principal axis of inertia. To further reduce the influence of the residual imbalance, the mechanical rotor bearings within a TMP are held in elastic "antivibration" rings that effectively dampen the residual unbalanced forces.

Dynamic multifrequency balancing of the rotor is generally done in several balancing planes. This relative complex procedure in the last years has become simplified by the use of computers and dedicated software. Modern TMPs have very low residual vibration amplitudes, below 0.02 μm. These low values are required for the use of a TMP with vibration-sensitive instruments, such as mass spectrometers and electron microscopes.

ROTOR MATERIALS

Most commercial TMPs use high-strength Al alloys as rotor material. Compared to other high-strength materials, such as Ti and steel alloys, these Al alloys have a lower specific weight, are much easier to machine, and have sufficient thermal stability for the operational temperatures even at the typical TMP bakeout cycles. The use of ceramics (Si_3N_4) has been reported for the use of TMPs in very strong magnetic fields [16].

A typical maximum stress in the root of the rotor blades of a high-speed TMP rotor at operating conditions is in the range between 50 N/mm^2 (newton per mm^2) and 150 N/mm^2, well below the 0.2% elongation limits of adequate Al alloys (e.g., AlZnMgCu 1,5 F50: 430 N/mm^2).

DRIVE SYSTEMS FOR TMPS

The drive rotor of a TMP is an integral part of the TMP rotor and together with its drive stator is located in the forevacuum area.

Today three different motor systems are used: DC motor, AC motor, and hysteresis motor. The somewhat more expensive DC motor has lower energy consumption and energy losses than the other motors. The motors are driven by solid-state frequency converters. Some of these converters can operate the TMPs at variable rotor speeds.

For special applications (e.g., high radiation with particle accelerators), motor-driven frequency converters are available.

2.6.2 MOLECULAR DRAG PUMPS (MDP)

The molecular drag pump (MDP) depends on the same principle as the TMP. Momentum is transferred from a rapidly moving rotor surface to the particles to be pumped. Therefore the theoretical considerations are very similar.

2.6.2.1 Design

Molecular drag pumps are designed on the basis of the Holweck principle [3]. These pumps include two different sections: (1) an entrance section with a row of TMP blades for maximum conductance to ensure a high pumping speed, and (2) subsequent molecular pumping stages (up to 5) having in principle a drum design with a spiral groove structure to ensure efficiency and a high compression (Figure 7).

2.6.2.2 Typical Performance of Commercial MDPs

Commercial MDPs have a high admissible foreline pressure, ranging from 1 torr to 40 torr, depending on design and number of stages, and therefore can use small simple, dry, and inexpensive forevacuum pumps (e.g., membrane pumps).

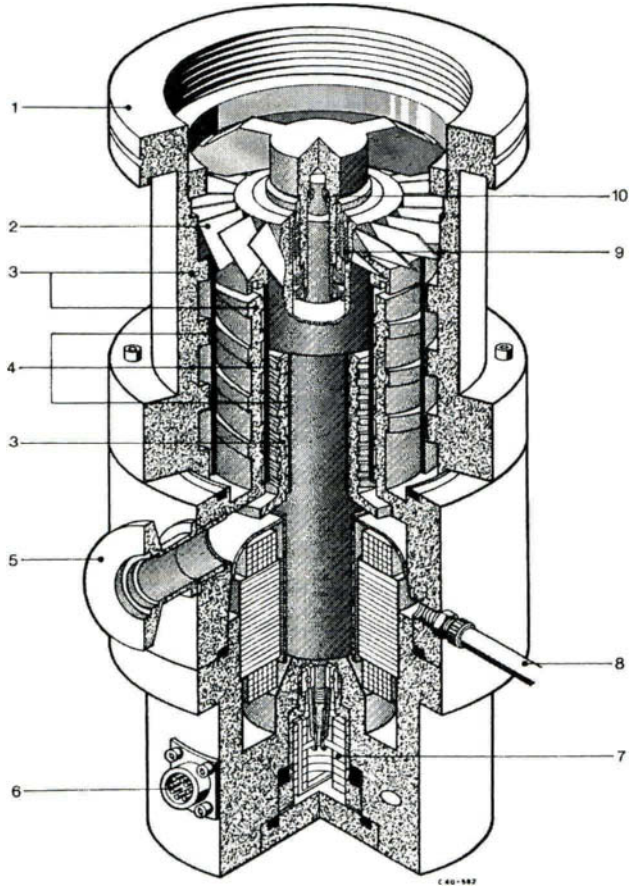
Typical data of commercial MDPs are shown in Tables 3 and 4.

ULTIMATE PRESSURE

Depending on how many molecular pumping stages are used in series, the attainable ultimate pressure is in the range of 10^{-7} torr to 10^{-5} torr using an adequate membrane backing pump.

Fig. 7.

Molecular drag pump.



1 HV flange
2 turbo disk for volume speed enhancement
3 threaded drag pump stator — five stages
4 pump rotor cylinders
5 FV flange
6 socket for electric connector
7 lubricant reservoir
8 purge gas entry
9 permanent magnet bearing
10 safety ball bearing

Courtesy of Pfeiffer Vacuum

Table 3

Compression for Commercial Molecular Drag Pumps

K_{max}	Manufacturer
Gas Pumped	[9, 10]
H_2	$2 \times 10^2 \ldots 1 \times 10^3$
He	$1 \times 10^3 \ldots 2 \times 10^4$
N_2	$1 \times 10^7 \ldots 1 \times 10^9$

Table 4

Pumping Speed for Molecular Drag Pumps

Pumping Speed	Manufacturer
Gas Pumped	[9, 10]
H_2 (%)	40 . . . 56
He (%)	53 . . . 67
N_2 (%)	100

2.6.3 COMBINATION OF PUMPS (TMP + MDP)

Turbomolecular pumps have a good pumping speed but an upper limit of backing pressure near 0.1 mbar. Compared with that, molecular drag pumps may operate at much higher backing pressure of about 10 torr and above but with clearly lower pumping speed.

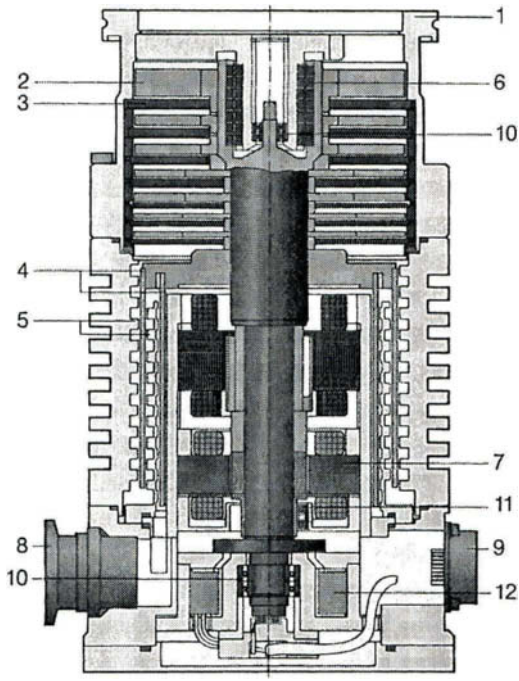
A combination of the two types of pumps will offer the advantage of using the high pumping speed of the turbomolecular pump in the first stage, followed by a molecular drag pump with lower speed but high compression ratio as a second stage. As a result, a backing pressure two orders of magnitude higher than with common turbomolecular pumps is obtained.

2.6.3.1 Design

This capability allows the use of simple small, dry, and inexpensive backing pumps (e.g., membrane pumps) and therefore permits a dry evacuation from atmospheric pressure down to the ultimate pressure.

Fig. 8.

Combination pump.



Turbomolecular part:
- 1 high-vacuum flange
- 2 first rotor disc stage
- 3 first stator disc dtage

Rotor suspension:
- 6 magnetic bearing
- 7 magnetic radial bearing
- 10 safety ball bearing

Molecular drag part:
- 4 cylindrical rotor stage
- 5 threaded stator stage
- 8 backing vacuum flange
- 11 position sensors
- 12 electric connection

Courtesy of Pfeiffer Vacuum

It is common practice to place all rotating elements of the combination pump on a single shaft as shown in Figure 8. If the components are arranged very close together, the gas duct from one to the following will be short, to assure good flow conductance. For the components of the combination pumps, design considerations and performance data as given in paragraphs 2.6.1.3, 2.6.1.4, and 2.6.2 are valid. For most applications of these pumps, a high gas throughput is usual. It is important to avoid any pressure rise caused by conductance limited apertures inside and between the pump components. Similarly, the temperature rise from heating by gas friction must be controlled. The admissible backing pressure is

higher if mechanical spaces in the drag channels are quite narrow. This creates the requirement to limit temperature increases. On the other hand, high operational temperatures are used to reduce undesired reactions or deposits from the pumped media. Thus, careful temperature control is an important design feature of combination pumps.

These combination pumps are available in different design configurations:

- TMP stages + multistage MDP (disk type, derived from Gaede): high forevacuum tolerance
- TMP stages + multistage MDP (drum type, derived from Holweck): high forevacuum tolerance
- complete TMP + multistage MDP (drum type, derived from Holweck): high forevacuum tolerance and very low ultimate pressure.

2.6.3.2 Typical Vacuum Data of Commercial Combination Pumps

Typical data of commercial combination pumps are given in Tables 5 and 6.

Table 5

Compression Data for Commercial Combination Pumps

K_{max}	Manufacturer
Gas Pumped	[9, 10, 11, 13]
H_2	$2 \times 10^3 \ldots 2 \times 10^7$
He	$7 \times 10^3 \ldots 2 \times 10^9$
N_2	$5 \times 10^7 \ldots 1 \times 10^{12}$

Table 6

Pumping Speed Data for Commercial Combination Pumps

Pumping Speed Ratios	Manufacturer
Gas Pumped	[9, 10, 11, 13]
H_2 (%)	24 ... 133
He (%)	50 ... 133
N_2 (%)	100

ULTIMATE PRESSURE

Depending on how many TMP and MDP stages are used in series, the attainable ultimate pressure is in the range of 5×10^{-11} torr to 1×10^{-9} torr.

2.6.3.3 Split-Flow TMPs

Split-flow TMPs [10, 13] are special multiport (two to three inlet ports) combination pumps for the application at, for example, multichamber and differentially pumped systems (e.g., analytical systems: GC-MS/LC-MS, etc.) [17].

Each pumping port conducts the gas at its different pressure separately into the appropriate stage of the combination pump. Together with already compressed other gas, it is then further pumped and exhausted into the common foreline.

2.6.4 EVALUATION OF COMBINATIONS OF BACKING PUMPS AND TMPS, ETC.

2.6.4.1 Nomenclature of the Different Types of Pumps

Depending on the manufacturer of these pumps, different designations are used for the same type of pump. Following is a list of trade names of the different TMPs, MDPs, and combination pumps:

Molecular Drag Pumps (MDP), Pumping Speed Range 7.5–300 L/s

Trade names:

- Drag pumps
- Molecular pumps
- Molecular drag pumps
- Spiromolecular pumps

Turbomolecular Pumps (TMP), Pumping Speed Range 35 L/s–10.000 L/s

Trade names:

- Turbomolecular pumps
- Turbo pumps

Combination Pumps (TMP + MDP), Pumping Speed Range 30–2.000 L/s

Trade names:

- Hybrid pumps
- Macrotorr pumps
- Turbodrag pumps
- Turbospiromolecular pumps
- Wide-range pumps
- Hy.cone pumps

2.6.4.2 Typical Data of Commercial MDPs, TMPs, and Combination Pumps

The following is a summary of the specific features of these pumps which must be considered when selecting a pump for a certain application.

Molecular Drag Pumps (MDP)

Bearings	Type	Lubrication
	Mechanical	Grease
	Mechanical/magnetic	Oil
Pressure range	From $<10^{-5}$ torr to >10 torr . . . 40 torr	
Backing pump required	Foreline pressure: <10 torr . . . 40 torr	

Turbomolecular Pumps (TMP)

Bearings	Type	Lubrication
	Mechanical	Grease, oil
	Mechanical/magnetic	Oil
	Magnetic	None
Pressure range	From $<10^{-10}$ torr to $>10^{-3}$ torr	
Backing pump required	Foreline pressure: $<10^{-1}$ torr	

Pump Combinations (TMP and MDP)

Bearings	Type	Lubrication
	Mechanical	Grease, oil
	Mechanical/magnetic	Oil
	Magnetic	None
Pressure range	From $<10^{-10}$ torr to >1 torr . . . 40 torr	
Backing pump required	Foreline pressure: <1 . . . 40 torr	

2.6.4.3 Typical Data of Backing Pumps

As all the TMPs, MDPs, and combination pumps mentioned need a backing pump, it is worthwhile to look at the existing types of backing pumps, in order to be able to evaluate which backing pump could be used together with the high-vacuum pump.

In principle, the selection of an adequate backing pump depends on its pumping speed, its pressure range, the quality of the forevacuum produced, its price, and its size.

Type	Ultimate Pressure	Quality of Vacuum
Membrane	<10 torr	Absolutely dry
Rotary vane	$<10^{-3}$ torr	Oil vapors
Piston	$<10^{-2}$ torr	Technically dry
Multi-Roots	$<10^{-1}$ torr	Technically dry
Claw-Roots	$<10^{-1}$ torr	Technically dry

Backing pumps producing an "absolutely dry" vacuum do not have any oil or grease in possible contact with the pumping area. "Technically dry" pumps do not have any oil or grease in the pumping area; however, they have grease or oil, e.g., in the bearing or feedthrough areas, which are kept away from the pumping area by dynamic seals such as labyrinths, etc.

2.6.4.4 Adaptation to Applicational Requirements

VENTING

After a power shutdown a TMP should be vented, or it will be contaminated by oil vapors as a result of pressure equalization between exhaust and inlet. By venting with a dry gas to atmospheric pressure, this effect can be suppressed and a contamination of the vacuum system avoided. Most TMPs, with the exception of the larger pumps, may be vented from the high-vacuum side at full rotor speed, if not otherwise contraindicated in the manufacturer's manual. Too strong a pressure rise will deform the rotor blades and make them contact the stator, damaging the pump. If the vacuum system cannot be vented directly, it is advisable to vent the TMP by means of a special vent port that opens into the compression stages. Venting a TMP from the contaminated foreline should be avoided by all means.

BAKING

To attain low ultimate pressures in the high-vacuum and ultra-high-vacuum range, the internal surfaces of a TMP (rotor, stator, housing) must be baked out.

Due to the temperature sensitivity of the aluminium alloy used for the rotors, there is a limit to the maximum value of the baking temperature. This maximum must be well below the critical temperature for the aluminium alloy with respect to its strength. Typical baking temperatures are in the range of 100°C to 140°C, using the bakeout systems and following the operation instructions supplied by the pump manufacturers.

COOLING

To dissipate the frictional heat from the bearing areas, the motor, and the heating by gas at high pressures, the TMPs must be cooled. Although convection cooling is sufficient for small pumps, larger pumps are equipped with fans for air cooling. For many pumps, water cooling is usual.

OPERATION IN MAGNETIC FIELDS

In the presence of magnetic fields, TMPs that have metallic rotors experience induction of eddy currents, which, due to heating, can cause serious problems concerning material strength. Therefore, TMPs with metallic rotors can be used in magnetic fields only if certain maximum values of the magnetic flux density will not be exceeded. These maximum values are specified by the manufacturers and, for static magnetic fields perpendicular to the axis of rotation, typically are in the range of 10 mT to 30 mT (tesla). For pulsed magnetic fields, different maximum values are valid. In case of higher flux densities, the TMP must be shielded magnetically, which can create problems with the distribution of the magnetic field itself. An experimental TMP with a ceramic rotor has been reported that has been tested with a magnetic flux density of 460 mT [16].

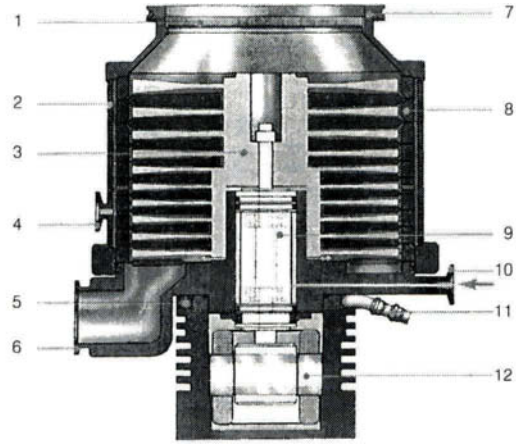
In this context keep in mind that the electronic drive systems and the control systems (magnetic suspension) of TMPs create magnetic stray fields in their vicinity. These stray fields can influence other electronic and magnetic systems. Depending on the position relative to the TMP, typical values for these stray fields are in the range of 1 μT to 50 μT.

PUMPING CORROSIVE GASES

Modern manufacturing techniques require pumping processes where the ultimate system pressure must be in the high-vacuum range and corrosive gases must be handled, e.g., below 10^{-3} torr. At these pressures the rotor material will not be attacked by the corrosives. However, the higher gas density at the forevacuum side in the rotor bearing area requires the use of special lubricants (Fomblin®, Krytox®, etc.), which are corrosion resistant.

Fig. 9.

TMP for pumping corrosive gases (purge gas).



1	splinter guard	7	HV flange
2	pump case	8	stator disk
3	rotor	9	spindle unit with ball bearings
4	venting connection	10	purge gas entry
5	water-cooling circuit	11	water-cooling entry
6	VV flange	12	electric driver [12]

In the semiconductor industry, Si and Al are etched in plasma etch machines, e.g., at pressures of above 10^{-3} torr. To avoid etching and corrosion of the TMP's aluminium rotor, the rotor and all internal parts of the TMP having contact with the corrosives are either made of corrosion-resistant materials or are specially coated (e.g., Ni-plated). Such coatings of the rotor can reduce the attainable ultimate vacuum of a TMP to 5×10^{-8} torr. Besides using special lubricants, for further protection TMPs for plasma etching are equipped with a purge gas system that admits inert gas (N_2, Ar, typically 30 sccm) directly into the bearing area and the forevacuum side of the TMP (Figure 9), where it creates a directed flow to the exhaust of the pump. This flow prevents the corrosive gases from entering the bearing area.

Using TMPs with magnetic suspension in these applications will reduce the problems mentioned.

PUMPING TOXIC OR RADIOACTIVE GASES

An additional range of problems must be solved in the case of a TMP pumping toxic or radioactive gases, e.g., tritium, with application in plasma fusion installations. Because of the safety hazards involved, the TMP must be extremely tight. With a "tritium" TMP, all seals to atmosphere are made of metal. The integral leak rate is below 10^{-9} torr L/s.

TMP IN COMBINATION WITH OTHER HIGH-VACUUM PUMPS

To increase the pumping speed of a TMP for hydrogen or water vapor and make use of its constant throughput and pumping speed, TMPs have been combined with Ti sublimation pumps or, recently, with cryopumps [18].

2.6.4.5 Precautions for the Operation of Pump Combinations for Clean Vacuum

In vacuum systems there are several sources of contamination, such as improperly cleaned surfaces or not using the proper materials for seals, feedthroughs, etc. Additional sources of contamination can be present in the pumping system. An oil-filled backing pump and the lubricant for the rotor bearings of a TMP are potential sources for contamination by oil vapors, which can be avoided with proper procedures. Depending on what components are in the pumping system, there are several ways in which a vacuum system may become contaminated unless some fundamental rules of operation are followed.

STARTING PROCEDURE

When roughing the system by the backing pump prior to starting the TMP, it is very easy to reach pressures below 10^{-1} torr in the roughing phase, allowing molecular backstreaming of roughing pump oil directly into the vacuum system. When in contrast both pumps are switched on together, by the time the pressure is low enough to allow this molecular backstreaming from the roughing pump, the TMP rotor will be accelerated and its compression sufficient to act as an efficient barrier for any backstreaming oil vapor.

Note: Whenever possible, the TMP and the backing pump should be started simultaneously at atmospheric pressure.

In applications that absolutely require a separate roughing phase before switching on the TMP, care should be taken that the roughing line does not enter molecular flow conditions. This could be achieved by pressure control systems or by gas purge systems set to pressures above 10^{-1} torr.

VENTING PROCEDURE

A TMP that is shut down should be vented with a dry, clean gas before its rotor decelerates so much as to allow backstreaming. This procedure is necessary for all combinations of lubricated TMPs with oil-sealed backing pumps.

For TMPs with greased bearings, problems of contamination are reduced to some extent, but care must also be taken while operating these pumps. Table 7

Table 7

Operational Modes that Need Precautions to Avoid Contamination of the Vacuum System

Pump or Combinations	Membrane Pump (absolutely dry)	Rotary Vane Pump (oil vapor)	Multi-Roots Claw-Roots (technically dry)
MDP (mechanical bearing)	Venting	Start/Venting	Venting
TMP (mechanical bearing)	n.a.	Start/Venting	Venting
TMP (magnetic bearing)	n.a.	Start/Venting	
TMP + MDP (mechanical bearing)	Venting	Start/Venting	Venting
TMP + MDP (magnetic bearing)		Start/Venting	

shows the modes of operation of several TMP pumping systems for which the recommended precautions must be observed to avoid contamination.

By observing these rules of operation and regarding the specifics of the different combinations of TMP and backing pump, these combinations can be used to produce and maintain a clean vacuum. However, with some of these combinations, wrong operation can lead to a contamination of the complete vacuum system.

2.6.4.6 Evaluation of Pump Combinations for a Specific Application

Evaluation of a pump combination for a certain application must consider the following:

- Special requirements of the application
- Possible pumps
- Attainable vacuum pressures
- "Quality of the vacuum"
- Price

MANUFACTURING PROCESSES FOR SEMICONDUCTORS

In the following, the evaluations of different pump combinations for use in manufacturing processes for semiconductors are given as examples. The typical manufacturing processes for semiconductors have special requirements concerning the vacuum pumping system:

- Pumping of corrosive vapors (purge gas)
- Abrasive media: dust, granules (membrane pumps not possible)
- The backing pumps must have no oil in the pumping system (chemical reactions, waste problems).

Table 8 shows the evaluation of pump combinations for application to semiconductor manufacturing processes in form of a matrix.

Table 8

Pump Combination Selection for Semiconductor Processes

	Membrane Pump		Rotary Vane Pump (oil)		Multi-Roots Roots-Claw	
MDP	$<10^{-5}$ S: No	HV/FV: O/AD HV:$ FV: $	$<10^{-5}$ S: No	HV/FV: O/O HV:$ FV: $$	$<10^{-5}$ S: No	HV/FV: O/TD HV:$ FV: $$$
TMP (mechanical bearing)	n.a.		$<10^{-9}$ S: No	HV/FV: O/O HV.$$ FV:$$	$<10^{-8}$ S: Part	HV/FV: O/TD HV: $$ FV: $$$
TMP (magnetic bearing)	n.a.		$<10^{-9}$ S: No	HV/FV: AD/O HV:$$$$ FV: $$	$<10^{-8}$ S: Yes	HV/FV: AD/TD HV:$$$$ FV: $$$
TMP + MDP (mechanical bearing)	$<10^{-9}$ S: No	HV/FV: O/AD HV:$$$ FV: $	$<10^{-9}$ S: No	HV/FV: O/O HV:$$$ FV: $$	$<10^{-9}$ S: Part	HV/FV: O/TD HV:$$$ FV: $$$
TMP + MDP (magnetic bearing)	$<10^{-9}$ S: No	HV/FV: AD/AD HV:$$$$ FV: $	$<10^{-9}$ S: No	HV/FV: AD/O HV:$$$$ FV: $$	$<10^{-9}$ S: Yes	HV/FV: AD/TD HV:$$$$ FV: $$$

Legend for the matrix elements:

Ultimate Pressure in torr

Quality of Vacuum Produced
High-/Forevacuum: HV/FV:

O: Contamination by hydrocarbons possible
TD: Technically dry (no oil vapors in the vacuum area of the vacuum pump, however, grease or oil in bearings, etc., close to the vacuum area)
AD: Absolutely dry (no oil vapors, etc., at all present)

For Use in Semiconductor Manufacturing: S

Yes: Pump combination can be used
Partial: Pump combination can be used with minor risk of contamination
No: Pump combination is not recommended for the use in this application

Price:

$ Low price
$$ High price
$$$ Very high price
$$$$ Extremely high price
High-vacuum pump: HV
Fore-vacuum pump: FV

2.6.5

THE USE OF TMP IN APPLICATIONS: SPECIFIC EFFECTS AND DEMANDS

Many applications for TMPs require special measures to attain the requested quality of the vacuum and a safe and reliable operation of the pumping system. The following section gives some specific effects of the most important applications for TMPs, etc., and the measures to cope with them.

Application	Mass Spectrometers	Leak Detectors	Electron Microscopes	Particle Accelerators	Fusion Installations, R & D	Etchers CVD	Load-lock/ Transfer	
Specific Effects	• = does apply (•) = may apply							Measures
High Inlet Pressures		•				•	(•)	Throttle valves, adequate fore-vacuum system, combination.
....	....	....	....	....	....	....	....	pumps.
Aggressive Gases	(•)				•	•	(•)	Purge gas, anti-corrosive coating, magnetic
....	....	....	....	....	....	....	....	suspension.
Solid Particles, Dust		(•)				•		Inlet filter, separators, purge gas, venting
....	....	....	....	....	....	....	....	procedure.
Deposition						•		Magnetic suspension, heating,
....	....	....	....	....	....	....	....	inlet baffle.
High Voltage				(•)				Electronic measures for pump and electronic
....	....	....	....	....	....	....	....	drive system.
Magnetic Field	(•)		(•)	(•)	•			Magnetic shielding (μ-metal).
....	....	....	....	....	....	....	....	
Mechanical Shocks	(•)	(•)			•		(•)	Magnetic suspension, damping elements, avoiding of hard im-
....	....	....	....	....	....	....	....	pacts (valves).
Vibration Sensitivity	•		•					Magnetic suspension, damping elements.

(*continues*)

Application	Mass Spectrometers	Leak Detectors	Electron Microscopes	Particle Accelerators	Fusion Installations, R & D	Etchers CVD	Load-lock/ Transfer	
Specific Effects	• = does apply (•) = may apply							Measures
Hydrocarbon Sensitivity	(•)	(•)	•	(•)	•	•	•	Magnetic suspension, grease-lubricated bearings, "dry" backing pump.
....	....	....	....	....	....	....	....	
Pressure Shocks	(•)	•)		(•)	•	•	•	Fast forevacuum valves, combination pumps, magnetic suspension, buffer volume, venting procedures.
....	....	....	....	....	....	....	....	
Ultimate Pressure	(•)		•	•	•			Baking, CF-flanges, high compression for H_2.
....	....	....	....	....	....	....	....	
Ionizing Radiation				•	•			Lead shielding, radiation-resistant materials.
....	....	....	....	....	....	....	....	
Electro-magnetic Compatibility	•		•	•	(•)			Electronic measures.

Application	Ion Implanter Source	Ion Implanter Beamline	Evacuation of TV Tubes	Aluminizing of TV Tubes	Coating, PVD	Sputtering	
Specific Effects							Measures
High Inlet Pressures	(•)				•	•	Throttle valves, adequate fore-vacuum system, combination pumps.
....	....	....	....	....	....	....	....
Aggressive Gases	•	(•)			•	•	Purge gas, anti-corrosive coating, magnetic suspension.

(*continues*)

Application	Ion Implanter Source	Ion Implanter Beamline	Evacuation of TV Tubes	Aluminizing of TV Tubes	Coating, PVD	Sputtering	
Specific Effects							Measures
Solid Particles, Dust			•	•	•	•	Inlet filter, separators, purge gas, venting procedure.
Deposition	•			•	•	•	Magnetic suspension, heating, inlet baffle.
High Voltage	•	•			(•)	•	Electronic measures for pump and electronic drive system.
Magnetic Field	•	•					Magnetic shielding (μ-metal).
Mechanical Shocks			•	•			Magnetic suspension, damping elements, avoiding of hard impacts (valves).
Vibration Sensitivity							Magnetic suspension, damping elements.
Hydrocarbon Sensitivity		•	(•)	(•)		•	Magnetic suspension, grease-lubricated bearings, "dry" backing pump.
Pressure Shocks			•	•	(•)	(•)	Fast forevacuum valves, combination pumps, magnetic suspension, buffer volume, venting procedures.
Ultimate pressure							Baking, CF-flanges, high compression for H_2.
Ionizing Radiation							Lead shielding, radiation-resistant materials.
Electromagnetic Compatibility	•	•					Electronic measures.

2.6.6 AVOIDING OPERATIONAL MISTAKES

While operating TMPs, etc., the following rules, if not in contradiction to directions from the manufacturer's manuals, must be observed carefully, to take full advantage of the features of these pumps.

2.6.6.1 Entering of Solid Particles

Solid particles entering the inlet of a TMP may damage rotor and bearings. A protection screen should be used. However, these screens reduce the pumping speed by approximately 20 to 30%.

Large quantities of dust entering the TMP can block the gaps between blades and pump body. Therefore, it is advisable to use dust filters at the inlet of the TMP. Here too, pumping speed losses caused by the low flow conductance of these filters will occur.

2.6.6.2 Inclined or Horizontal Position of the Pump

If the pump is mounted in a horizontal position, the outlet flange (forevacuum side) should point downward, to avoid condensates (water vapor, oil vapor from the forevacuum tubulation) from accumulating in the bearing area of the pump (risk of damage).

2.6.6.3 Baking of the Pump

While baking a pump to attain low pressures, the maximum permissible temperatures, according to the manufaturer's specification, must be respected. Overheating can lead to bearing and rotor failures (rotor material loses its strength).

Do not bake the pump when the system is not baked. The gases or vapors released from the pump during bakeout will condense on the cooler surfaces of the system.

2.6.6.4 Replacement of Lubricant

The time intervals for replacing the lubricant, indicated by the manufacturer, should be respected to assure full performance and lifetime of the bearings.

2.6.6.5 Backing Pressure

With conventional TMPs the backing pressure should typically not exeed 0.5 torr. With higher foreline pressures, the pump heats up (gas friction) and the rotor and/or the bearings could be damaged. In addition, if the rotor is not in the molecular flow range, the pumping speed and the compression are reduced.

The use of rotary vane pumps (oil-sealed) should be restricted to two-stage pumps only, since single-stage pumps of this type with opened gas ballast will bring the risk of not attaining the necessary vacuum level in the backing line.

2.6.6.6 Dimensions of Inlet Connections

It is recommended that the high-vacuum connection have at least the same diameter as the inlet flange of the pump. Smaller diameters lead to significant pumping speed losses caused by the low conductances.

For example, for a TMP with a nominal pumping speed of 500 L/s for N_2 and an inlet flange of 150 mm I.D., the pumping speed losses for N_2 using a 300-mm-long connection of different inner diameter are as follows:

High-Vacuum Connection		Effective Pumping Speed	Pumping Speed Loss
I.D.	Length		
150 mm	300 mm	370 L/s	26%
100 mm	300 mm	100 L/s	80%
40 mm	300 mm	20 L/s	96%

REFERENCES

1. W. Becker, *Vak. Tech.* **7** (1958) 149.
2. W. Gaede, *Ann. Phys.* **41** (1913) 337.
3. M. Holweck, *C.R. Acad. Sci. Paris* **177** (1923) 43.
4. M. Siegbahn, *Arch. Math. Astr. Fys.* **30B** (1944) 17.
5. C. H. Kruger and A. H. Shapiro, in *Transactions of the 7th National Vacuum Symposium*, AVS, New York, 1960, pp. 6–12.
6. L. Maurice, S. Sagot, Le Vide 111 (1964) 109.
7. K. H. Bernhardt, *J. Vac. Sci. Technol.* **A1** (1983) 136.
8. J. Henning, *Vacuum* **21** (1971) 523.
9. Produkt Leitfaden, Alcatel Hochvakuumtechnik GmbH, Annecy 1995.

10. Komponenten für die Vakuumtechnik, Balzers-Pfeiffer GmbH, Balzers 1996.
11. Vakuum Katalog, Edwards Hochvakuum GmbH, Marburg 1993.
12. Katalog HV 300, Teil B, Leybold AG, Cologne 1994.
13. Vacuum Products Catalog, Varian Ass. Inc., Lexington 1995/1996.
14. K. H. Mirgel, *J. Vac. Sci. Technol.* **9** (1972) 408.
15. L. Maurice, *Jap. J. Appl. Phys. Suppl.* **2** (1975) 21.
16. Y. Murakami, T. Abe, S. Mori, N. Nakaishi, and S. Hata, *J. Vac. Sci. Technol.* **A5** (1987) 2599.
17. A. Conrad and O. Ganschow, *R&D Magazine*, April 1994, p. 35.
18. J. E. de Rijke and W. A. Klages, Jr., *Solid State Technology*, April 1994, p. 63.

CHAPTER 2.7

Pumps for Ultra-High Vacuum Applications

Jack H. Singleton
Consultant

An ultra-high vacuum (UHV), typically defined as a pressure below 10^{-8} torr, can be obtained using many of the high-vacuum pumps already described in Chapters 2.4, 2.5 and 2.6, but the successful application of *any* pump for UHV use requires that the influx of gas into the system from all sources — including outgassing, permeation through the walls, and leakage from the atmosphere — be rigidly controlled.

The equilibrium pressure, P (torr), in a system is the result of the balance between the rate of influx of gas from all sources, Q (torr liter/sec), and the effective pumping speed of the pumps, S (liter/sec):

$$P = Q/S$$

In a typical, leak-tight, unbaked system using a cryopump, turbomolecular pump, or liquid-nitrogen-trapped diffusion pump, pressures in the low (10^{-7}) to high (10^{-8} torr) range can readily be obtained after a few days pumping. Over longer pumping periods, the base pressure continues to fall, but at a much slower rate, and pressures below the 10^{-8} torr range are hardly ever attained in an unbaked system. Outgassing from the system walls, typically in the range of 10^{-9} to 10^{-10} torr liter/sec cm^2, is the limiting factor. The achievement of UHV conditions, say in the 10^{-10} torr range, simply by using a bigger pump would require an increase in speed of the order of 10^5, which is generally impractical. The simplest practical

ISBN 0-12-325065-7
$25.00

approach is to design a vacuum system that can be baked out at a temperature of at least 200°C higher than its normal operating temperature so as to reduce the rate of outgassing by a factor of about 10^5.

2.7.1 SYSTEM DESIGN FOR ULTRA-HIGH VACUUM

To select an appropriate pump for ultra-high vacuum use, it is useful to provide a few basic rules for the design of an entire system that will achieve the ultimate pressure within 24 to 48 hours of startup.

1. As much as possible of the system should be bakeable to at least 200°C higher than the normal operating temperature, so that materials used for construction must be stable at the bakeout temperature.

2. Minimize the surface area of the system, because the outgassing rate is proportional to area. The use of porous materials, such as anodized aluminum, should be avoided.

3. Wall materials that are permeable to atmospheric gases should be eliminated, or used very sparingly. This particularly applies to the use of O-ring seals, which provide reliable and simple construction for unbaked systems. Such seals not only limit the temperature to which the system can be baked, but they also permit permeation of atmospheric gases through the elastomer into the vacuum system, providing a rate of influx that is second only to that from outgassing. During a bakeout, the rate of permeation actually increases, falling to its original value as the system cools to normal operating temperature. Conversely, the rate of permeation can be reduced by cooling the seals, although this is generally not a practical procedure. Demountable O-ring sealed flanges can be easily replaced by metal-gasketed flanges such as the Conflat® geometry, eliminating permeation. The selection of valves is particularly important; permeation from atmosphere can be avoided by using a metal-gasket bonnet seal and a bellows-sealed shaft seal. For the most rigorous conditions, the internal seal on the nose of the valve is a metal seal, such as copper, silver, or gold. Elastomers may be used to provide somewhat easier sealing (lower force) but at the expense of a limitation in the bakeout temperature: Viton, Kalrez®, and polyimide elastomers have been used, with usable temperatures of approximately 200, 250, and 300°C, respectively, with the valve baked in the open position.

The efficacy of the procedures just recommended is dramatically illustrated by the work of Alpert and his colleagues [1] in developing practical techniques for achieving UHV. Their original glass vacuum systems were outgassed by baking overnight at ~400°C, while pumping with a liquid-nitrogen-trapped diffusion

pump, then isolated from that pump, so that the only remaining pumping action was that of the Bayard-Alpert ionization gauge, which provided a speed of ~0.1 liter/sec when operated at an electron emission of 10 mA; pressures in the low 10^{-10} torr range were routinely achieved. Note that in this work, the *entire* system on which the gauge was pumping was baked at 400°C so that very low total outgassing rates were achieved. In subsequent work, the vacuum systems were continuously pumped by a liquid-nitrogen-trapped glass diffusion pump; this provided a very limited speed of about 0.1 liter/sec, largely limited by the small conductance of the all-metal Alpert valve [1] that was used as the system isolation valve. The connection from that valve to the diffusion pump could not be effectively baked out, and therefore had a much higher outgassing rate than the baked part of the system; the fact that UHV was achievable despite this outgassing load must be ascribed to relatively high speed of the diffusion pump, typically ~30 liter/sec. Clearly the greater part of the pumping capacity was used to control the outgassing load from a small section of the entire system.

2.7.2

THE SELECTION OF PUMPS FOR ULTRA-HIGH VACUUM APPLICATIONS

All the high-vacuum pumps described in Chapters 2.4, 2.5, and 2.6 can be successfully used to exhaust UHV systems. However, it must be noted that many commercial versions of the pumps cannot be baked to temperatures that are required for rapid outgassing. It is therefore essential to minimize the section of the system that cannot be effectively outgassed, and to provide adequate pumping speed to handle the higher outgassing influx from this section. These pumps are increasingly used for UHV applications, especially in systems that are rarely vented to atmospheric pressure. In such applications, the initial pumpdown of the system may be protracted because surfaces in and adjacent to the pump remain at too low a temperature for *rapid* outgassing, but once a satisfactory ultimate pressure has been reached, the system remains continuously in the UHV range until a failure occurs, or routine maintenance is required. For example, in thin-film deposition production systems and surface analysis systems, samples are introduced to and removed from the vacuum chamber by use of a load-lock system, thus minimizing gas influx at all times.

The following brief comments discuss the use and suitability of the pumps, described in Chapters 2.4, 2.5, and 2.6 and in Sections 2.7.3 and 2.7.4, for UHV applications.

2.7.2.1 Diffusion Pumps

Pressures in the low 10^{-10} or high 10^{-11} torr range can be obtained using a three- or four-stage pump in conjunction with an *efficient* liquid-nitrogen-cooled trap. To prevent oil backstreaming, such a trap must ensure that any oil molecule will make two or more collisions on a surface that is continuously maintained at liquid nitrogen temperature. Pressures in the 10^{-9} torr range can be obtained using only a water-cooled baffle and a low-vapor-pressure pumping fluid of either the polyphenyl ether or silicone type (see Chapter 2.4). Diffusion pumps are not selective, being effective for all gases, and are commercially available with very large pumping speeds, and therefore are a good general-purpose pump for all system sizes. However, satisfactory operation requires continuous maintenance of the electrical heating, water or air cooling, a forepressure below some critical level (often ~0.5 torr), and liquid nitrogen cooling of the trap. Failure of *any* of these items can result in oil contamination of the system. Clearly, preventive maintenance is crucial.

Oil contamination of diffusion-pumped systems is most frequently a result of initially roughing the system to too low a pressure using an untrapped oil-sealed mechanical pump. It can be readily avoided by crossing over to the diffusion pump when the foreline pressure is still in the viscous flow region, or alternatively, by using a properly maintained, effective foreline trap.

The limitations on bakeout described earlier, are particularly severe for a diffusion pump/liquid nitrogen trap assembly, so achievement of the ultimate pressure may be an unacceptably long process.

Recommendation: Because of the multiplicity of factors that can affect the performance of this pump, the difficulty in outgassing the pump and trap, and the serious potential for oil contamination, diffusion pumps are *not* recommended for UHV applications. Exceptions to this recommendation might include the requirement for low capital expense (noting that running expense is high) or very large pumping speed.

2.7.2.2 Cryogenic Pumps

A cryopump provides a nonselective, intrinsically clean method for UHV pumping. Very large pumping speeds are readily available, and the simplicity of the device permits reliable operation. The capacity for pumping helium, hydrogen, and neon is very limited, compared to that for all other gases, but this is rarely of concern in a UHV system. The main drawback of such pumps is that the pumped gases are released very rapidly within 10 minutes of a power interruption, so that

a rapid-acting, leak-tight, bakeable valve is essential to isolate the system during such an emergency, and for use during a scheduled regeneration of the pump.

The limitations of bakeout, described earlier, apply also to the cryopump, although the problem is far less severe than for the diffusion pump. The fact that a large part of the cryopump assembly operates at low temperatures automatically assures that the outgassing rate decreases once the pump is started.

The procedures used to rough out both cryopump and vacuum system must avoid contamination, so that the same considerations already described in Section 2.7.2.1 apply.

Recommendation. This pump is of value where very large speeds are essential. The main problems are associated with the limited ability for bakeout, and the rapid release of previously pumped gas in the event of a power failure. A cryopump can pump only small quantities of helium before regeneration is required and should never be used if a substantial influx of this gas is ever present.

2.7.2.3 Turbomolecular Pumps

A turbomolecular pump provides a nonselective, clean method for UHV pumping, combining simplicity of operation and high reliability. Current pumps are limited to a maximum speed of 10,000 L/s, but this is not a limitation for most UHV applications. Bakeout of the vacuum system is much less of a limitation than for the two preceding pumps. For example, for one commercial pump a temperature of 160°C is permissible at the mounting flange.

Although the pumps are effective for all gases, a limitation should be noted in the pumping of hydrogen, one of the principal residual gases in a UHV system. The compression ratio for this gas is always less than for all other gases, frequently being as low as 600 in a standard turbomolecular pump, resulting in a limit to the minimum residual hydrogen pressure that can be achieved. For example, if it is essential to maintain a hydrogen partial pressure of 1×10^{-10} torr in the vacuum chamber the foreline partial pressure of hydrogen must be maintained at $\sim 6 \times 10^{-8}$ torr or less. Some oil-sealed mechanical pumps cannot meet this requirement, and in fact generate hydrogen by degradation of the oil. In critical cases, the addition of a titanium sublimator pump (see Section 2.7.4.1) to the vacuum chamber will resolve the problem. An alternate and preferable approach is to use a wide-range turbopump, where the turbo stage is backed up by a molecular drag stage, providing compression ratios for hydrogen as high as 10^7; an additional advantage is that such pumps can use an oil-free diaphragm forepump, thus eliminating any possibility of oil contamination from that source.

Recommendation. For general UHV applications, the turbomolecular pump is probably the best choice among pumps described in Chapters 2.4, 2.5, and 2.6. The growing reliability of the bearings systems, the availability of magnetic bear-

ing pumps, and the combination turbomolecular drag pumps are important factors in this recommendation. The ability to subject the pump to even a limited bakeout permits better outgassing of the entire system, especially because this bakeout can be performed with the pump in operation. The major difficulty to be considered is that a power failure results in loss of pumping within a minute or two, as the pump rotor slows down, and a rapid-acting, leak-tight, bakeable valve must be present to isolate the UHV chamber from the pumping line in such an eventuality.

2.7.2.4 Sputter-Ion (Getter-ion) Pumps

Sputter-ion pumps, which are discussed in detail in Section 2.7.3, are particularly suited to UHV applications provided that the throughput of gas is relatively low. Because of their simple construction, and lack of moving parts, they are highly reliable, easily degassed, and provide virtually fail-safe operation.

Their disadvantages include high selectivity in pumping different gases, a serious decrease in pumping speed at the lowest pressures, and high initial expense. They are not available in very large pumping speeds, and are not appropriate if very large quantities of gas must be pumped, primarily because such applications result in short operational life.

Recommendation. The sputter-ion pump is probably the best choice for general UHV applications if it is essential that a low pressure be maintained at all times. It provides by far the closest approach to fail-safe operation, and the operating life is very long at low pressures. However, if a system must be frequently cycled to atmospheric pressure, or if a significant gas load must be handled by the pump — for example, argon in a sputtering system — the operational life will be reduced, and a turbomolecular pump may be a better choice.

2.7.2.5 Sublimation Pumps

These pumps are discussed in detail in Section 2.7.4.1. They provide very-high-speed pumping of chemically reactive gas at a relatively low capital and operating cost. They do not pump the rare gases or methane and similar highly stable organic molecules, and must therefore be used in conjunction with a second pump, most commonly a sputter-ion pump, which is effective for pumping such gases.

Recommendation. The sublimation pump is invaluable in providing very high speeds *for reactive gases.* The pump contains no moving parts, which provides high reliability. It is normally an integral part of the vacuum chamber and is therefore effectively outgassed during bakeout. Most frequently used in conjunction

with a sputter-ion pump, it is also useful in conjunction with a standard turbomolecular pump when a very low residual hydrogen partial pressure is required.

2.7.2.6 Nonevaporable Getter (NEG) Pumps

These pumps are discussed in detail in Section 2.7.4. They find application both in very large systems, and in very small, sealed-off devices. If the gas influx is primarily hydrogen, deuterium, or tritium, high-speed pumping can be achieved at ambient temperature. They are becoming widely used as distributed pumps in very large high-energy physics vacuum systems. The pumps are useful for all chemically reactive gases.

Recommendation. NEGs are excellent for pumping hydrogen and its isotopes and all reactive gases, particularly where a simple, reliable device, operating at relatively low or, in some cases, ambient temperature, is required. The combination of a NEG with a sputter-ion pump can maintain hydrogen pumping to very low pressures. The total quantity of gas that can be pumped before a NEG is saturated is quite low as compared to most pumps previously discussed; they are not suitable for high-throughput applications.

2.7.3 SPUTTER-ION PUMPS

In this section, two types of pumps that are particularly well suited for UHV applications, sputter-ion and getter pumps, will be described in detail. They are classified as capture pumps, because the gases that are pumped are trapped within the pump structure, most of them being irreversibly immobilized as stable chemical compounds. The pumps can be baked to at least as high a temperature as a turbomolecular pump, while operating, and their simple structure permits effective, fairly rapid outgassing under these conditions. For the ultimate in rapid cycling from atmospheric pressure, they can be baked as high as 450°C, simultaneously with the vacuum chamber,[1] so that very rapid outgassing of *all* surfaces within the vacuum system can be achieved. At the higher bakeout temperatures, neither sputter-ion or getter pumps are normally operated, so that the system must be pumped using a subsidiary turbomolecular or similar "clean" pump. The advantage of such baking is clearly illustrated by the fact that routine operation of a new system in the 10^{-10} torr range has frequently been demonstrated after a single 16-hour bake at 400°C.

[1] For bakeout above 250°C, the magnet and connecting cable may have to be removed (see section 2.7.3.2).

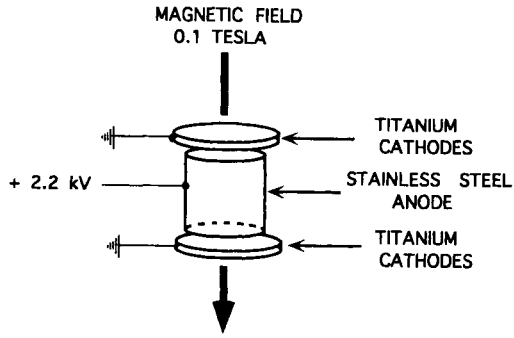
When sputter-ion and getter pumps are used, the system is fully isolated from the atmosphere, and this, coupled with the fact that a UHV system must be free from significant leaks, ensures a degree of fail-safe operation that is unmatched by any other type of pump. If a power failure shuts off the pumps, only a relatively modest pressure rise will normally occur, and the system will rapidly return to its normal operating pressure once power is restored.

2.7.3.1 Operating Mechanism of Sputter-Ion Pumps

The first sputter-ion pump, which became commercially available in 1958, was the so-called diode type [2,3]. Figure 1 is a drawing of a small single-cell sputter-ion pump. It has a pumping speed of about 0.2 liter/sec, and is typical of an "appendage" pump, which is used to maintain a low pressure in a sealed-off electronic device throughout its useful life. A cylindrical stainless steel anode, about 1 cm in diameter and 1 cm long, is maintained at a potential of ~2.3 kv with respect to a pair of cathodes, which are positioned a few millimeters from each end of the cylinder, in contact with the pump envelope, at ground potential. A permanent magnet provides a field of about 0.1 tesla, directed along the axis of the cylinder. Any electron that is produced in the body of the pump, perhaps by field emission or cosmic radiation, is accelerated toward the anode, but is constrained to move in a circular orbit by the magnetic field, and thus is trapped and cannot reach the anode. If the electron collides with a gas molecule, causing ionization, a second electron is produced and a cascade process quickly generates an electron density of the order of 10^{10} electrons/cm^2, which is sustained over a wide pressure range [4]; this high density of electrons provides efficient ionization of gas entering the pump.

The electrical and magnetic fields, for the electrode geometry shown in Figure 1, appear to be oriented parallel to each other, but in fact the space charge

Fig. 1.



Appendage single-cell sputter-ion pump.

caused by the high density of electrons has the effect of depressing the potential along the axis of the anode to that of the cathode, so that a radial electrical field results. This particular geometry, with the electrical and magnetic fields orthogonal to each other, is known as a *Penning Discharge*, and was described by Penning, in 1937, in a paper [5] discussing a cold cathode discharge device for the measurement of pressure (the Philips gauge).

The collision of an electron and a gas molecule generates both an electron and a positive ion. The ions are accelerated in the electrical field to a high energy, collide with the cathodes, which are typically a chemically reactive metal such as titanium, and are neutralized. Many are buried in the cathodes to a maximum depth of about a micron, and this burial process provides one primary pumping mechanism. A smaller number of the ions are reflected, with some loss of energy, as energetic neutrals, which are no longer constrained by the electrical and magnetic fields and bury themselves in any surface within the pump that is in line of sight. This process provides a second pumping mechanism, which is particularly important as the dominant process for pumping noble gases, such as argon.

A third consequence of the collision of high-energy ions with the cathode is that cathode material is sputtered from the point of impact, producing an atomically clean layer on all surfaces that are within line of sight. Neutral molecules of reactive gases, such as nitrogen and oxygen, which collide with these sputtered films have a high probability (0.1 to 1) of being chemisorbed (gettered) on the first impact, and this process is the third primary pumping process in the sputter-ion pump. Note that this process is restricted to the pumping of chemically reactive gases, and is not effective for the noble gases, such as argon and helium, or for methane.

In a new pump, the trapping of gases by burial in the cathodes, the first process just described, is initially a major contributor to the overall pumping mechanisms. The ions are neutralized at the surface and penetrate a micron or less. The ion impact area is the central portion of the cathodes, and becomes more tightly focused as the pressure falls below the 10^{-6} torr range, covering 20% or less of the cathode. Consequently, the volume of the cathodes in which the atoms are buried is limited, and with one exception, hydrogen, the buried atoms are immobile. With continued pumping, the concentration of buried gas atoms in the impact area steadily increases, and thus these atoms become an increasing fraction of the sputtered species. With a constant gas influx, the rate of resputtering of the gas approaches its rate of burial, and at steady state there will be *no* net pumping by this process. The time taken to reach steady state is not well defined but is on the order of hours at a pressure in the 10^{-6} Torr range. Clearly, the high initial pumping speed in a pump with new cathodes will decrease toward a long-term equilibrium speed.

The process just described is responsible for an important problem of the diode sputter-ion pump. This is the memory effect. Sputtering from the cathodes re-

leases small amounts of previously pumped gases, accounting for the presence of impurities in a vacuum chamber that can be quite unexpected, especially when the previous pumping history of an ion pump is not known. Such impurities are, of course, more significant in a low-impurity UHV environment. Impurities will be most prevalent whenever a system is first pumping down, since the initial higher operating pressure results in a higher sputtering rate from the cathodes. If pumping is continued at relatively high pressure, the impurities rapidly disappear, being replaced by the new gas species that is being pumped. However, in the more typical operating procedure for a UHV system the gas influx is not sustained because, as the pressure falls steadily toward the ultimate operating level, the previously pumped species in the cathodes are not completely displaced, remaining as a continuing source of impurities.

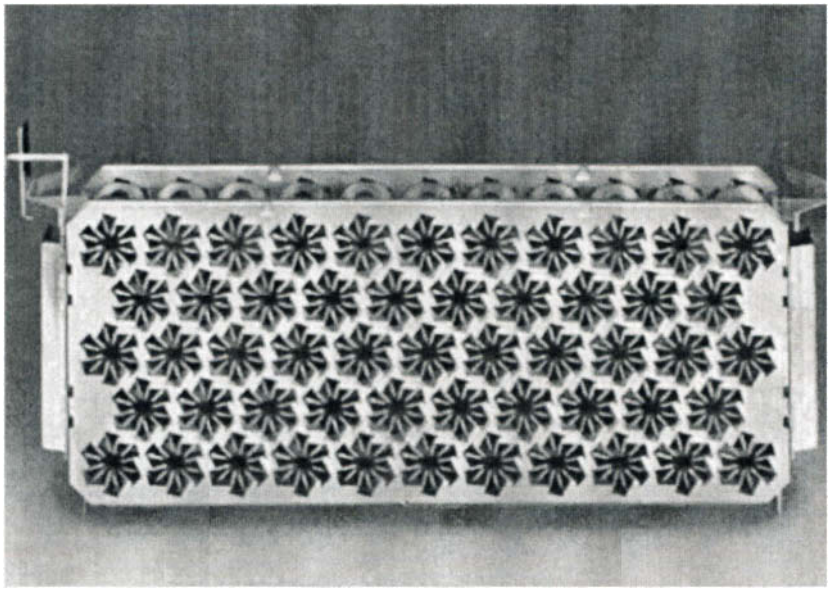
A second problem of the *diode* pump relates to the pumping of argon. This is an important gas because it is frequently used in sputtering processes in the vacuum chamber itself. In addition, the low level of argon in the atmosphere, ~7 torr, can occasionally lead to serious pumping problems when pumping air. Argon is not a chemically reactive gas and cannot be pumped by chemisorption (i.e., chemical reaction at the surface) on the sputtered titanium film. The sole mechanism for pumping is by burial in any surface, and the primary location in a *new* pump is in the cathodes. The fact that the pumping of argon, in a diode pump with titanium cathodes, is limited to burial, is reflected in a low pumping speed, typically 1%, for this gas as compared to the speed for reactive gases. Early experiences with the diode sputter-ion pump uncovered a problem when pumping argon, either alone or in mixtures [6].

After some critical amount of argon has been pumped, the pumping process sometimes fails dramatically, with a rapid release of argon into the system, which is then followed by a resumption of pumping. Once initiated, this regurgitation process is repeated at regular time intervals. Clearly, the purely physical containment of argon within the bulk of the cathodes sometimes reaches an unstable level. One successful solution to the instability of diode pumping is the substitution of one of the titanium cathodes by tantalum, producing the so-called differential ion, or DI pump[2] [7]; this substitution also results in an increase in the pumping speed for argon, up to ~20% of that for reactive gases. Clearly, if argon is an important gas load in a system, the use of a DI pump should be considered.

An explanation of the argon-pumping mechanism is given in an elegant paper by Jepsen [8], which suggests that efficient pumping of argon results when an argon ion, reflected at a cathode as a neutral atom, retains a significant fraction of the kinetic energy of the incident ion. The energetic neutral so produced is unaffected by the electrical and magnetic fields, and can bury itself in any surface of

[2]The differential ion pump was introduced by Ultek under the trademark Ultek D-I Pump. For convenience it will be referred to as the DI pump.

Fig. 2.



Anode and sputter-cathode of a StarCell™ triode sputter-ion pump. Each star-shaped section is coaxial with an individual anode cell. Reprinted with permission from Varian Vacuum Products.

the pump within line of sight of the ion impact point. In contrast to the central areas of the cathodes, where the initial ion impact occurs, most surfaces available for this burial are not subject to ion bombardment, and re-sputtering is not a problem. Such a burial is reinforced by subsequent deposition of sputtered titanium over the "grave site." The enhanced pumping, which results from the use of a tantalum cathode in the DI pump, is a consequence of the higher atomic mass ratio for tantalum/argon (181/40), as compared to that for titanium/argon (48/40), which ensures that a much larger fraction of the argon atoms is reflected, at significantly higher energies. The Jepsen hypothesis has been experimentally confirmed by Vaumeron and DeBiasio [9].

An important variant of the diode pump is the introduction of a second pair of electrodes between the anode cell and the pump wall, the so-called triode ion pump [6,10]. These additional electrodes are constructed of titanium to form an open structure, having an "egg crate", mesh, strip, or StarCell® geometry [11]. Figure 2 shows this additional electrode in the StarCell® geometry. These electrodes serve as the cathode of the pump, while the walls of the pump, usually stainless steel, are operated at anode potential (for safety reasons, the wall of the pump must always be at ground potential, which means that the polarity of the power supply for a diode pump is the reverse of that for a triode pump). The advantages of the triode geometry are twofold:

1. Ions strike the cathode grid at a low angle, which results in a higher sputtering yield. The cathode atoms are mostly sputtered in a forward direction to deposit on the wall of the pump. Because the pump wall is maintained at cathode potential, any ions that pass *through* the open grid structure of the cathode find themselves in a retarding potential field that causes them to reverse direction, back to the cathode; they do not have sufficient energy to reach the wall of the pump and thus cause no resputtering from that surface. As compared to a diode pump, the release of previously pumped gas (the memory effect) is negligible.

2. There is a large increase in the number of ions that are neutralized and scattered in the forward direction, to bury themselves in the pump walls, where they are subsequently further covered by sputtered titanium. The pumping speed for argon is high, ~30% of that for nitrogen.

The triode pumping configuration provides very effective, stable pumping for argon and for the other noble gases, so that it provides an alternative to the DI pump. A disadvantage is that, in many versions, the total quantity of titanium available for sputtering is substantially lower than in the corresponding diode pump. Consequently, the operating life is shorter, a factor to be considered if substantial throughputs of gas must be pumped for sustained periods, as in sputtering applications. The pumping of hydrogen is of particular interest, first because this is one of the most prevalent gases in a UHV system, and second, because of the problems that arise from the pumping characteristics for this gas. The pumping of hydrogen is very efficient at pressures above $\sim 10^{-8}$ torr [12], and essentially all hydrogen ions that strike the cathodes are pumped, as compared to ~30% for nitrogen ions [12,13]. The hydrogen first forms a solid solution in the titanium until the solubility limit is reached, at around 0.1 atom % [14]. Additional hydrogen forms a stable hydride, eventually reaching the composition TiH_2. When operating at pressures below 10^{-5} torr, the rate of hydrogen diffusion into titanium, even at room temperature, exceeds the rate of arrival of hydrogen ions. Consequently, the buried hydrogen atoms diffuse away from the point of impact, producing a more uniform distribution in the pump cathodes.[3]

Resputtering of hydrogen from the cathodes is therefore of minor importance until very large quantities have been pumped. For this reason, the pumping of hydrogen in the cathodes remains an important pumping mechanism throughout the life of the pump, in contrast to the limited amount of ion burial for all other species. Over the life of a pump, the capacity for pumping hydrogen is much greater than for other reactive gases, because it makes use of the entire volume of the cathode, whereas other reactive gases are predominantly pumped by chemisorption on the surface of sputtered material, most of which is sputtered from a relatively small area of each cathode. It must be noted that the pumping efficiency for hydrogen drops significantly as the pressure falls below 10^{-8} torr [14]. In this re-

[3] At high hydrogen pressures, the diffusion may be more complex. See the extended discussion by Welch [3].

gion, the use of a supplementary pump, such as a titanium sublimation pump or a non-evaporable getter pump [15], proves beneficial (see Section 2.7.4.).

Hydrogen is also pumped by chemisorption on sputtered titanium or tantalum, and once adsorbed, diffusion into the bulk metal will occur. However, the strength of the hydrogen–titanium bond is far less than that for all other reactive gases, including nitrogen and oxygen, and hydrogen chemisorption on sputtered titanium surfaces is greatly inhibited in the presence of these gases. A surface covered with a monolayer of nitrogen will chemisorb very little hydrogen, effectively blocking the diffusion of hydrogen into the bulk titanium [12]. For this reason, the pumping of hydrogen in the presence of other reactive gases depends predominantly on the burial of hydrogen ions in the titanium cathode. These factors account for a further complication in the pumping of hydrogen. In the unusual condition of pumping essentially *pure* hydrogen [12,13], the titanium cathode *surfaces* are progressively cleaned as the surface layer of gases such as nitrogen is sputtered away, permitting hydrogen *molecules* to first chemisorb on the atomically clean surface and then to diffuse, as atoms, into the titanium. This direct gettering of molecules increases the effective pumping speed by as much as a factor of 3, and the increase in speed is sustained when the power supply is turned off, slowly decreasing as the surface become covered with other chemisorbed gases. This enhanced pumping for hydrogen occurs automatically after a few hours pumping pure hydrogen at pressures in the high 10^{-6} torr range. It can also be achieved by pumping argon in the pump for a relatively short period of time [13]. Argon has a much higher sputtering yield than does hydrogen, cleaning the surface more rapidly, and additionally, argon does not adsorb, thus leaving an atomically clean surface.

The formation of titanium hydride by extensive pumping of hydrogen, causes substantial lattice expansion, with severe cracking [13] and warping of the cathodes, which may eventually cause an electrical short circuit to the anode. For hydrogen service, it has been customary to use much thicker titanium cathodes to increase the capacity for hydrogen, and to reinforce the cathodes to minimize warping. At the present time, the more common solution to pumping large quantities of hydrogen is to use a nonevaporable getter (NEG) as a supplementary pump, mounted either internal or external to the sputter-ion pump. NEGs are discussed in Section 2.7.4.4.

There remains yet another problem associated with hydrogen that has been pumped by a sputter-ion pump. The cathodes become loaded with hydrogen by pumping hydrogen or water. The hydrogen pumping is a reversible process, because of the low binding energy for this gas, which can be driven out from the cathodes by increasing the temperature. At 25°C titanium hydride has a dissociation pressure of the order of 10^{-12} torr, but this increases to ~ 1 torr at 400°C. Clearly, one can remove the hydrogen from the cathodes, in a controlled way, by a bakeout at 400°C, but this same process can inadvertently occur when starting a

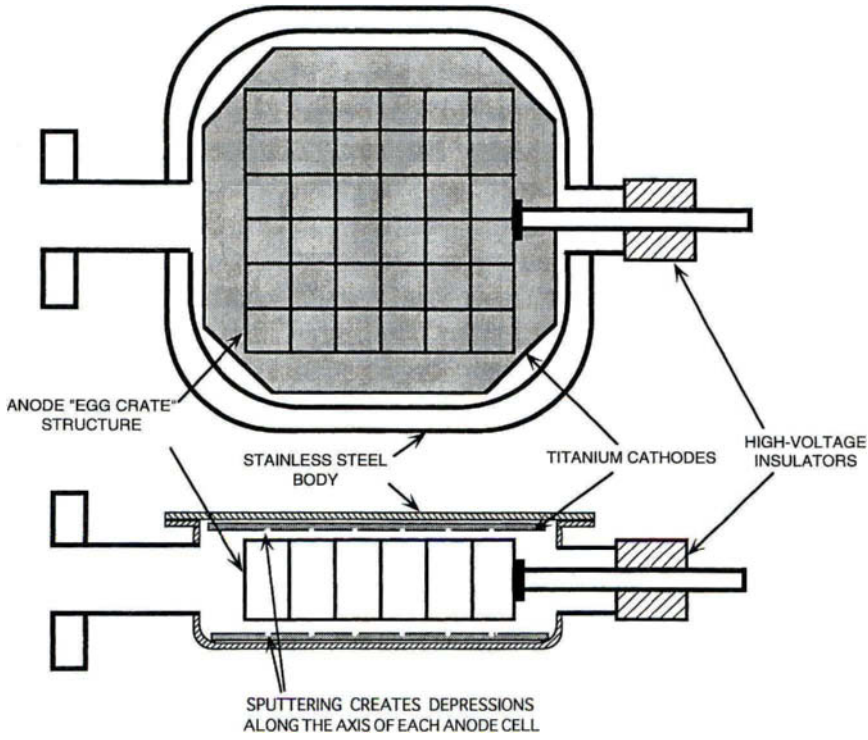
diode pump after exposure to atmospheric pressure. If a pump is started when the pressure is above $\sim 10^{-3}$ torr , a glow discharge is initiated, causing increased outgassing and the pressure will increase rapidly unless the roughing system has adequate pumping speed. The power dissipated in the pump at this stage of startup is very high; pump walls become dangerously hot to the touch, and internal temperatures of 500 and 900°C [16,17] have been measured. Such conditions initiate rapid evolution of hydrogen from the cathodes, further increasing the pressure, and the pumpdown "hangs up." Much grief caused during pumpdown can be avoided if the pump is roughed down well below 1×10^{-3} torr *before* turning on the power to the pump. In this pressure range a confined Penning discharge is immediately established, glow discharge induced outgassing from the walls is avoided, and sputter-ion pumping starts immediately. Note that the startup problems are reduced by using a triode pump, rather than a diode, because the potential distribution prevents ion bombardment of the wall, so there is no outgassing from this process [16].

One other class of gases to be considered, organic compounds, including the hydrocarbons, constitute an impurity of great concern in certain systems. Higher-molecular-weight organic compounds are dissociated, depositing inert carbon, while the hydrogen is pumped into the cathodes [18]. Methane, an impurity found in most UHV systems, is very effectively pumped, but the pumping ceases immediately when the power to the pump is turned off, whereas other chemically reactive gases continue to be gettered on the sputtered titanium deposits until the surface is fully covered by adsorbed gas. Methane is *not* chemisorbed by titanium or tantalum, and must be decomposed by electron bombardment before pumping can occur. It is not clear whether the decomposition is complete, with the deposition of carbon. The sudden prominence of methane as a principal residual gas, when a pump is turned off, can give a false impression of the rate of generation of this gas in the vacuum system. There is ample evidence that methane can be produced in the sputter-ion pump itself [19], or on a hot titanium sublimator [20]. However, it should not be assumed that methane is a universal impurity in sputter-ion pumped systems; an attempt to demonstrate its production in a sputter-ion pump, by operating over a wide range of hydrogen pressures, was not successful [21]. The only detectable source of methane in this particular system was the Bayard-Alpert ionization gauge, and the methane disappeared if the gauge was turned off.

2.7.3.2 The Construction of Sputter-Ion Pumps

The single-cell sputter-ion pump just described provides a nominal speed (typically measured for nitrogen, at a pressure of around 10^{-6} torr) of approximately 0.25 liter/sec, suitable only for very small pumping applications, such as a sealed-

Fig. 3.



Schematic of the first commercial sputter-ion pump.

off electronic device. In order to provide higher pumping speeds, the pump must be constructed using multiple cells. The first commercial pump was rated at 8 liter/sec and contained an "egg crate" anode structure of thirty-six 0.5-inch square cells, 0.5-inch deep [2]. A schematic drawing of this type of pump is shown in Figure 3. The pump was mounted between the poles of a massive Alnico horseshoe magnet, which had high magnetic fringing fields (~0.02 tesla at 2 cm from the pump casing), limiting its use in some applications. The newer version of this pump uses a compact ferrite magnet assembly having very low stray magnetic fields (down to the earth's magnetic field, 2 cm from the pump casing). For pumps having speeds up to several hundred liters per second, the anode cells are distributed in several pumping modules, crammed into the smallest permissible space, to make efficient use of the magnetic field. Because the volume available for the discharge affects the pumping speed, it must be as large as possible. The spacing between anode and cathode elements is kept small, while still providing sufficient conductance for gas flow between the mounting flange and each pumping

module to optimize the pumping speed [22]. Early pump designs commonly used Alnico magnets, allowing the pumps to be baked to the order of 400°C with the magnet in place; in some designs the magnets were mounted internally, and so could not be removed. Modern designs use ferrite magnets, which can be baked to ~350°C, and the external stray fields have been very substantially reduced. Samarium-cobalt magnets are also used, permitting a maximum bakeout temperature of ~250°C.

The large magnets required to provide fields in excess of 0.1 tesla, produce strong forces between the magnetic elements, and their disassembly should not be treated lightly. Most magnet structures are external to the pump vacuum envelope, although in some early designs the magnets were placed in hermetically sealed cans (to assure low outgassing rates) within the vacuum enclosure. The field strength of the magnets does not always conform to specification and should be checked if the pumping speeds appear low.

The anode structures in the original commercial pumps were made from titanium strip, but stainless steel appears to be universally used in later pumps. Cathodes of early diode pumps used $\frac{1}{8}$-inch titanium, but this has been replaced by $\frac{1}{16}$-inch or thinner sheet. In models specifically designed for pumping hydrogen the original thickness is retained. In the DI pump, the tantalum cathode is usually 0.040 inch thick. Titanium cathodes in triode pumps have appeared in many forms, from a simple mesh, through parallel narrow strips, to the spark-machined StarCell™ electrodes.

In many triodes, the amount of titanium that can be sputtered, before the integrity of the cathode is impaired, is less than is the case for diode cathode elements. It must, however, be emphasized that the life of even the heaviest diode elements is limited and it is not uncommon to find the cathode sputtered through in the center of the discharge impact area. This occurs in pumps that have been operated well above the UHV pressure range for sustained periods. If sputtering is continued beyond cathode failure, a small leak will eventually develop through the pump wall.

The insulators used in a pump must be designed so as to be protected from short-circuit by sputtered cathode metallic films. In general, this requires only line-of-sight shielding for the low-pressure operation for which these pumps are intended.

2.7.3.3 The Pumping Speed of Sputter-Ion Pumps

The variation in the pumping speed of a sputter-ion pump is a very complex function, depending on gas species, pressure, and time. Typical pumping speed specifications for a number of gases are given in Table 1. Note the significant differences between (1) the standard and DI versions of the diode pumps and (2) the

Table I

Pumping Speeds (in liters/sec) for Diode and Triode Sputter-ion Pumps Using Titanium Cathodes, and for a Diode Pump Using Titanium and Tantalum Cathodes ("differential" ion pump).

Gas	Diode	Triode
H_2	270	210
N_2	100	100
O_2	60	60
H_2O	100	100
CH_4	90	90
Ar	1 (20)[a]	30
He	10 (25)[a]	30

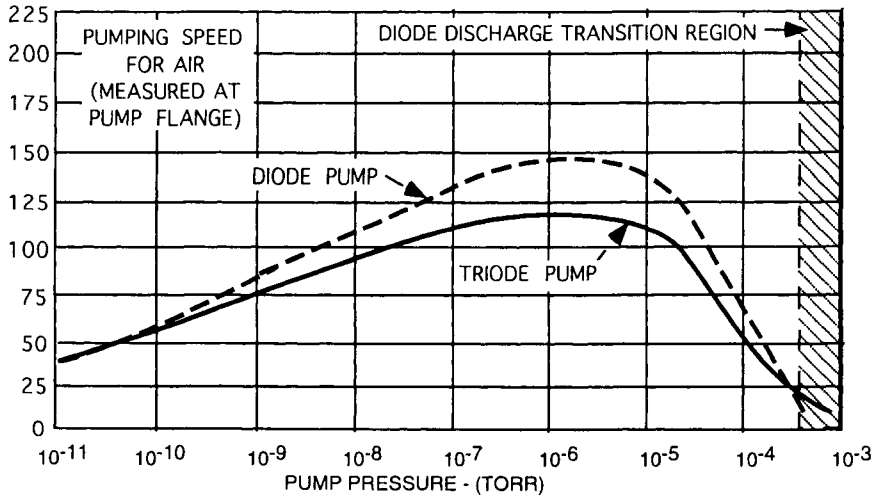
[a]Differential ion pump using titanium and tantalum cathodes.

triode pumps. Pumping speeds are often measured at pressures where the speed is a maximum, in the 10^{-5} to 10^{-7} torr range.

We have already noted that the speed of a new (i.e., unused) pump will include a large contribution (30 to 40%) due to ion burial in the cathodes, and that this contribution will decrease as the impact area becomes saturated with pumped gas. In recent years, pump specifications give an indication of the potential magnitude of this drop, which will occur within hours of starting the pump, if the gas load is sustained at a high level, say in the 10^{-6} torr pressure range, or may continue as a slow decrease extending over years if the pumping is predominantly in the UHV range.

Of even greater importance is the significant decrease in the pumping speed of a sputter-ion pump, which occurs as the pressure falls. It results from a decrease in the electron density sustained in the anode volume of each pump cell. In the early pump designs, with anode cells the order of 0.5-inch diameter or 0.5-inch square, the pumping speed, measured for nitrogen at 1×10^{-8} torr, was only 10% of that at 10^{-5} torr. Substitution of anode cells with diameters exceeding 0.75-inch, has greatly improved the low-pressure performance of many pumps; the speed falls off fairly steadily with decreasing pressure, to around 50% at 10^{-10} torr. Figure 4 shows typical pumping speed curves for both diode and triode pumps. When a common pump body and magnet assembly is used for both a diode and a triode configuration, the triode must use a shorter-length anode, so as to leave space for inserting the additional sputter-cathode structure. This results in a lower pumping speed for each individual cell; in some cases, compensation is introduced by decreasing the diameter of the anode cells in the triode elements so as to accommodate a greater number of cells; this change causes some decrease in the pumping speed at very low pressures, compared to the pumping speed of the corresponding diode pump.

Fig. 4.



The pumping speed (arbitrary units) of diode and triode sputter-ion pumps as a function of pressure (approximate). Reprinted with permission from Varian Vacuum Products.

A word of caution is necessary in discussing pumping speeds. In general these are measured under steady-state conditions when pumping nitrogen. The individual pumping speeds for *most* gases may have a different magnitude, but follow the same pattern as for nitrogen, when measured at steady state. We have already seen that pure hydrogen is a very different case, falling off more rapidly than other gases at low pressures, and having a possibility of much higher speeds at high pressure, if gettering of molecules occurs. But there are also time dependent changes in speed for all gases when the rate of gas flow is suddenly changed, and there are also synergistic effects with some mixtures of gases. The pumping speed of a sputter-ion pump for individual gases is not clearly defined, in sharp contrast to the performance of turbomolecular, cryo-, and diffusion pumps. If the operation of a sputter-ion pumped system demands a well-defined pumping speed for each gas that may be used, it is essential to use a conductance-limited pump so that the effective speed is determined by the size of the conductance rather than by the pump itself.

2.7.3.4 The Operational Life of Sputter-Ion Pumps

The nominal operating life of a sputter-ion pump is limited by the quantity of titanium or tantalum that can be sputtered before the cathodes are destroyed. In a diode pump, this occurs when the cathodes are sputtered through in the ion impact areas, which are located along the axis of each anode cell. In the triode it is reached when the cathodes begin to disintegrate, often causing an electrical short

in the pump. The rate of sputtering is proportional to the operating pressure, giving, for example, a nominal life of up to 60,000 hours at 10^{-6} torr. This implies a life of only 600 hours at 10^{-4} torr; clearly the sustained operation of a pump in this pressure range is impractical. At the other end of the scale, one can expect extremely long life for operation at 10^{-10} torr. Factors that lead to reduced life include the flaking of sputtered deposits off the walls of the pump, leading to electrical shortcircuits, and the growth of whiskers on the electrodes, leading to localized discharges when the whiskers serve as field-emission electron sources. Flaking occurs because of poor bonding of the titanium film to the pump walls, or to underlying titanium layers. Such failure rarely occurs in a pump that is continuously operated in a UHV condition, but is common in pumps that are frequently cycled to atmospheric pressure. Exposure to the atmosphere allows chemisorption on the titanium surface, followed by physical adsorption, particularly of water vapor. On subsequent pumpdown, the chemisorbed layers are never removed, and the physisorbed water may not be completely removed, in the absence of a bakeout. The contaminated surface of the sputtered deposit provides only weak bonding for subsequent titanium deposits, leading to flaking.

2.7.3.5 Pressure Measurement Using Sputter-Ion Pumps

In an ion pump, the ion current provides a continuous measure of the pressure in the pump itself, and therefore serves to monitor the overall performance of the system. Accurate pressure indications are not always obtained, especially at the low end of the range. This is particularly a problem in triode sputter-ion pumps, where field emission currents are large enough to swamp the ion currents that provide the pressure indication. Pierini and Dolcino [23] believe that this field emission is a consequence of the sharp edges left on the sputter cathodes by conventional manufacturing techniques, and they have shown that cathodes fabricated using a spark erosion technique (Figure 2) have very low leakage currents; the question of how well this improved performance persists throughout the life of a pump was not addressed. Although the measurement of pressure by an ion pump is useful, it should be emphasized that in many cases it is not an adequate substitute for measurement of pressure in the vacuum chamber itself using an ion gauge or, preferably, a partial pressure analyser.

2.7.3.6 Practical Applications of Sputter-Ion Pumps

Sputter-ion pumps are ideally suited for use on ultra-high-vacuum chambers where cleanliness is paramount and where very low pressures must be maintained over long periods. Under these conditions pumps give high reliability and stable

operation, with long life. Bakeout of a complete system, including the ion pump, assures rapid degassing, and the achievement of the system ultimate pressure within one or two days (depending on the bakeout temperature). Of primary importance is the fact that system operation is virtually fail safe with a system isolated from the atmospherc; for example, an all-metal, leak-tight system, left without pump operation in excess of one year, pumped down to an ultimate in the low 10^{-9} torr range within hours of turning the sputter-ion pump back on.

At low pressures, sputter-ion pumps require very little power to operate and their use on portable systems, using a battery-operated power supply, is entirely feasible. Sputter-ion pumps also find great application in particle accelerators and in electron storage rings where they must reliably maintain pressures of 10^{-10} torr or less in vacuum chambers, which can be on the order of miles in length. In some cases the pump elements are inserted into the vacuum chamber and operate in the high magnetic fields of the device itself. Distributed pumping systems are particularly appropriate where the vacuum chamber is a very long, small-diameter tube, where the conductance of the tube is so low that the pumping points must be fairly close together.

Sputter-ion pumps are *not* well suited for use on systems that operate for long periods with a high throughput of gas at pressurcs in the 10^{-5} torr range or higher, or where there is a substantial flux of organics. In such instances the use of turbomolecular or cryopump would be a more appropriate choice.

2.7.3.7 Rough Pumping Sputter-Ion Pumps

In any UHV system, successful operation depends on assembling a relatively clean vacuum system, and maintaining that cleanliness during the initial pumpdown of the system prior to starting the sputter-ion pump. To facilitate easy starting, it is recommended that diode sputter-ion pumps be routinely roughed out to a pressure of 1×10^{-3} torr, or less, before turning on the power. This assures that the discharge in the pump will initiate in the confined Penning mode rather than as an unconfined glow discharge. The advantage is that the normal pumping mechanisms are immediately effective, assuring a rapid startup, and permitting immediate isolation from the roughing system. When a glow discharge is initiated by starting the pump at higher pressures, it provides little pumping, but a great deal of outgassing occurs, and the glow discharge can spread far beyond the pump volume. This occurs because the wall of a diode pump, and hence the walls of the rest of any metal vacuum chamber on which it is mounted, serve as a cathode and are subjected to ion bombardment. Substantial outgassing develops, which has some advantage in cleaning up the pump, but the outgassing often overwhelms the roughing pump, and as the pressure rises the glow discharge intensifies. The increased power dissipated in the pump causes the sputter-ion pump cathode tem-

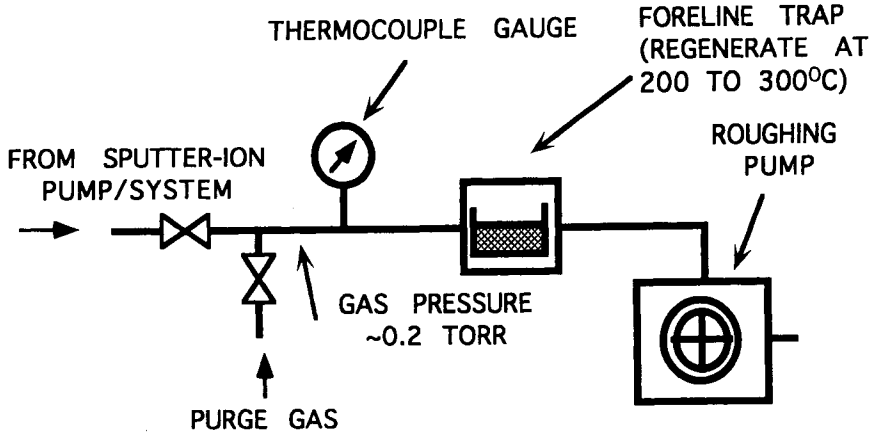
perature to increase rapidly. Cathode temperatures above 400°C have been measured, causing dissociation of the titanium hydride, the rapid release of hydrogen putting further demands on the roughing system. The quantities of hydrogen involved can be large in pumps that have been in service for a long time, especially those that have been cycled frequently to atmospheric pressure (with consequent exposure to moisture). The starting pressure for a triode pump is not so critical as for the diode pump [16], for any glow discharge is confined to the pump body; the walls of the pump operate at anode potential and are not subjected to ion bombardment, thus limiting the outgassing.

From a practical viewpoint, if the roughing pressure starts to increase when the power to a diode pump is turned on, do not wait more than a minute or two before shutting the power off, or the body of the pump will become exceedingly hot. With the pump power off, continue roughing until the pump has cooled, and only then try to restart the pump, after the pressure has fallen below that which existed when the pump was originally started.

An oil-sealed mechanical pump can readily achieve the pressure necessary to start an ion pump, but at 1×10^{-3} torr, the foreline will normally be in the free molecular flow region, allowing backstreaming of pump oil into the UHV system. Use of an efficient foreline trap will minimize oil contamination, provided that the trap is well maintained, and reactivated as required; unfortunately, it is difficult to establish any simple way of determining when reactivation is necessary short of using a mass spectrometer to look for the presence of oil on the system side of the trap. The useful life of a trap depends on too many factors, and a conservative approach is advisable. To be quite sure of the most effective trapping, consider reactivating the trap immediately before each use; this is not particularly onerous, because many UHV systems are only rarely cycled up to atmospheric pressure. An interesting procedure has been suggested by Santeler [24] using a system similar to that shown in Figure 5. Whenever the mechanical pump is started, the foreline is maintained at a pressure in the viscous flow region[4] using a gas purge, such as dry nitrogen, introduced on the low-pressure side of a Molecular Sieve™ (synthetic zeolite) or activated-alumina-filled foreline trap. In this condition, oil backstreaming from the mechanical pump is suppressed, and the foreline trap remains clean. The sputter-ion pump is initially roughed out with the gas purge flowing, and once equilibrium is reached, the gas purge is turned off, allowing the mechanical pump to pump to its base pressure, with the protection of a virtually uncontaminated foreline trap. After the sputter-ion pump has been started, it is isolated from the roughing pump, the gas purge is reinstated, once again suppressing oil backstreaming, and keeping the foreline trap clean. The mechanical pump can then be vented, turned off, and left at atmospheric pressure until it is again required. Using this procedure, the period between successive reac-

[4] Viscous flow will dominate so long as $P > 0.45/D$, where P is the pressure in torr, and D is the internal diameter of the foreline, in centimeters.

Fig. 5.



The use of a foreline trap, with a gas purge system, for roughing down a sputter-ion pump.

tivations of the foreline trap can be very long indeed. Note also that a gas purge should always be used during the bakeout of the foreline trap for reactivation, thus avoiding any danger of oil contamination of the low-pressure side of the trap.

The so-called dry pumps, which operate without the use of oil, provide an alternative to the oil-sealed mechanical pump. Scroll pumps are now available with ultimate pressures in the 10^{-3} torr region, but other types of "dry" pumps must be used in combination — for example, a diaphragm pump backing a molecular drag pump — to achieve satisfactory low pressures. In multiple ion pump installations, a properly operated turbomolecular pump system is frequently used for roughing. However, this is not often economically feasible in facilities where only one or two ion pumps are used. In these cases, the classic ultra-clean roughing technique is to use a sorption pump.

The sorption pump is simply a vacuum vessel containing a very-large-surface-area adsorbent that can be cooled to liquid nitrogen temperature. On cooling, the adsorbent physically adsorbs a large quantity of atmospheric gas, providing clean pumping for the sputter-ion pump and vacuum system. The adsorbent commonly used is a synthetic zeolite, or Molecular Sieve™, the atomic structure of which contains uniform internal pores having an internal diameter on the order of angstroms. An activated alumina adsorbent can also be used. A typical pump may contain a surface area of several hundred acres, and a pumping capacity on the order of 10^5 torr-liter; a single pump can therefore reduce the pressure in a 100-liter system below 0.1 torr. Adsorption coverage is a function of temperature, and of the partial pressure of the particular gas being pumped. It is a reversible process, and a pump is reactivated simply by warming to ambient temperature, venting the desorbed gas to the atmosphere. The entire procedure is repeated as required. If faster roughing down is important, several precooled pumps can be installed so

as to avoid the delay introduced by the process of regeneration. It must be emphasized that the saturated pump contains sufficient gas to develop a pressure of many atmospheres in the pump volume on warmup, and it is mandatory to include a pressure-relief valve on the pump to avoid overpressure during either a routine, or accidental regeneration.

Although the operating principle of the sorption pump is very simple, a specific mode of operation is necessary to reach pressures below 1×10^{-3} torr. The reason is that, although most gases in the atmosphere are strongly adsorbed on surfaces at liquid nitrogen temperature, giving equilibrium pressures well below the 10^{-4} torr range, three common gases—helium, hydrogen, and neon—are not effectively physically adsorbed. These gases are present in the atmosphere at a total pressure of $\sim 2 \times 10^{-2}$ torr, and even if a regenerated sorption pump is used to pump the chamber to an equilibrium condition, well over 50% of these three gases will remain in the gas phase. Consequently, the ultimate pressure will remain in the high 10^{-3} torr region, too high to comfortably start a well-used diode pump. Fortunately a simple procedure can remove most of these gases. When the first sorption pump is opened to a vacuum chamber at atmospheric pressure, the gas flow from the chamber to the pump will be in the viscous flow regime, and typically will continue in that state until the pressure in the roughing line falls to a few tenths of a torr. In the viscous flow regime, *all* components of the atmosphere are swept along into the pump, and the viscous flow of gas will prevent any untrapped gas from diffusing upstream, and escaping from the pump. In short, any neon, helium, and hydrogen in the pump volume will be trapped there, *so long as the pump is valved off while viscous flow still dominates.* If the first round of sorption pumping is used to exhaust the chamber to only 7.6 torr, then roughly 7.6/760, or 1% of the untrapped gases will be left in the vacuum chamber, equivalent to a total pressure of these gases of $\sim 2 \times 10^{-4}$ torr. Successive stages of sorption pumping can then be used to reduce the pressure of the remaining *adsorbable* gases to the required level.

Sorption pumping is an inherently clean roughing technique. It is relatively cumbersome, but the speed of a typical pump is the order of 5 liter/sec, i.e., close to that of a 10-cfm mechanical pump. In practice, an ultra-high-vacuum system is rarely vented to the atmosphere, and a cumbersome procedure can usually be tolerated if it occurs only at intervals of several months! If any zeolite powder is blown into the vacuum chamber, it can play havoc with the sealing surfaces of any all-metal valves, but this problem can be avoided with appropriate handling of the pumps.

2.7.3.8 Starting Up and Running Sputter-Ion Pumps

For most UHV systems, bakeout of the vacuum chamber is performed using the sputter-ion pump as the sole pump, and no attempt is made to fully outgas the

pump itself. The pump runs fairly warm during starting, providing some outgassing, and heat from baking out the rest of the system results in some further increase in the pump temperature. If the lowest possible system pressure is required in a reasonably short period of time, the entire pump and vacuum chamber can be baked simultaneously, using a supplementary pumping system whose cleanliness mirrors that desired in the UHV system. If a bakeout requires a temperature above 250°C, it may be necessary to remove the magnets and high-voltage power cable. In the discussion that follows, we assume that only the vacuum chamber will be baked, while using the sputter-ion pump as the pump.

Once a sputter-ion pump has been roughed down, the pump can be turned on. At the highest operating pressure, $\sim 1 \times 10^{-3}$ torr, the electrical impedance of the pump is very low and exceedingly large currents would flow if the power supply maintained a voltage in the kilovolt range. This could result in serious damage to the pump. The power supplies are deliberately designed to limit the available power. The voltage and current are severely limited, and consequently, the pumping speed is initially very low. If the pressure begins to fall when the ion pump is turned on, the roughing pump can be immediately valved off, proceeding slowly enough to confirm that the pressure continues to fall as the speed of the roughing pump is reduced. As pressure continues to fall, the sputter-ion pump voltage will rise steadily, reaching the maximum, design level when the pressure is around 10^{-6} torr. With a pump in good condition, the pumpdown will proceed smoothly, at a rate determined by the gas load from the vacuum system. With a well-used pump, the sound of momentary electrical arcing may be detected, accompanied by a groan from the power supply, and a sudden pressure burst. Arcs may be initiated at sharp edges of flakes of titanium, or at whiskers on the cathodes. The location of the discharges can be readily seen if a window provides a view of the pump electrodes. In many cases, arcing episodes gradually diminish as the pressure falls, and the pumps perform satisfactorily at the system operating pressure, but the arcs give fair warning that the pump should be reconditioned, or replaced at a convenient opportunity. On other occasions, a flake of titanium may bridge the electrodes in a pump, either partially or totally shorting out the pump. A smart blow on the side of the pump can sometimes resolve the problem satisfactorily, at least on a temporary basis, and operation can continue without problem until the next time the system is pumped down following venting to atmosphere.

It is appropriate in the present context to note the advantage to be gained by always venting a vacuum chamber to atmosphere using a *dry*, gas, nitrogen being the most common choice. This minimizes the exposure of the system to atmospheric moisture, which will tenaciously adsorb on the system walls to a depth of one or more monolayers, the actual thickness depending on the humidity of the air. Water is always the last gas to be removed from a chamber during pumpdown and is normally the cause of slow attainment of the ultimate pressure in a vacuum system. To maximize the advantage gained by nitrogen venting, minimize the time that the chamber is exposed to the atmosphere.

Once the pressure in the vacuum system falls into the 10^{-6} torr range the vacuum chamber can be baked out. Few problems are associated with this process, assuming that the system is free from leaks. The major gases to be removed are water, carbon monoxide, and hydrogen. Organics, such as methane may well be generated in the pump, but the rate of production is usually small, and should not affect the proper operation of the system. O-rings should not normally be used to seal against atmosphere, and those used as internal valve seals should be Viton, Kalrez®, or polyimide, which are thermally stable and permit bakeout to approximately 200, 250, and 300°C, respectively.

The type of sputter-ion pump selected for a new vacuum system will depend on the major gases which will be present. The pumping speeds of three types of pump, for common gases, is given in Table 1. The size of pump must be selected with reference to the throughput that is expected for each gas and the pumping speed for that gas at the operating pressure. The major common gases of concern are argon and hydrogen. Experience shows that unstable pumping may develop with pure argon, or with an argon-containing mixture, for example, air. In some studies, argon instabilities occurred periodically whenever a particular quantity of gas had been pumped. Once instabilities developed, the time between successive instabilities was short at high throughput, but increased on lowering the gas flow. If argon is to be consistently used in a system, it is prudent to choose the most reliable pump type. The DI type of diode pump or the StarCell™ triode are probably the best choices for pumping substantial quantities of argon, because the quantity of titanium in these pump types is considerably greater than in a conventional triode pump, and therefore better operating life can be expected.

Sputter-ion pumps are very efficient for pumping hydrogen, with the exception of very low pressure operation where the hydrogen pumping speed can drop off even more precipitously than for other gases [14]. Serious problems occur when pumping so much hydrogen that distortion of the cathodes occurs. For example, at sustained pressures in the 1×10^{-5} torr range. Pumps have been built using extra-thick titanium cathodes, reinforced to resist warping and fitted with water cooling, to avoid problems of dissociation of the titanium hydride that can occur if the cathodes run too hot in a heavily loaded pump. One additional approach to limit pump overheating at high hydrogen loads is to operate the pump at a lower voltage (<4.8 kv) [25]. If standard pumps must be used, it should be noted that the capacity for pumping hydrogen in a tantalum cathode is substantially less than in a titanium cathode, so the conventional titanium cathode diode pump is likely to perform better than the DI pump when *large* throughputs of hydrogen are anticipated. However, getters are now available with outstanding hydrogen-pumping capacity, and they should certainly be considered to supplement a sputter-ion pump; such getters are discussed in Section 2.7.4.4.

Although the sputter-ion pump is an effective general-purpose device, it has a relatively high capital cost and is often combined with a relatively low-cost getter

pump, most commonly a titanium sublimation pump. Sublimation pumps provide very high pumping speeds for reactive gases, including for example, oxygen, nitrogen, hydrogen, carbon monoxide, and water vapor. They are particularly useful for sustained operation in the UHV region. The sublimation pump can be started after the vacuum chamber has reached a pressure in the 10^{-6} torr range, or less, and can readily provide sufficient speed to reduce the system ultimate pressure by a decade or more. Alternately, a sublimation pump can provide reserve pumping capacity when a deliberate influx of a reactive gas is to be introduced into the vacuum system. The characteristics of sublimation pumps are described in detail in Section 2.7.4.

For any process that requires the maintenance of very low pressures and low organic impurity levels, a sputter-ion pump can give excellent performance and long operating life. When samples must be transferred between atmosphere and the system, on a regular basis, the use of a load-lock chamber for the transfer avoids any venting of the main vacuum chamber, permitting one to retain most of the advantages inherent in these pumps.

If a sputter-ion pumped system must be vented to atmosphere on a regular basis, it is advantageous to provide a valve to isolate the pump from the chamber, so that it can be left evacuated, and even running, while the chamber is vented using dry nitrogen gas. In a further refinement of this approach, the vented vacuum chamber is continuously purged using dry nitrogen, to minimize air intrusion to the greatest extent possible; any components or samples added to the system are pre-baked and kept meticulously clean. Such stringent measures can result in a far more rapid pumpdown of the vacuum chamber, limiting the extent of exposure to the atmosphere and the necessity for a subsequent severe bakeout.

An operating sputter-ion pump can generate a copious source of electrons and ions during the initial pumpdown, inducing unexpected system behavior, for example, erroneous ion gauge pressure measurements. At lower pressures the pump can introduce electrons, ultraviolet radiation, and X-rays into the vacuum chamber. Fortunately, if the pump is mounted or shielded so that there is no direct line of sight into the system, these problems become negligible. A further perturbation introduced by the pump is stray magnetic fields; the use of better-confined magnet assemblies has reduced this problem to a relatively low level.

The power supply of a pump is designed for a particular size of pump, and is designed to limit the power output under the high-load conditions that prevail during startup. It is unwise to substitute either a larger or smaller power supply, without consulting with the manufacturer. The dissipation of too much energy during the early stages of pumpdown could lead to destructive overload, and the use of too low a voltage will reduce the pumping speed of the pump.

Take adequate precautions to ensure safe operation of a sputter-ion pump, because of the potentially lethal currents and voltages involved. A damaged power cable should never be used, and it is strongly recommended that an independent

electrical ground be provided for the pump body (and for the rest of the vacuum chamber, if there is any insulating barrier between the pump and chamber). It should be recognized that, when a diode pump is first started, the general glow discharge may spread into the rest of the vacuum system. Electrons can travel remarkable distances and if there is any wire or structure in the chamber that is attached to an electrical feedthrough, that feedthrough should never be left electrically "floating." It can accumulate an electrical charge quite high enough to provide a nasty electrical shock when touched. *Ground all unconnected feedthroughs!*

2.7.3.9 Troubleshooting and Maintenance of Sputter-Ion Pumps

Since there are no moving parts in a sputter-ion pump, there are only a few common problem areas. Failure of the power supply is relatively uncommon. Most supplies have automatic overload protection when the power drawn by the pump exceeds the design value, as for example would result if a pressure surge above 10^{-3} torr initiated a general high-current glow discharge. When the pump is started, this overload protection must be turned off, but should be restored as soon as the pressure begins to fall, say into the 10^{-4} torr range. In a pump that is marginally operational, momentary arcs within the pump could trigger the protective circuit, which should therefore be left off when these problems arise, but the pump must be watched carefully through this startup phase, until arcing ceases once the pressure has fallen far enough.

After long use, many pumps develop whiskers, which serve as field emission sources for electrons in the presence of high electrical fields. The conditions for the formation of these whiskers are not well defined, but they produce localized discharges, giving higher than expected pump currents. The presence of discharges can be inferred by comparing the pressure in the vacuum chamber, using an ion gauge, with that indicated by the pump current. When the pump is operating normally, its indicated pressure will be lower than that in the system, but if localized discharges have developed, the reverse will be true. Discharges do not provide significant pumping, and if the current draw is high, may result in the voltage applied to the pump falling below its design value.

Two procedures have been used to remove whiskers in an ion pump. The first is known as "Hi-Potting," a procedure widely used in the conditioning of high-voltage vacuum devices. It consists of the momentary application of a current-limited potential on the order of 20 kv to the pump, while under vacuum. This procedure can remove whiskers by ion bombardment. A second procedure is to run the ion pump at an oxygen pressure in the 10^{-5} to 10^{-6} torr range in the hope of destroying the whiskers by chemical reaction. The procedures are not always

successful and should be used as a last resort if replacement of the pump is not possible.

The problems that arise from the flaking of thick titanium deposits from the pump walls have already been alluded to. Denison [26] found that flaking problems in *sublimation* pumps could be greatly reduced by using a 300°C bakeout after each exposure to atmosphere, presumably removing adsorbed water. A bakeout is not feasible with a sputter-ion pump unless a supplementary pumping system can be employed, but the precautions just outlined, of either leaving the pump valved off during system venting, or of venting the pump with dry nitrogen, should reduce flaking problems.

When whiskers or flakes are present, they can produce an electrical short. Also, if the pump has been run at high pressures, there may be titanium deposits on the high-voltage feedthrough or anode insulators. The pumping efficiency will usually be reduced if electrical leakage problems cause excessive current draw from the power supply. In extreme cases, this may result in an obviously low pump voltage, while in other cases the pressure indication from the pump current will be higher than normal. The problem may be identified by disconnecting the power supply and measuring the electrical resistance from the high-voltage terminal to ground. Check the resistance with both polarities of the meter, since it may be low in one direction and high in the other. Note that exposure of the pump to atmospheric pressure will often increase the resistance dramatically, so resistance checks are best made while the pump is still under vacuum.

A series of sharp knocks on the side of the pump will sometimes displace a short, and whiskers can be removed by the procedures just described. In many cases, however, the pump must be either replaced or cleaned. The former is the preferred route assuming that time and money are available, but an on-site cleanup is possible *provided adequate precautions are taken.* Details of recommended procedures can be found in old sputter-ion pump manuals dating from the 1960s. The process starts with disassembly, followed by mechanically cleaning the pump walls. At this stage some parts, such as insulators, should be replaced. If the cathodes are seriously eroded, they should be replaced. Take care when removing any sputtered titanium films; they may contain a great deal of unreacted metal, which may glow or burn when scraped free from the wall. Chemical etching of pump parts is commonly required, involving the use of very aggressive etching baths, including nitric and hydrofluoric acids. All necessary precautions to protect the eyes and skin are essential in these procedures, which are not recommended for the uninitiated.

Finally, it is important to be aware that sputter-ion pumping falls off to negligible values very rapidly at pressures above 10^{-3} torr. Consequently, if the pressure in a system is measured using the ion pump current, very low readings may indicate *either* a very low *or* a very high pressure.

2.7.4 GETTER PUMPS

The term *gettering* has its origin in the techniques used in early radio tubes to obtain and maintain an acceptable degree of vacuum for the operation of the tubes. These devices had to be manufactured on a time scale that did not permit adequate outgassing, and so a reactive metal surface was included in the tube to react with any residual gases, essentially getting them out of the way. Commonly used getters included heated filaments of reactive metals such as titanium or zirconium, and evaporated films of the alkaline metals, particularly barium. The one thing that all these "getters" have in common is their rapid reaction to form a stable chemical compound with common residual gases, such as oxygen, nitrogen, and water vapor. It is interesting to note that the much smaller residual amounts of argon and helium, which cannot be pumped by chemical reaction, are undoubtedly pumped as ions, after ionization by electron bombardment. Such processes were widely studied in attempts to explain the pumping of these gases in ionization gauges [27]. In current vacuum technology, barium gettering remains a crucial step in the production of TV picture tubes and cathode-ray tubes, and evaporated films of titanium find wide use in UHV pumping systems; bulk or nonevaporable getters, commonly known as NEGs, are finding applications in general vacuum use and in sealed-off devices. Note, however, that getters are supplementary pumps that can only remove chemically reactive gases; they must be combined with another pump that can handle any influx of noble gases, or of methane and other chemically stable organic compounds.

2.7.4.1 Sublimation Pumps

Sublimation pumps are devices that deposit a film of a reactive metal on a selected substrate either continuously or periodically. Titanium is the preferred metal and is deposited from a high-temperature source. The titanium sublimes from the source, going directly from the solid to the vapor phase. This fortunate circumstance permits the direct electrical heating of a titanium wire in order to provide an appropriate sublimation rate, without melting the wire.

The removal of gas by a sublimation pump involves a process of chemisorption on a titanium surface. Most gases are sorbed *only* on the surface (the one exception being hydrogen). As the surface coverage increases, the rate of pumping decreases, becoming negligible once the surface is fully covered by a single adsorbed layer. The probability of a reactive gas atom or molecule being adsorbed on first impact, *on an atomically clean titanium surface*, is between 10 and 100%,

depending on the particular gas. As the surface becomes covered, this probability drops by one or two orders of magnitude! When a vacuum chamber is in the ultra-high vacuum range, say 1×10^{-9} torr, it will take two to three hours for a monolayer to form, so that a brief period of sublimation each hour will replenish the clean titanium layer, maintaining a high average rate of pumping. The time necessary to replenish the film depends on the rate of sublimation, and on the area to be covered. A larger surface area translates into higher pumping speeds, but requires a larger total quantity of titanium to be deposited each time. If a sublimation pump is operated at higher pressures, the periods between operation of the sublimator must be decreased, and at the highest practical rate, the sublimator must operate continuously, so as to maintain the rate of sublimation slightly higher than the rate at which the gas reacts with the titanium. It is important to note that, even at the highest pressures at which a sublimator is likely to be operated, say 1×10^{-5} torr, the mean free path of a nitrogen molecule is around 500 cm, so that the probability of nitrogen reacting with titanium in the gas phase is quite negligible.

There is an upper-pressure limit of around 10^{-4} torr for the effective operation of a sublimator. On exposure to the atmosphere, a sublimator surface is almost instantaneously covered with a strongly bound, chemisorbed layer, mainly of oxygen and nitrogen. This layer is not desorbed at the operating temperature but diffuses into the bulk titanium, leaving a clean surface on the titanium filament on which further chemisorption can occur. When the chemisorbed layer is present, sublimation of titanium is strongly inhibited. Consequently, if the filament is heated when the system pressure is too high, the chemisorption of gas on the hot surface will exceed the rate of diffusion into the bulk, thus maintaining a full surface layer [28]. A similar phenomenon has been reported for yttrium sublimation [29]. Because sublimation is suppressed, a layer of titanium is not deposited on the system walls, and essentially all the gettering of gas must occur on the small area of the hot sublimator. Consequently, the effective pumping speed remains very low. In contrast, if the system pressure is in the low 10^{-5} or 10^{-6} torr region, the rate of diffusion into the filament will exceed the rate of gettering on its surface, with the result that sublimation proceeds. The surface area of deposited titanium soon far exceeds that of the hot filament, and the bulk of the gettering is transferred from the hot-filament surface to the deposited titanium surface. Gettering *always* occurs on the hot filament, but at lower operating pressures, it is too small to sustain significant surface coverage on the sublimator, and proper operation is assured.

A typical titanium sublimation temperature is around 1550°C, giving a vapor pressure of about 1×10^{-4} torr and a rate of sublimation of around 0.05 g/hour for every square centimeter area of hot titanium; this rate is very roughly equivalent to the pumping (or gettering) capacity required to handle a throughput of

1×10^{-3} torr liter/sec. Most sublimators are only used at operating pressures below 10^{-6} torr, thus assuring a reasonable operating life before the sublimator must be replaced. The construction of practical sublimators is discussed later.

2.7.4.2 The Practical Application of Sublimation Pumps

The selection of a titanium sublimation source is determined by the requirements for both pumping speed and throughput. If the system always operates in the UHV range, and the throughput of gas is small, then the simplest type of device and manual control will give adequate life even at high pumping speeds. Where the throughput is substantial, then a device capable of dispensing appropriate quantities of titanium, preferably with automatic control of the sublimation rate, may prove essential.

When the titanium is deposited on a surface at room temperature, it provides *maximum* pumping speeds of the order of 3 and 9 liter/sec.cm^2, for hydrogen and oxygen, respectively [30]. Cooling the titanium film to liquid nitrogen temperature increases the pumping speeds by a factor between 2 and 5, depending on the gas. Higher pumping speeds can be easily achieved by increasing the area on which the titanium is deposited, but note that a larger surface area will necessitate an increase in the rate of sublimation to maintain the higher rate of pumping.

The maximum pumping speed is obtained when the titanium surface is clean. In practical applications, the surface will always be partially covered by an adsorbed layer, so the pumping speed will be only some fraction of this maximum value, determined by the frequency at which the surface is replenished by sublimation. Harra [30] provides guidance for the case where continuous sublimation is used. For the case where the pressure is very low, and periodic sublimation is used, it is necessary to assume some effective pumping speed of the titanium, which is less than the speed for the clean film [30], say 20 to 30%.

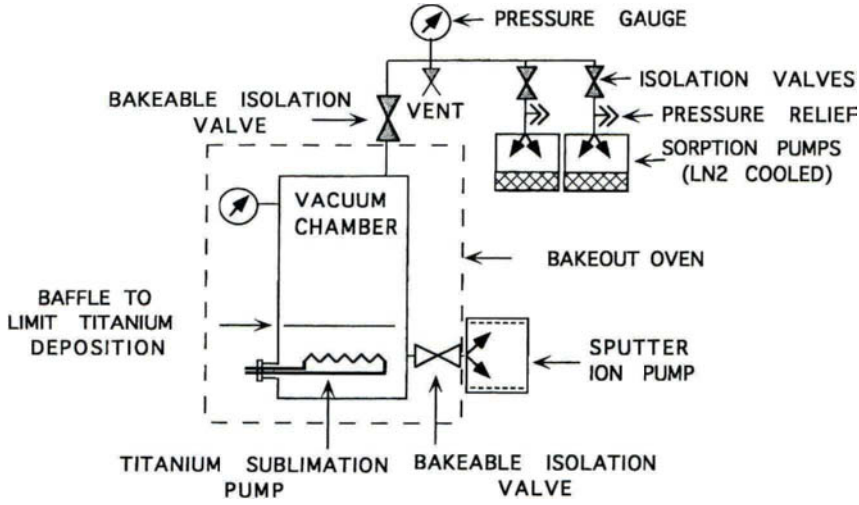
The required area of titanium in a particular system can be calculated from the equation

$$\text{Titanium area} = Q/P \times S \text{ cm}^2$$

where Q is the estimated influx of gas from all sources, including outgassing and leakage, in torr liter/sec
P is the requires pressure, in torr
S is the gettering speed per square centimeter of titanium for the gas being pumped

Particular care must be taken to assure that the conductance between the pump and the vacuum chamber does not seriously limit the effective speed of the pump.

Fig. 6.



Schematic of vacuum system using sputter-ion and titanium sublimation pumps in parallel, with sorption pump roughing system.

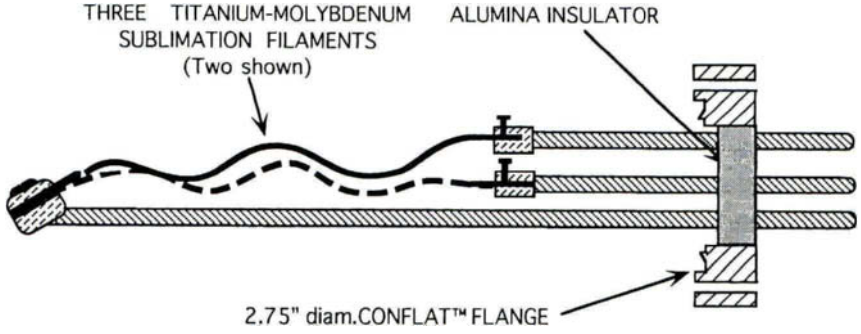
The sublimator is mounted inside the vacuum chamber so that the sublimed titanium film is deposited on the chamber walls, or on baffle plates, positioned to protect the working volume of the chamber from titanium deposits. A typical system combining a sputter-ion pump and a titanium sublimation pump is shown schematically in Figure 6.

For the reasons just discussed, do not heat the sublimator until the system pressure is well below 10^{-4} torr, so as to avoid unnecessary contamination of the source. This is not a serious limitation, since the sublimation pump is most commonly employed as a supplementary pump.

One of the only serious problems of the sublimation pump is the flaking of the sublimed titanium film from the surfaces on which it is deposited. This is precisely the same problem already discussed for sputter-ion pumps. It can be minimized by avoiding unnecessary venting to the atmosphere, or, by baking the sublimator after such exposure [26], before restarting the sublimation process.

In normal operations, the amount of titanium sublimed is invariably in excess of that necessary to sustain the required pressure, but this titanium remains accessible for hydrogen pumping. On exposure to the atmosphere, oxygen adsorption on the surface occurs, effectively blocking access to any unreacted titanium in the bulk. Steinberg and Alger [31] showed that hydrogen gettering on such titanium films can be restored by the deposition of a new film of titanium on the exposed surface. Presumably such a film permits dissociative chemisorption of hydrogen followed by diffusion of the hydrogen atoms into the bulk.

Fig. 7.



Titanium sublimator pump cartridge.

WIRE SUBLIMATOR SOURCES

The simplest form of sublimator consists of a feedthrough carrying two or more replaceable filament sources, which are heated, in turn, until all the titanium has been sublimed. A drawing of this type of sublimator is shown in Figure 7. Typically, each filament can sublime over 1 g of titanium before failure. Most filaments are made from a 85% titanium–15% molybdenum alloy. When heated at constant power, this alloy is more robust, and evaporates titanium at a more constant rate, over its life, than does a comparable pure titanium source [32,33]. A typical sublimator equipped with three filaments can dispense over 3 g of titanium when each filament is used, in turn, to exhaustion. With a rate of sublimation adjusted so as to just balance the rate of pumping, roughly 210 torr liter of gas can be pumped before the sublimators must be replaced.

SUBLIMATION FROM TITANIUM SPHERES

Substantially greater quantities of titanium can be provided from hollow titanium spheres using radiant heaters [34]. In one form, the heater is a tungsten filament mounted inside the sphere, and up to 35 g of titanium can be sublimed from a sphere of roughly 3.5 cm diameter. With these thin-walled spheres the structural integrity is a concern, because a phase change in titanium, at around 880°C, develops stresses in the titanium, which, with repeated temperature cycling will eventually cause failure [34]. For this reason, the power supply for the sublimator is programmed to hold the titanium at a stand-by temperature in excess of 880°C, where the sublimation rate is negligible, and to cycle up to a selected temperature above 1500°C when sublimation is required. Failure occurs when sublimation produces a hole in the sphere. Radiation losses through the hole cause the sphere to run at too low a temperature for usable rates of sublimation. One disadvantage

to the use of the titanium sphere is that it must be kept at an elevated temperature at all times, and this heats up the chamber walls, increasing the rate of outgassing.

SUBLIMATION BY ELECTRON BOMBARDMENT HEATING

For very-high-capacity pumping, titanium evaporation using an electron beam evaporator can be useful.

2.7.4.3 Control Units for Sublimation Pumps

The simplest form of power supply for a wire or sphere sublimator is simply an adjustable power source operated from a timer circuit; the time between heating cycles, and the heating time, can be adjusted to provide an *average* pumping speed. This type of regulation does not provide a fixed pumping speed, so that the pressure in a system slowly rises between periods of sublimation, and in addition, there is invariably a small pressure spike whenever sublimation starts. It is a trivial matter at pressures below 10^{-8} torr, because the amount of titanium required to sustain pumping is often only 1 or 2 grams per year, and the time between sublimation periods is so long as to provide a relatively stable pressure environment for experimentation or process work. However, if a large throughput of gas is present, then a high rate of sublimation will be needed, and this requires more careful control to sublime only enough titanium to maintain the required pressure. Although this process can be accomplished manually, it is far more efficient to employ a feedback control, with a pressure gauge as the sensing element, combined with careful selection of the periodicity, time, and temperature of sublimation. Under the best of circumstances, the control of a sublimation pump is somewhat crude, as compared to the inherent self-regulation of a sputter-ion pump.

2.7.4.4 Nonevaporable Getter Pump (NEG)

The nonevaporable or bulk getter removes gas by chemical interaction in precisely the same way as a sublimation pump, but pumping is sustained by the process of diffusion into the bulk of the getter material, rather than by the deposition of fresh getter material. A NEG is often operated at elevated temperature, which accelerates the rate at which chemisorbed surface species diffuse into the bulk of the getter. The pumping speed is determined by the rate of diffusion away from the surface. However, the speed decreases as the concentration of pumped gas in the solid increases, becoming very slow indeed as saturation is approached.

The early application of bulk getters was largely restricted to use in sealed-off vacuum devices; an excellent account of these early applications is given by

Reimann [35]. Titanium and zirconium were frequently selected, using a simple electrically heated filament mounted in the device, and operated at temperatures in excess of 800°C, to provide a high rate of gettering for gases such as oxygen, nitrogen, and the carbon oxides [36,37]. Gettering of these gases produces compounds so stable that they are not decomposed even at temperatures where the filaments melt. Because the rate of pumping is limited by the rate of diffusion, a higher speed for most gases except hydrogen is achieved simply by increasing the temperature, limited only by the danger of filament failure by sagging or melting. The problem with this approach is that the optimum temperature for pumping hydrogen (including that present in water vapor) is 400°C or less [36,37] At higher temperatures, the dissociation pressure of hydrogen in the metals increases. If the hydrogen partial pressure in the system is high, the pumping of hydrogen continues, but the ultimate pressure to which hydrogen can be pumped is limited by the dissociation pressure. At 400°C more reactive gases preferentially adsorb on the surface, displacing most hydrogen already there, and these adsorbed gases only slowly diffuse into the bulk. Consequently, when using a pure metal such as titanium or zirconium, the pumping of hydrogen — a dominant gas in the residual atmosphere of many devices — is extremely limited in the presence of most other reactive gases.

The widespread practical use of NEGs can be traced back to the innovative work of della Porta and his colleagues [38] in developing an alloy of zirconium, 84% Zr–16% Al, designated St 101. This alloy has the typical gettering capabilities of zirconium for reactive gases, but with the advantage of significantly lower gettering temperatures. A second major advantage is that hydrogen is gettered even in the presence of the more reactive gases. With the commercialization of this development, NEGs are available that can simultaneously pump all the principal, chemically reactive residual gases in a vacuum system when operated at ~400°C. The alloy is produced as a fine powder that is coated, without any binder, on a substrate [38]. The total quantity of getter is limited, so that the practical applications are for relatively low throughput conditions.

When a NEG is exposed to the atmosphere, during installation in a system, it is covered with an adsorbed layer of gas, making it unusable for gettering; it must be activated by heating to a temperature in the 500 to 800°C range, under vacuum, to allow the surface layer to diffuse into the bulk. The efficiency of activation increases with temperature and time. Since the gettering occurs even during the activation, reducing the amount of getter available for subsequent pumping, the pressure is kept as low as possible during activation. The published specifications for a particular getter pump using this material, quote an initial pumping speed for nitrogen of ~110 liters/sec, when operating at 400°C, and a practical pumping capacity of ~25 torr liter before reactivation is required. Reactivation by heating to ~800°C allows diffusion of the near-surface sorbed gas into the bulk, restoring the getter to nearly its initial performance. Reactivation can be re-

peated, giving a total useful pumping capacity of ~255 torr liter,[5] but the performance slowly degrades with each successive activation. This getter is particularly useful for pumping hydrogen and its isotopes, providing pumping speeds a factor of 5 higher than for nitrogen. Hydrogen pumping is effective even at ambient temperature (but at reduced speed), so long as the partial pressures of gases such as oxygen remain low.

Following the success of the original zirconium/aluminum getter alloy, NEGs have been developed that permit the pumping of all common gases, at still lower temperatures, and useful performance can now be achieved at ambient temperature. NEGs having a highly porous structure are very effective at low temperatures, providing a very large surface area accessible to the gas phase to provide sustained pumping. The large area compensates for the fact that the rate of diffusion into the bulk is limited at low temperature. These NEGs can be reactivated by periodically heating to a higher temperature, accelerating diffusion into the bulk, and freeing the surface for additional gettering. One main advantage of room temperature operation is obviously the absence of any power input during operation. One important NEG is an alloy of zirconium, vanadium, and iron, designated St 707 [39]. The temperature for full activation of these getters can be as low as 400–500°C, and even lower temperatures can still provide substantial gettering potential. An St 707 getter, capable of pumping nitrogen at approximately the same speed as the St 101 material just described, would have a practical pumping capacity roughly four to five times greater.

The ability to operate at ambient temperature is important in many large and complex systems designed to operate well below 10^{-10} torr; in many systems only mild degassing temperatures can be used, the components being degassed before assembly; the absence of any heating of the getters minimizes the system operating temperature, and helps keep the rate of degassing down to acceptable levels.

For use at higher throughputs, NEGs are also manufactured as self-supported porous structures, providing greatly increased quantities of accessible getter material.

2.7.4.5 Practical Application of NEGS

NEGs are used in a wide range of applications, including small dewars, sealed-off electronic devices, metal halide discharge lights, and large particle accelerators. They are particularly valuable for their ability to pump hydrogen, and so have

[5] To understand the limited capacity of this getter, compare this performance to that claimed for a sputter-ion pump, rated at 60,000 hours life at 1×10^{-6} torr, which implies a pumping capacity of 2×10^{4} torr liters.

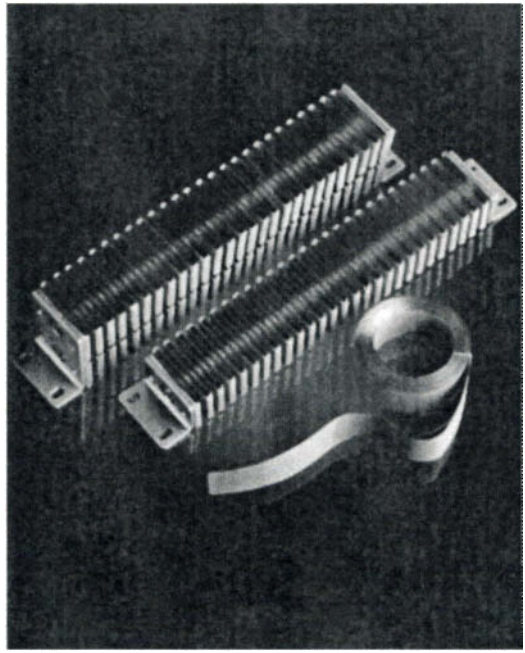
been used in fusion test reactors. They have also proved successful in particle accelerators, which have very long vacuum chambers and operate at very low pressures, providing a distributed pumping system that operates at ambient temperature, and has no moving parts.

NEGs are available as self-contained pumping units, fitted with integral heaters that range from tiny pellets with dimensions of the order of millimeters, having pumping speeds of a liter/sec or less, to pumps with speeds up to 750 liter/sec. Figure 8 shows an example of a large getter pump. Where very large pumping

Fig. 8.

A nonevaporable getter pump (NEG). The getter cartridge, consisting of four rows of accordion-pleated getter strips is on the right. The cartridge support and heater is on the left. Reprinted with permission from SAES Getters/USA., Inc.

Fig. 9.



Getter modules with integral heaters, for mounting inside the vacuum chamber. Reprinted with permission from SAES Getters/USA., Inc.

speeds are required, they are available as modules, such as those shown in Figure 9, which can be mounted inside the vacuum chamber so as to provide pumping close to the source of the gas load; in this function they have been used in place of cryogenic panels, to provide very high total pumping speeds.

A NEG can only pump reactive gases, so that supplementary pumping *must* be provided to remove noble gases, or methane. A positive characteristic is that they do not contribute any significant impurities in most applications.

The effective pumping speed is very dependent on the operating temperature and on the particular gas species to be pumped. Where hydrogen or its isotopes are the principal gas load, high pumping speeds can be maintained at ambient temperature, and large amounts of gas can be pumped. For such applications, the life of the getter is limited by embrittlement of the getter, rather than by the saturation.

Operation at elevated temperature provides more sustained pumping for all gases. Conversely, operating porous getters at low temperatures may require periodic heating to reactivate the getter by allowing the pumped gas to diffuse from the near-surface region into the bulk getter. The speed of pumping of all getters decreases with the quantity of gas that has been pumped, so careful attention must be paid to the effect of this decline for each application.

The combination of a NEG with a sputter-ion pump provides a significant pumping speed for hydrogen, which is particularly advantageous for very-low-pressure applications, and is also a very satisfactory method of pumping large quantities of hydrogen. The NEG is built into the envelope of some sputter-ion pumps, but a separate NEG unit can be combined with any sputter-ion pump to provide the same advantages. Hydrogen is reversibly sorbed by all getter materials, and hence the getter can be regenerated by heating, while using a supplementary pumping system. The regeneration feature makes NEG pumping of hydrogen, and its isotopes, very useful even where large quantities of gas must be pumped. For such cases, self-supporting thicker cross sections of getter are preferable to those forms in which the active material is coated as a thin layer on a substrate.

Note that NEG pumps, like sublimation pumps, have limited capability for pumping very large quantities of gas, especially as compared to the essentially unlimited capacity of a turbomolecular or diffusion pump.

Because the getter material in a NEG is only used up as gas is pumped, the pumps are self-regulating, in the same way as a sputter-ion pump. This places the NEG at a distinct advantage to the sublimation pump, where optimum life of the pump requires careful management of the rate of sublimation.

REFERENCES

1. D. Alpert, *Handbuch der Phys.*, **12**, (1958) 609.
2. L. D. Hall, *Rev. Sci. Instr.* **29**, (1968) 367.
3. K. M. Welch, *Capture Pumping Technology* (Pergamon Press, NY, 1991).
4. W. J. Lange, *J. Vac. Sci. Technol.* **7**, (1969) 228.
5. F. M. Penning, *Physica* **4**, (1937) 71.
6. W. M. Brubaker, *Trans. 6th Nat. Vac. Symp. (1959)* (Pergamon Press, New York, 1960), p. 302.
7. T. Tom and B. D. James, *J. Vac. Sci. Technol.*, **6**, (1969) 304.
8. R. L. Jepsen, *Proc. 4th Int. Vac. Congress (1968)* (Institute of Physics and the Physical Society, London, 1968), p. 317.
9. J. A. Vaumoron and M. P. DeBiasio, *Vacuum* **20**, (1970) 109.
10. A. R. Hamilton, *Trans. 2nd Int. Cong. and 8th Nat. AVS Symp.* (Pergamon Press, New York, 1962), p. 388.
11. M. Pierini and L. Dolcino, *J. Vac. Sci. Technol.* **A1**, (1983) 140.
12. J. H. Singleton, *J. Vac. Sci. Technol.* **6**, (1969) 316.
13. S. L. Rutherford and R. L. Jepsen, *Rev. Sci. Instr.*, **32**, (1961) 1144.
14. J. H. Singleton, *J. Vac. Sci. Technol.*, **8**, (1971) 275.
15. P. della Porta and B. Ferrario, *Vuoto* **1**, (1968) 2.
16. Thomas Snouse, *J. Vac. Sci. Technol.*, **8**, (1971) 283.
17. R. L. Jepsen, *Cooling Apparatus for Cathode Getter Pumps*, U.S. Patent 3,331,975, July 16, 1967.
18. J. E. Kelly and T. A. Vanderslice, *Vacuum*, **11**, (1961) 205.
19. David Lichtman, *J. Vac. Sci. Technol.*, **1**, (1964) 23.
20. A. K. Gupta and J. H. Leck, *Vacuum* **25**, (1975) 362.
21. J. H. Singleton, unpublished measurements.

22. R. Zaphiropoulos and W. A. Lloyd, *Proc. 6th Nat AVS Vac. Symp. 1969* (Pergamon Press, New York, 1960), p. 307.
23. M. Pierini and L. Dolcino, *J. Vac. Sci. x Technol.*, **A1**, (1983) 140.
24. Donald J. Santeler, private communication. Formerly at Process Applications, Inc., Oak Ridge, Tennessee.
25. ULTEK Bulletin SIB No. 200-9A, September 10, 1968.
26. D. R. Denison, *J. Vac. Sci. Technol.*, **4**, (1967) 156.
27. P. A. Redhead, J. P. Hobson, and E. V. Kornelsen, *The Physical Basis of Ultrahigh Vacuum* (Chapman and Hall, London, 1968), pp. 209–216.
28. M. V. Kuznetsov, A. S. Nasarov, and G. F. Ivanovsky, *J. Vac. Sci. Technol.* **6**, (1969) 34.
29. D. R. Denison, *Proc. 4th. Int. Vac. Cong.* (Institute of Physics and the Physical Society, London, 1968), p. 377.
30. D. J. Harra, *J. Vac. Sci. Technol.* **13**, (1976) 471.
31. R. Steinberg and D. L. Alger, *J. Vac. Sci. Technol.* **10**, (1973) 246.
32. G. M. McCracken and M. A. Pashley, *J. Vac. Sci. Technol.* **3**, (1966) 96.
33. R. W. Lawson and J. W. Woodward, *Vacuum* **17**, (1967) 205.
34. D. J. Harra and T. W. Snouse, *J. Vac. Sci. Technol.* **9**, (1972) 552.
35. Arnold L. Reimann, *Vacuum Technique* (Chapman and Hall, London, 1952).
36. V. L. Stout and M. D. Gibbons, *J. Appl. Phys.* **26**, (1955) 1488.
37. J. H. Singleton, *Residual Gases in Electron Tubes* (Academic Press, London and New York, 1972), p. 213.
38. P. della Porta, T. Giorgi, S. Origlio, and F. Ricca, *Trans. 2nd Int. Cong. and 8th Nat. AVS Symp.* (Pergamon Press, New York, 1962), p. 229.
39. C. Pisani and P. della Porta, *Suppl. Nuovo Cimento* **5**, (1967) 261.

This Page Intentionally Left Blank

PART 3

Vacuum Measurements

3.1 The Measurement of Low Pressures 257
- 3.1.1 Overview 258
- 3.1.2 Direct Reading Gauges 260
- 3.1.3 Indirect Reading Gauges 265
- 3.1.4 Calibration of Vacuum Gauges 286

References 288

3.2 Mass Analysis and Partial Pressure Measurements 290
- 3.2.1 Overview and Applications 290
- 3.2.2 Inlet Systems 300
- 3.2.3 Ion Generation and Ion Sources 303
- 3.2.4 Ion Separation and Analyzers 308
- 3.2.5 Detection of Ions 323

References 326

3.3 Practical Aspects of Vacuum System Mass Spectrometers 335
- 3.3.1 Historical Insight 335
- 3.3.2 Expected Gases in a Vacuum System 336
- 3.3.3 The Ion Generation Process 340
- 3.3.4 Techniques for Analysis 351
- 3.3.5 Calibration of Vacuum System Mass Spectrometers 364
- 3.3.6 Some Applications 370

References 374

3.4 Mass Flow Measurement and Control 376
- 3.4.1 General Principles of Mass Flow Measurement 376
- 3.4.2 Overview of Thermal Mass Flow Controller Technology 378
- 3.4.3 Performance Characteristics 382
- 3.4.4 Troubleshooting 386

References 387

This Page Intentionally Left Blank

CHAPTER 3.1

The Measurement of Low Pressures

Ron Goehner
Electron Technology Division

Emil Drubetsky
The Televac Division

Howard M. Brady
Electron Technology Division

William H. Bayles, Jr.
The Televac Division

Electron Technology and Televac are Divisions of the Fredericks Company

Measurements at or below atmospheric require knowledge of the expected pressure range and the accuracy and/or repeatability required by the processes taking place in the vacuum chamber.

Table 1 gives an indication of the pressure categories often encountered in vacuum applications. It is useful to remember that 1 gram molecular mass of any gas occupies 22.4 liters of volume at a standard temperature of 0° C and 1 atmosphere pressure. Also, 1 gram molecular mass contains 6.02×10^{23} molecules. Thus at this standard temperature and pressure 1 cm^3 of any gas will contain 2.5×10^{19} molecules. At reduced pressure, there will be proportionally fewer molecules.

ISBN 0-12-325065-7
$25.00

Table 1

Vacuum Categories

Pressure			Number Density	Mean Free Path	Surface Collision Frequency	Time[a] fo Monolay Formatio
	(Pa)	(Torr)	(cm^{-3})	(cm)	($cm^{-2}\ sec^{-1}$)	(sec)
One atmosphere	101.3×10^{3}	760	2.52×10^{19}	6.6×10^{-6}	2.8×10^{23}	3.3×10
Rough vacuum	1.33×10^{-1}	10^{-3}	3.3×10^{13}	5.0	4.4×10^{17}	2.5×10
High vacuum	1.33×10^{-4}	10^{-6}	3.3×10^{10}	5.0×10^{3}	4.4×10^{14}	2.5
Very-high-vacuum	1.33×10^{-7}	10^{-9}	3.3×10^{7}	5.0×10^{6}	4.4×10^{11}	2.5×10
Ultra-high-vacuum	1.33×10^{-10}	10^{-12}	3.3×10^{4}	5.0×10^{9}	4.4×10^{8}	2.5×10

[a]Assuming adhesion efficiency of 100% and a molecular diameter of 3×10^{-8} cm.

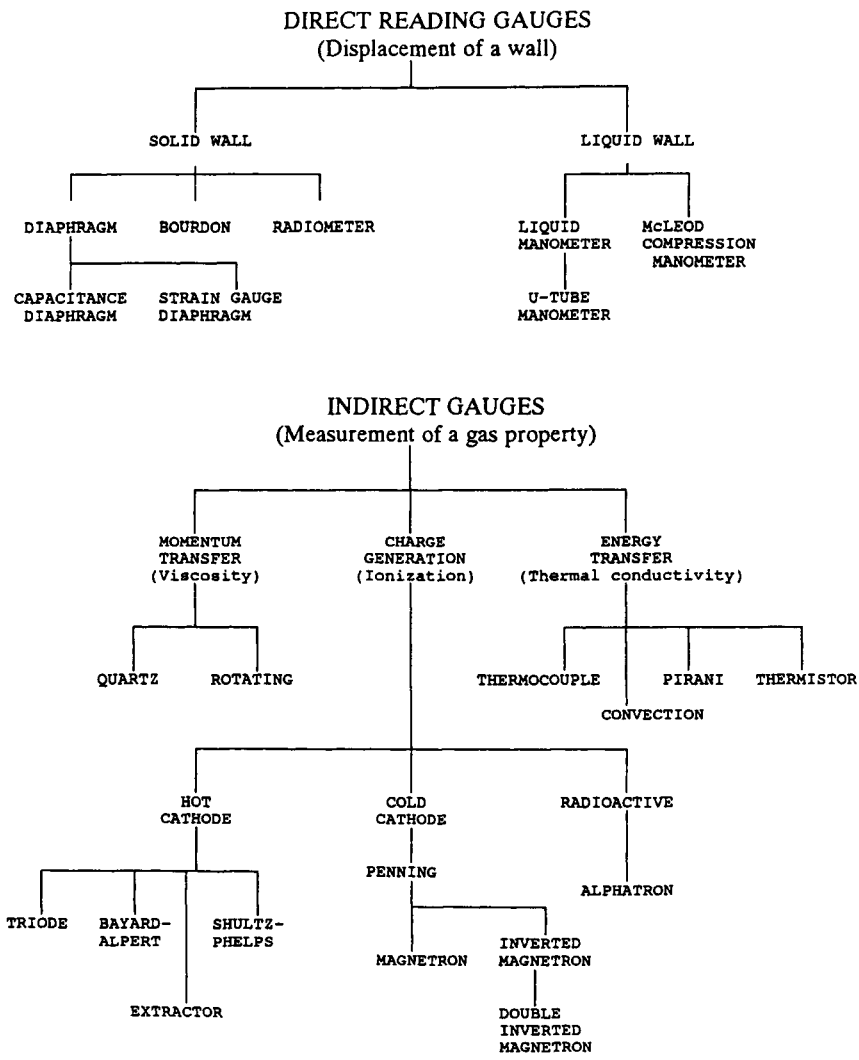
The time required to form a monolayer on a clean surface is of particular interest. From Table 1 one can see that at atmospheric pressure a monolayer is formed in nanoseconds. It takes a pressure of 10^{-6} torr or less to have a few seconds before a monolayer is formed. A more complete treatment of the quantities shown in Table 1 together with the defining equations that determine how these quantities vary with changes in pressure, volume, and molecular species, is given in Part 1 of this handbook.

3.1.1 OVERVIEW

The pressures that must be measured in most vacuum systems cover many orders of magnitude. In some simple applications the pressure range may be from atmospheric (1.01×10^{5} Pa, 760 torr) to 1×10^{-3} Pa (7.5×10^{-6} torr) or about 8 orders of magnitude. Other applications such as semiconductor lithography and high-energy physics experiments may require an ultimate vacuum of 7.5×10^{-9} torr or lower, which implies a pressure range of 11 orders of magnitude below atmospheric pressure. These great ranges require more than one gauge to give reasonable measurements.

The development of vacuum-measuring instruments (commonly called *gauges*) over the past 50 years has used transducers (or sensors), which can be classified as either direct reading or indirect reading [1]. Figure 1 schematically represents some of the more common gauges. Transducers that measure pressure by means

Fig. 1.

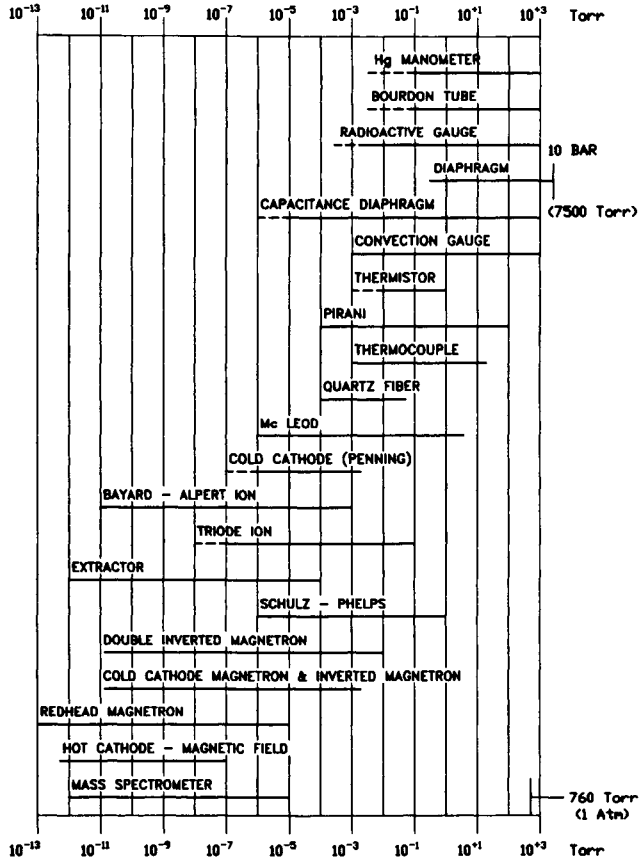


Classification of Pressure Gauges

of a force on a surface are called *direct* reading gauges, whereas those which determine pressure by means of any property of the gas that changes with density are called *indirect* reading gauges.

The operating pressure range for various types of vacuum gauges is shown in Figure 2.

Fig. 2.



Pressure Ranges for Various Gauges

3.1.2

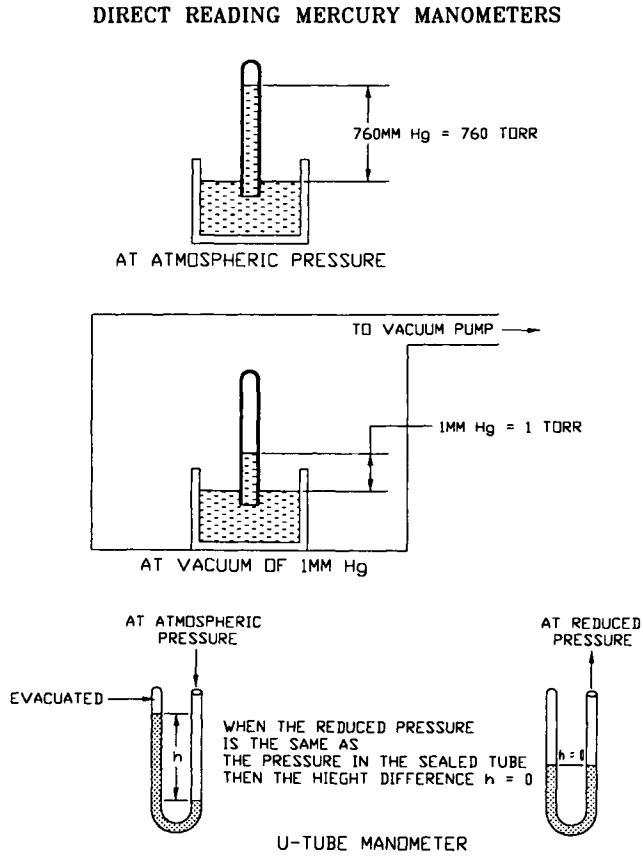
DIRECT READING GAUGES

Direct reading gauges can be divided into those that use a liquid wall and those that use a solid wall. In either case the gauge measures the pressure by reading the force exerted on a surface by thermally agitated molecules and atoms.

3.1.2.1 Liquid-Wall Gauges

There are essentially two common gauges that use a liquid wall: the liquid column manometer and the McLeod gauge.

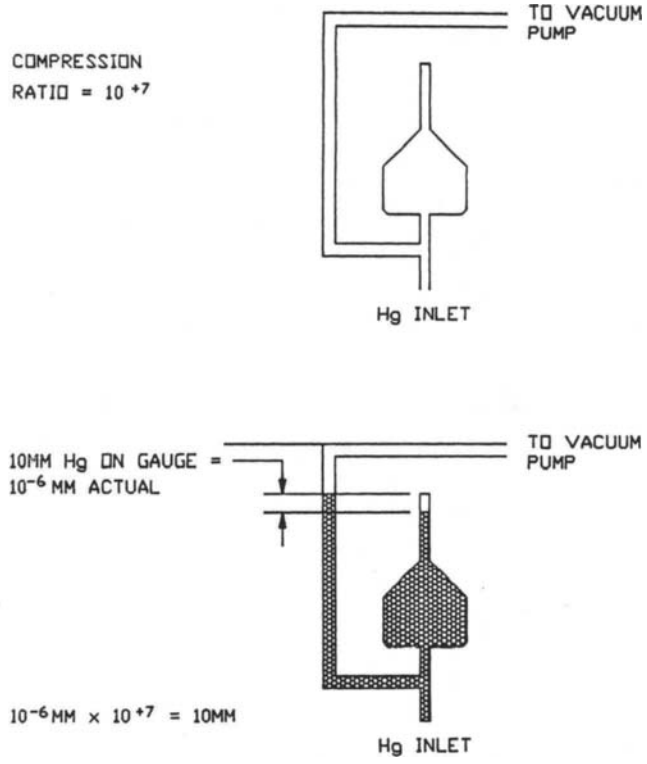
Fig. 3.



Mercury Manometers

Perhaps the simplest vacuum gauge is the liquid column manometer. This gauge (see Figure 3) consists of a straight or U-shaped glass tube exhausted and sealed at one end and filled partly with mercury or a low-vapor-pressure liquid such as diffusion pump oil. As the space above the mercury pool is evacuated in the straight tube, the length of the column of mercury decreases. In the case of the U-tube, as the free end is evacuated the two columns tend to assume the same height. The difference in height is a measure of the pressure at the open end. If the liquid is mercury, the pressure is directly measured in torr (mm of Hg). It is limited to pressures no lower than ~1 torr (133 Pa). If the liquid is a low-density oil, the U-tube is capable of measuring a pressure as low as ~0.1 torr. Although this gauge is an absolute direct reading instrument, the use of toxic mercury or creeping low-density oils that will in time contaminate the vacuum system, make it a poor choice.

Fig. 4.



McLeod Gauge

Due to the pressure limitation of the U-tube manometer the McLeod gauge [2] was developed. Using Boyle's law, this gauge greatly increased the range of vacuum measurement (see Figure 4). It uses a mercury manometer in which a volume of gas is compressed before measurement. The sensitivity of the McLeod gauge may be increased by increasing the compressible volume or by decreasing the diameter of the compression capillary. However, if the compression volume is too large, the weight of the mercury required to fill it sets a limit. If the capillary is made too small, the mercury will stick to the walls. These restrictions challenge the glass-blower's art [3]. This device when used in a liquid-nitrogen-trapped system by an experienced person, can be used as a primary standard. If as in Figure 4 the gas is at 10^{-6} torr and there is a compression ratio of 10^{7}, the difference between the columns will be 10 mm. The fragility of the glass and toxicity of the mercury argue against use by any but those skilled and careful individuals interested in a relatively inexpensive standard. Unless precautions such as liquid

nitrogen trapping are used any condensible vapors present will form a liquid on compression, giving a false low reading.

3.1.2.2 Solid-Wall Gauges

There are two major solid-wall gauge types — capsule and diaphragm.

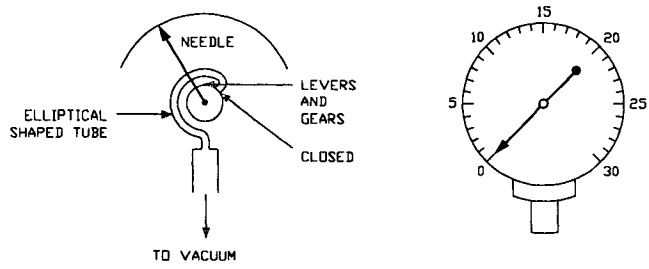
BOURDON GAUGES

The capsule-type gauges depend on the deformation of the capsule with changing pressure. The common pressure gauges on compressors and gas cylinders use this principle to measure pressures above atmospheric to several thousand psi. The sensitivity of this type of gauge is low at pressures below atmospheric. One variant, the Bourdon gauge (Figure 5), enjoys commercial popularity. In this case the capsule is in the form of a thin-walled tube bent in a circle with the open end attached to the vacuum system. The atmospheric pressure deforms the tube, and an attached pointer mechanism is deflected giving, in this case, a linear indication of the pressure that is independent of the nature of the gas. Various manufacturers supply gauges capable of measuring pressures as low as 1 torr. However, changing atmospheric pressure causes inaccuracies in the readings. Capsule gauges are rugged, inexpensive, and simple to use. Also, the capsule can be made of materials inert to corrosive vapors. Compensated versions of the capsule and Bourdon gauge have been developed that circumvent the inaccuracy caused by changing atmospheric pressure [4].

DIAPHRAGM GAUGES

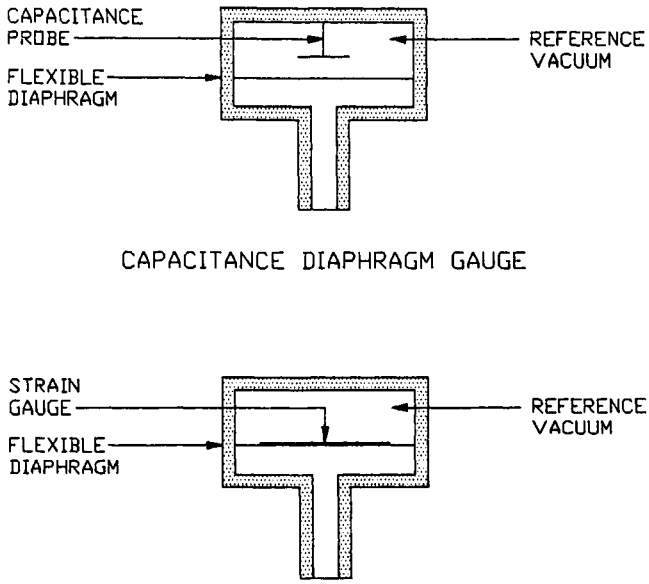
If compensated capsule or diaphragm mechanisms are married to suitable electronic measuring circuits, increased performance is achieved. One such arrangement is the capacitance diaphragm gauge (capacitance manometer).

Fig. 5.



Bourdon Gauge

Fig. 6.



Diaphragm Gauge

Figure 6 shows the salient features of the capacitance diaphragm gauge. A flexible diaphragm forms one plate of a capacitor, and a fixed probe the other. A change in pressure deforms the flexible diaphragm that changes the capacitance, which is converted to a pressure reading. The sensitivity, repeatability, and simplicity of this gauge enables this type of direct reading gauge to be used with appropriate heads as a standard from 10^{-6} torr to atmospheric pressure. A dynamic range of four or five orders of magnitude is obtainable in a single head [5].

Figure 6 also illustrates schematically the strain gauge type of diaphragm gauge. In this case deformation of the diaphragm causes a proportional output from the attached strain gauge. Sensitivities and dynamic range tend to be less than those of the capacitance diaphragm gauge. However, the price of the strain gauge type diaphragm gauge tends to be lower.

Both these gauges, with their high sensitivities, are prone to errors caused by even small temperature changes. Temperature-controlled heads or correction tables built into the electronics have been used to minimize this problem. Other sources of error in all solid-wall gauges are hysteresis and metal fatigue.

3.1.3

INDIRECT READING GAUGES

Indirect reading gauges measure some property of the gas that changes with the density of the gas.

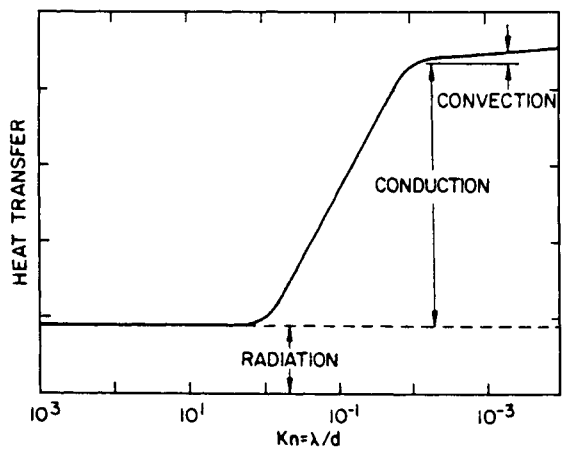
3.1.3.1 Thermal Conductivity Gauges

Thermal conductivity gauges use the property of gases to exhibit reduced thermal conductivity with decreasing density (pressure). The thermal conductivity decreases from a nearly constant value above ~1 torr to essentially 0 at pressures below 10^{-2} torr. The gauge controllers are designed to work with a specific sensor tube and substitutions are limited to those that are truly functionally identical. Figure 7 shows how the heat transfer regimes are related to the Knudsen number (see Part 1 of this handbook) which can be related to the geometry of the sensor and then to pressure.

PIRANI GAUGES

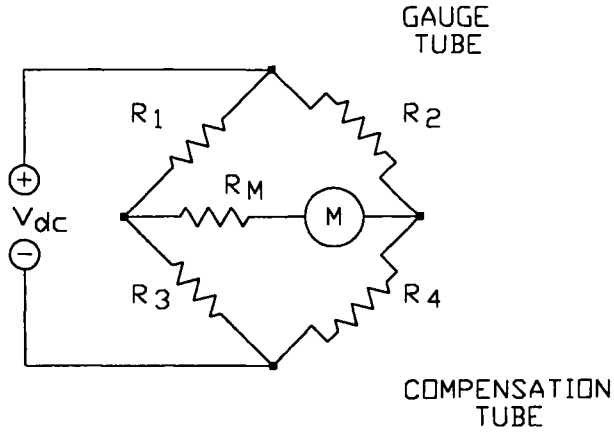
The Pirani gauge is perhaps the oldest indirect gauge that is still used today. In operation a sensing filament carrying current and producing heat is surrounded by the gas to be measured. As the pressure changes, the thermal conductivity changes, thus varying the temperature of the sensing filament. The temperature

Fig. 7.



Heat Transfer Regimes in a Thermal Conductivity Gauge. Reprinted with permission from A Users' Guide to Vacuum Technology, John F. O'Hanlon, p. 57. Copyright 1980, John Wiley & Sons.

Fig. 8.



Basic Pirani Gauge Circuit. Reprinted with permission from A Users' Guide to Vacuum Technology, John F. O'Hanlon, p. 58. Copyright 1980, John Wiley & Sons.

change causes a change in the resistance of the sensing filament. The sensing filament is usually one leg of a Wheatstone bridge. The bridge may be operated so that the voltage is varied to keep the bridge balanced; i.e., the resistance of the sensing filament is kept constant.

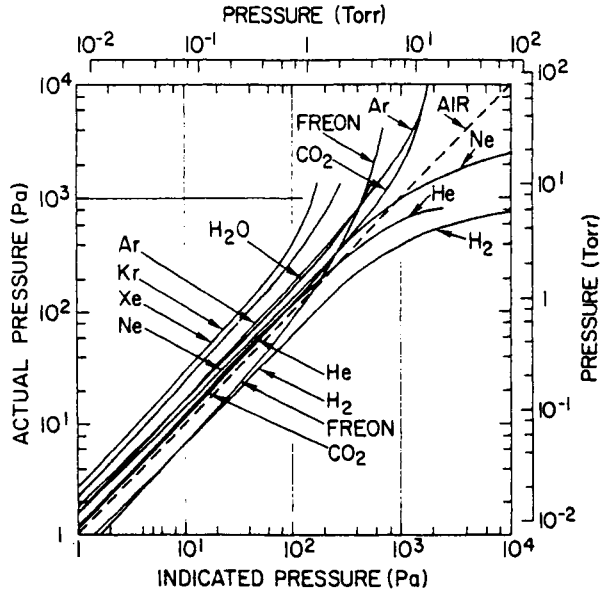
This method is called the *constant temperature method* and is deemed the fastest, most sensitive, and accurate. To reduce the effect of changing ambient temperature, an identical filament sealed off at very low pressure is placed in the leg adjacent to the sensing filament as a balancing resistor. Because of its high thermal resistance coefficient, the filament material is usually a thin tungsten wire; in fact a 10-watt light bulb works quite well [6] (see Fig. 8).

A properly designed, compensated Pirani gauge with sensitive circuitry is capable of measuring to 10^{-4} torr. However, the thermal conductivity of gases varies, which causes a variation in gauge response. These variations can be as large as a factor of 5 at the low-pressure end and as high as 10 at the high-pressure end of the scale (see Figure 9). If the composition of the gas is known, use can be made of the calibration curves supplied by the manufacturer. Operation in high partial pressures of organic molecules — oils, for example — is discouraged.

THERMOCOUPLE GAUGES

The thermocouple is another example of an indirect reading thermal conductivity gauge. The thermocouple gauge sensing element consists of a filament, usually of resistance alloy, heated by the passage of a constant current (see Figure 10). A thermocouple is welded directly to the midpoint of the filament or preferably to a conduction bridge at the center of the heated filament. This provides a means of directly measuring the temperature. With a constant current through the filament,

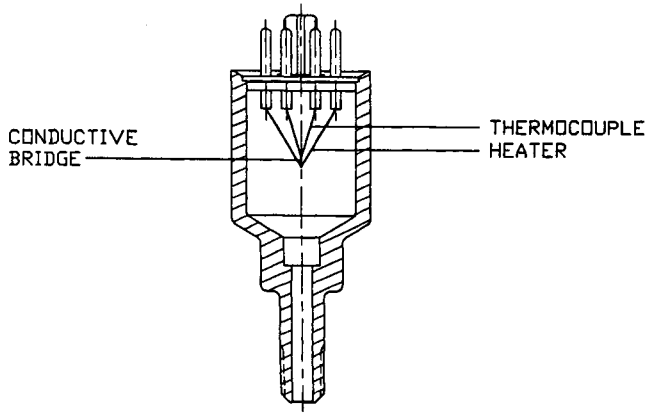
Fig. 9.



Calibration Curves for Pirani Gauge. Reprinted with permission from Leybold-Heraeus G.M.b.H., Postfach 51 07 60,5000 Köhn, Germany.

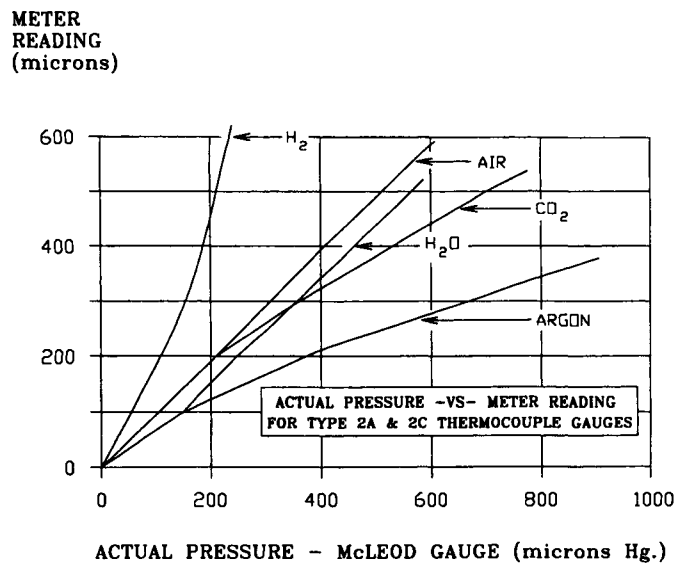
the temperature increases as the pressure decreases. This happens because there are fewer molecules surrounding the filament to carry the heat away. The thermocouple output voltage increases as a result of the increased temperature and is an inverse indication of the pressure. The thermocouple gauge can be operated in the constant temperature mode just like the Pirani gauge.

Fig. 10.



Thermocouple Gauge

Fig. 11.



Calibration Curves for Thermocouple Gauge

The same considerations concerning gas composition stated for the Pirani gauge apply to all thermal conductivity gauges. Figure 11 shows the calibration curves for a typical thermocouple gauge. Like the Pirani, the thermocouple gauge can be optimized for operation at various pressure ranges.

As with the Pirani gauge, operation in high partial pressures of organic molecules—oils, for example—is to be avoided. One manufacturer preoxidizes the thermocouple sensor for stability in "dirty" environments and for greater interchangeability.

THERMISTOR GAUGES

A device closely related to the Pirani gauge is the thermistor gauge. In this gauge, a single thermistor is used as one leg of a bridge circuit. The inverse resistive characteristics of the thermistor element unbalances the bridge as the pressure changes causing a corresponding change in current. A sensitive micrometer calibrated in pressure units measures the current. The thermistor gauge covers about the same pressure range as the thermocouple but may have a slight edge in withstanding contamination by oils and other organic compounds. As is true for other thermal conductivity gauges, the exact calibration depends on the nature of the gas measured. In a well-designed bridge circuit the plot of current vs. pressure is practically linear in the range of 10^{-3} to 1 torr [7]. Modern thermistor gauges use constant temperature, as discussed earlier.

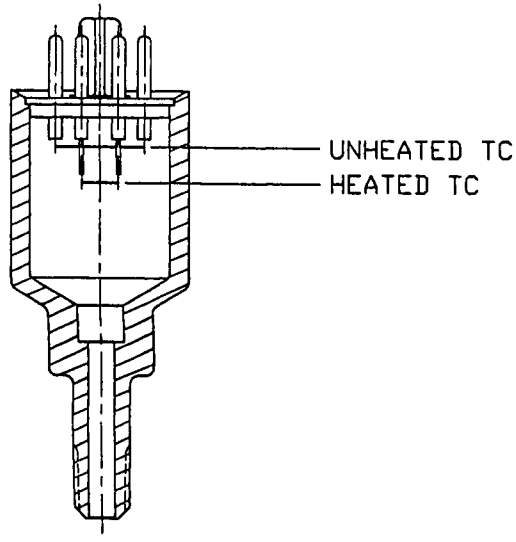
CONVECTION GAUGES

As noted, the high-pressure operating range of thermal conductivity gauges is normally limited to 1 torr. This is due to the near-constant thermal conductivity at pressures higher than 1 torr. However, there is a small contribution to heat transfer caused by convection. Several manufacturers have developed gauges that use convection to extend the usable range to atmospheric pressure and slightly above [8,9,10,11,12]. Because this process uses convection, orientation is critical.

The Convectron™ uses the basic structure of the Pirani with special features to enhance convection cooling in the high-pressure region [13]. To use the gauge above 1 torr (133 Pa), the sensor tube must be mounted with its major axis in a horizontal position. If the only area of interest is below 1 torr, the tube may be mounted in any position. As mentioned, the gauge controller is designed to be used with a specific model sensor tube, and since extensive use is made of calibration curves and lookup tables stored in the controller, no substitution is recommended.

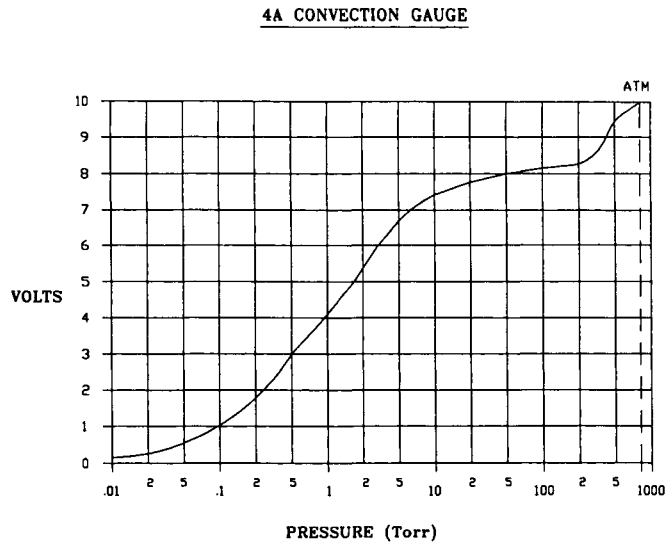
The Televac convection gauge uses the basic structure of the thermocouple gauge except that two thermocouples are used [14]. As in any thermocouple gauge, the convection gauge measures the pressure by determining the heat loss from a fine wire maintained at a constant temperature. The response of the sensor depends on the gas type. A pair of thermocouples is mounted a fixed distance from each other (see Figure 12). The one mounted lower is heated to a constant temperature by a variable current power supply. Power is pulsed to this lower

Fig. 12.



Convection Gauge

Fig. 13.



Output Curve for Convection Gauge

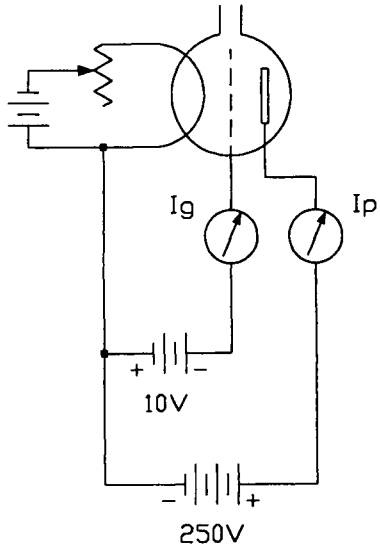
thermocouple, and the temperature is measured between heating pulses. The second (upper) thermocouple measures convection effects and also compensates for ambient temperature. At pressures below ~2 torr (270 Pa) the temperature in the upper thermocouple is negligible. The gauge tube operates as a typical thermocouple in the constant temperature mode. Above 2 torr, convective heat transfer causes the upper thermocouple to be heated. The voltage output is subtracted from that of the lower, requiring more current to maintain the wire temperature. Consequently, the range of pressure that can be measured (via current change) is extended to atmospheric pressure (see Figure 13). Orientation of the sensor is with the axis vertical.

The use of convection gauges with process control electronics allows for automatic pumpdown with the assurance that the system will neither open under vacuum or be subject to overpressure during backfill down to air conditions. These gauges with their controllers are relatively inexpensive. In oil-free systems, they afford long life and reproducible results.

3.1.3.2 Hot-Cathode Ionization Gauges

Hot-cathode ionization gauges include the simple triode as well as Bayard-Alpert and other designs.

Fig. 14.



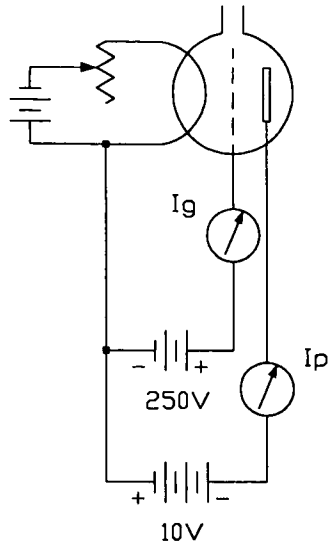
Typical Triode Connection

TRIODE HOT CATHODE IONIZATION GAUGES

The triode electron tube—a device consisting of a thermionic emitting cathode (in this case a filament), a grid, and an anode (old-timers call it a plate)—has been used as an indirect vacuum gauge for at least 80 years [15]. In fact this application is an extension of the diagnostic "gas test" used to this day by gauge tube manufacturers [16]. In the gas test a triode, or any thermionic gauge tube with three or more elements, is connected as an amplifier (see Figure 14 for a typical triode connection). The voltages are arranged so that the grid is at −10 volts or more with respect to the cathode. A fairly high voltage (+250 volts) is placed on the anode or plate, and the filament voltage is adjusted to allow 10 mA to flow from cathode to anode. Under these conditions the current flowing to the grid is clearly caused by positive ions formed by collision of the electron stream with gas molecules in the cathode to plate space. If the triode has cylindrical symmetry, the ion/collector current can easily be estimated [17]. Assuming the filament to plate spacing is 1 cm, the pressure in the tube is 10^{-5} torr and the electron (anode) current, the current flowing in the cathode to plate (anode) circuit, is 10 mA, the mean free path (mfp) of an electron in air at 10^{-5} torr is 4×10^3 cm and the probability of ionization is 10%. Using Equation (1), we can then say the ion current collected by the grid is 2.5×10^{-7}.

$$I_i = (I_e)\,(D_{f-p})\,(R_i/L_c) \tag{1}$$

Fig. 15.



Alternate Triode Connection

where I_i = ion current (A)
I_e = electron current (A)
R_i = ionization probability (decimal) [ions/cm^3/torr]
L_e = electron mfp (cm)
D_{f-p} = filament-to-plate distance (cm)

A more rigorous treatment of triode gauge performance is given by Morgulis [18] and Reynolds [19] as recorded by Leck [20].

It was discovered that if the triode were connected as in Figure 15 so that the grid is positive and the plate is negative all with respect to the filament, the ion current collected by the plate for the same electron current to the grid was greatly increased [21]. The ion current increases because the electron path is much longer as electrons miss the grid wires, turn around, and miss again. The electrons oscillate in this fashion until they strike the grid.

It is in this second higher-sensitivity mode that the triode gauge is used today. The linearity of ion current to molecular density (pressure) has been demonstrated by several investigations [15,21,22]. This linearity allows one to define a sensitivity factor S such that

$$I_i = S \times I_e \times P \tag{2}$$

where I_i = ion current (A)
I_e = electron current (A)
P = pressure
S = sensitivity (in units of reciprocal pressure)

Another, and perhaps more confusing sensitivity that was used by the industry in the past, was to quote the sensitivity in microamps per micron of Hg at a certain electron current. In this manner, the micrometer in the ion collector circuit is able to read directly in torr with a suitable range multiplier. For example, the VG1A, perhaps the most common triode gauge of its time, had a sensitivity to dry air of 100 microamps/micron at an electron current of 5.0 milliamps.

The triode gauge with suitable controller providing regulated grid and collector voltages together with a servo circuit to ensure constant electron current and a sensitive current meter at the collector is capable of measuring pressures from 10^{-3} to 10^{-8} torr. Schulz-Phelps triode gauges allow a pressure range of 10^{-1} to 10^{-6} torr [23]. The stability and linearity of triode gauges are such that they have been used as a secondary standard [24], however as with all indirect reading gauges, the triode sensitivity is dependent on the gas being measured [25].

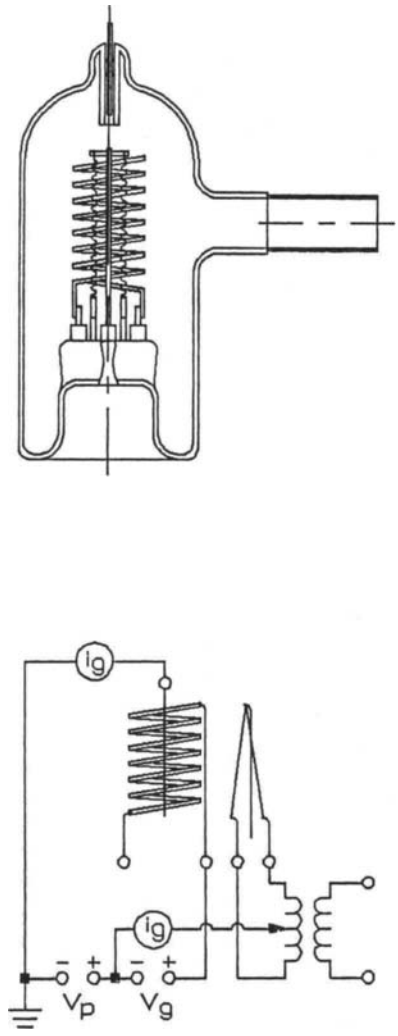
In all but special applications that require operation at relatively high pressures, the triode gauge has been replaced by the Bayard-Alpert gauge.

BAYARD-ALPERT HOT-CATHODE IONIZATION GAUGES

In the 1940s the gauge capable of measuring the lowest pressure was the triode gauge. It soon became apparent that the so-called pressure barrier at 10^{-8} torr was caused by a failure in measurement rather than pumping [26]. Nottingham correctly deduced that this artificial barrier was caused by electrons striking the grid, causing low-energy X-rays, which in turn struck the ion collector (plate), emitting photoelectrons. The electron emission is indistinguishable from positive ion collection and was calculated to be of the order of 10^{-8} torr for the triode gauges then in use [27]. In 1950 a simple solution to this problem was proposed by Bayard and Alpert [28]. This is now the most widely used gauge for general UHV measurement.

The Bayard-Alpert gauge is essentially a triode gauge reconfigured so that only a small amount of the internally generated X-rays strike the collector. Figure 16 shows the essential features of the gauge and a simple circuit. One can see that the cathode has been replaced by a thin collector located at the center of the grid. The cathode filament is now outside the grid and spaced several millimeters from it. One advantage to the Bayard-Alpert design was its ability to use the same controller as the triode gauge taking into account any sensitivity differences. The nude version of the Bayard-Alpert gauge is shown in Figure 17. At about the same time the Bayard-Alpert gauge was gaining popularity, Weinriech offered a solution to the burnout of the tungsten filaments then prevalent [29,30]. The use of platinum metals coated with refractory oxides allowed the gauge to withstand sudden exposure to atmosphere with the filament hot. The common combinations used are either thoria or yttria on iridium. It was soon noticed that Bayard-Alpert and even triode gauges of identical structure and dimensions but with different filaments, e.g. tungsten vs. thoria iridium, had different sensitivities, tungsten filament ver-

Fig. 16.

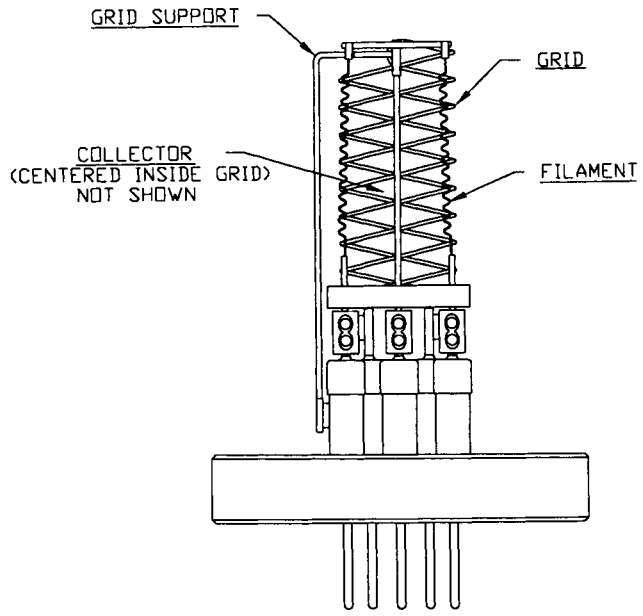


Bayard-Alpert Hot Cathode Ionization Gauge

sions being 20 to 40% more sensitive than the iridium of the same construction.

As discussed earlier, low-energy X-rays striking the ion collector and emitting electrons set a limit to the lowest pressure that can be measured. A variety of innovative schemes to reduce this X-ray limit have been proposed. Special designs with extremely fine collectors have been made that extend the high-vacuum end of readability to 10^{-12} torr but at the expense of the high-pressure end [31,32].

Fig. 17.

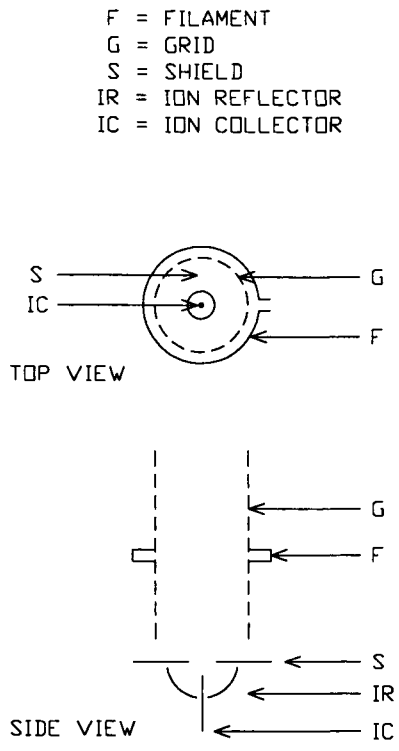


Nude Ionization Gauge

Redhead [33] designed a modulated gauge with an extra electrode near the ion collector. Measuring the ion current at two modulator potentials allows the X-ray current to be subtracted increasing the range to 5×10^{-12}. Other gauges use suppressor electrodes in front of the ion collector [34,35]. However, at this time the most widely used UHV hot-cathode gauge for people who need to measure 10^{-12} torr is the extractor [36]. This gauge is shown schematically as Figure 18. In this gauge, the ions are extracted out of the ionizing volume and deflected or focused onto a small collector. More recent designs have been developed [37,38] and a variant using a channel electron multiplier [39] has reduced the low-pressure limit to 10^{-15} torr.

Although the Bayard-Alpert gauge is the most widely used in its pressure range, it does suffer from some disadvantages. Inspection of Equation (1), which holds in a general way for Bayard-Alpert gauges, indicates that the ion current, as well as being pressure dependent, is geometry dependent and changes in geometry, tube to tube, or during the life of the tube, will change the sensitivity. Redhead [40] has reported on these sensitivity changes, as has Ohsako [56]. Other investigators have reported on the inaccuracy and instability of Bayard-Alpert gauges, with widely differing results [41,42,43,44,45,46]. Designs have been put forth that reduce or eliminate these problems [47,48]. As mentioned earlier, the sensitivity of the Bayard-Alpert gauge is dependent on the nature of the gas being mea-

Fig. 18.



Extractor Ionization Gauge

sured. Flaim and Owenby [49] presented data on the relative sensitivity of various gases that are presented by O'Hanlon [50]. These data, with additions from commercial sources, are shown as Table 2.

3.1.3.3 Cold-Cathode Ionization Gauges

In 1936 Penning [51] invented the cold-cathode discharge gauge as a method to measure pressures below 10^{-3} torr. In a Crookes tube below 10^{-3} torr, the mean free path is so great that little ionization takes place. Penning increased the probability of ionization by placing a magnetic field parallel to the paths of ions and electrons, which forces both particles into helical trajectory.

In one configuration of this gauge (see Figure 19), an anode is placed midway between two parallel-connected cathodes. The anode is a loop of metal wire whose plane is parallel to that of the cathodes. A potential difference of a few kilovolts is maintained between the anode and the cathodes. In addition, a magnetic field is applied between the cathodes by a permanent magnet, usually external to the

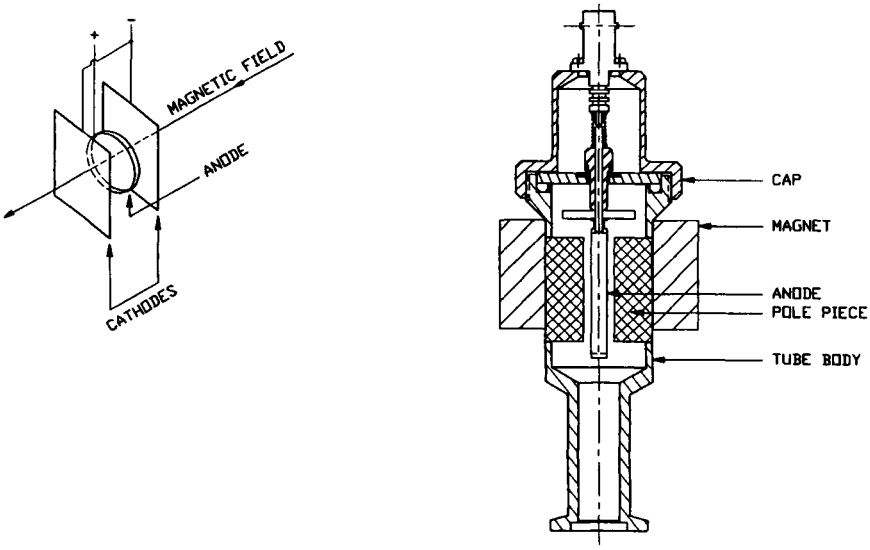
Table 2

Average Ionization Gauge Sensitivity Ratios "r" for Various Gases Normalized for Nitrogen

Gas	Reynolds[19]	Dushman & Young[76]	Wegener & Johnson[77]	Riddiford[78]	Shultz[79]	O'Hanlon[1]	ETI[80]
He	0.16	0.15		0.26	0.21	0.18	0.13
Ne	0.23	0.24			0.33	0.25	
Ar	1.30	1.19		1.06	1.49	1.20	1.47
Kr		1.86					
Xe		2.73					
H_2		0.46	0.52	0.38	0.42	0.48	0.42
N_2	1.00	1.00	1.00	1.00	1.00	1.00	1.00
O_2			0.85	1.14		0.85	0.77
Hg	2.75	3.44				3.50	
Dry Air				0.81			0.90
CO			1.07			0.48	1.01
CO_2			1.37				1.09
H_2O			0.89			0.90	
SF_6					2.54		

gauge body. Electrons emitted from either of the two cathodes must travel in helical paths due to the magnetic field, eventually reaching the anode, which carries a high positive charge. During the course of this long electron path, many of the electrons collide with the molecules of residual gas, creating positive ions, which

Fig. 19.



Penning Gauges

travel more directly to the cathodes. The ionization current thus produced is read out on a sensitive current meter in units of pressure.

This gauge, with its increased ruggedness as compared to the glass envelope hot-cathode ionization gauge, has found favor in recent years for industrial applications. Applications include use in leak detectors, vacuum furnaces, and electron beam welders. The cold-cathode gauge is also known as the Penning gauge or the Phillips ionization gauge (PIG).

The Penning gauge is rugged, simple, and inexpensive. However, the current is not a linear function of the pressure. The range of the Penning gauge is only 10^{-3} to 10^{-6} torr, and there is some instability and lack of accuracy. This gauge does not have the accuracy and stability of hot-cathode gauges. Penning and Nienhauis [52] improved performance by using a cylinder for the anode with cathode plates at each end and fitted with a cylindrical magnet. This slightly more complex gauge operates in the range of 10^{-2} to 10^{-6} torr (1 to 10^{-4} Pa). Even with this improvement, the Penning designs have a problem: At pressures below 10^{-5} or 10^{-6} torr, the discharge-initiating operation may fail to strike. Applying higher voltage does assist starting, but leads to field emission from the ion collector, which then sets the limit of low-pressure measurement.

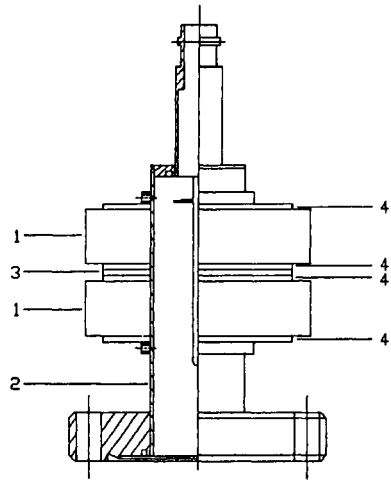
Redhead extended the range with a magnetron design [53] and later with Hobson by using an inverted magnetron design [54]. These designs extended the low-pressure range to 10^{-12} torr [55] or better. With all these improvements, problems with instability, hysteresis, and starting, though improved, still exist [57,58]. Modern magnetrons use a more simplified design without an auxiliary cathode.

More recently a double inverted magnetron was introduced [59]. This gauge has greater sensitivity (amp/torr) than the other types. It has been operated successfully at $\approx 1 \times 10^{-11}$ torr. The gauge (see Figure 20) consists of two axially magnetized annular-shaped magnets (1) placed around a cylinder (2) so that the north pole of one magnet faces the north pole of the other one. A nonmagnetic spacer (3) is placed between the two magnets, and thin shims (4) are used to focus the magnetic fields. This gauge has to date been operated at -10^{-11} torr, stays ignited and reignites quickly when power is restored at this pressure. Essentially instantaneous reignition has been demonstrated by use of radioactive triggering [60].

3.1.3.4 Resonance Gauges

One example of a resonance-type vacuum gauge is the quartz friction vacuum gauge. Although not widely and commonly used for general vacuum measurement, one such example is reviewed [61].

Fig. 20.



Double Inverted Magnetron

A quartz oscillator can be built to measure pressure by a shift in resonance frequency caused by static pressure of the surrounding gas or by the increased power required to maintain a constant amplitude. Used in this way, it is effective in measuring from near-atmospheric pressure to about 0.1 torr. A second method is to measure the resonant electrical impedance of a tuning fork oscillator. This method is shown to be proportional to the friction force applied by the ambient gas. Test results for this device show an accuracy within $\pm 10\%$ for pressure from 10^{-3} to 10^{3} torr. The lengths of the tuning fork oscillator varied from 2.5 to 20 mm.

3.1.3.5 Molecular Drag Gauges (spinning rotor) Gauges

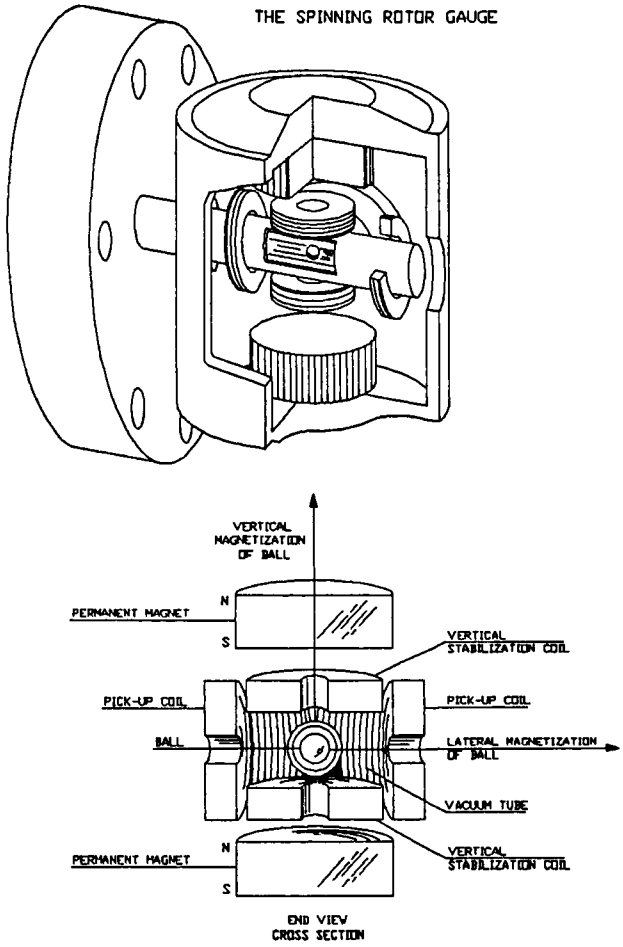
The idea of measuring pressure by means of the molecular drag of rotating devices was introduced by Meyer [62] and Maxwell [63] in 1865. In their devices the rotors were tethered to a wire or thin filament. Holmes [64] introduced the concept of the magnetic rotor suspension, leading to the spinning rotor gauge. Beams [65] et al. disclosed the use of a magnetically levitated, rotating steel ball to measure pressure at high vacuum. Fremerey [66] has traced the interesting historical development of this gauge.

The spinning rotor gauge, more usually called the molecular drag gauge (MDG), has become more popular since its commercial introduction in 1982 [67]. The useful range of the MDG overlaps the low-pressure range of the capacitance manometer and the high-pressure range of hot-cathode ion gauges, thus filling an

important gap. The linearity of the MGD is better than that of the ion gauge at high pressure (10^{-2} to 10^{-4} torr) and is theoretically more stable than other gauges at lower pressures [68].

In a MDG, the rate of change of the angular velocity of a freely spinning ball is proportional to the gas pressure and inversely proportional to the mean molecular velocity. Angular velocity is determined by removing the driving force and then measuring the AC voltage induced in the pickup coils by the magnetic moment of the ball (see Figure 21).

Fig. 21.



Molecular Drag Gauge (Spinning Rotor)

Briefly, a small rotor (steel ball bearing) about 4.5 mm in diameter is magnetically levitated and spun up to about 400 Hz by induction. The ball, enclosed in a thimble connected to the vacuum system, is allowed to coast by turning off the inductive drive. Then the time of a revolution of the ball is measured by timing the signal induced in a set of pickup coils by the rotational component of the ball's magnetic moment. Gas molecules exert a drag on the ball, slowing it at a rate set by the pressure P, its molecular mass m, temperature T, and the coefficient of momentum transfer σ, between the gas and the ball. A perfectly smooth ball would have a value of unity. There is also a pressure-independent residual drag (RD) caused by eddy current losses in the ball and surrounding structure. There are also temperature effects, which cause the ball diameter and moment of inertia to change.

In the region of molecular flow, the pressure can be determined from

$$P = \frac{\pi \rho a \bar{c}}{10 \sigma_{\text{eff}}} \left(\frac{-\omega'}{\omega} - RD - 2 \alpha T' \right) \tag{3}$$

where ρ is the density of the rotor
a is the radius of the rotor
ω'/ω is the fractional rate of slowing of the rotor
$\bar{c}$ is the mean gas molecular velocity
α is the linear coefficient of expansion of the ball
T' is the rate of change of the ball's temperature

All the terms in the first part of the equation can be readily determined except for the accommodation coefficient σ, which depends on the surface of the ball and the molecular adhesion between the gas and the surface of the ball. The accommodation coefficient σ must be determined by calibration of the MDG against a known pressure standard or, if repeatability is more important than the highest accuracy, by assuming a value of 1 for σ. Dittman et al. [68] have made repeated measurements of σ on many balls over several years. The values obtained ranged from 0.97 to 1.06 for 68 visually smooth balls. Assuming a value of 1 for σ would not introduce a large error and would according to Fremerey [66] allow the MDG to be considered a primary standard. The controller [68] contains the electronics to power and regulate the suspension and drive, detect and amplify the signal from the pickup coils and then time the rotation of the ball. It also contains a data processor that stores the calibration data and computes the pressure.

McCulloh et al. [69] have studied the zero stability of the MDG and have concluded that the largest contribution to instability is temperature change of the rotor. Changes in accommodation coefficient have been studied by Fremerey [66] and Cosma et al. [70] and, although the value will not change more than a few percent, it is difficult to predict how a specific ball will react to differing gas species.

These difficulties notwithstanding, Dittman et al. [68] have indicated that the short-term stability of MDG is much less than 1 percent, and the long-term stability is of the order of a few percent. Although these results were obtained under near-ideal laboratory conditions, careful users will be able to reproduce these results. The MDG is perhaps the best available transfer standard for the pressure range of 10^{-2} to 10^{-7} torr (1 to 10^{-5} Pa), but as it is designed for laboratory use in controlled relatively vibration free environments its use in industrial applications is severely limited. The MDG is expensive. Typically it requires several minutes to achieve accurate results under essentially static conditions and hence will be found most often in research and standards labs.

3.1.3.6 Mass Spectrometers

The theory and practical applications of mass spectrometers are discussed in Section 3.3.5 of this handbook. Calibration of mass spectrometers can be accomplished by equating the integral (the total area under all the peaks, taking into account the scale factors) to the overall pressure as measured by another gauge (cold-cathode, BA, or MDG). Once calibrated, the mass spectrometer can be used as a sensitive monitor of system pressure. It is also important when monitoring system pressure to know what gases are present. Most gauges have vastly different sensitivities to different gas species.

3.1.3.7 Vacuum-Measuring Instruments

SINGLE-MEASUREMENT INSTRUMENTS

Several manufacturers sell vacuum instruments that make single measurements or that measure multiple sensors of the same type (see Figure 22). These are available with either analog or digital displays, analog "recorder" outputs and relays. Some also have output to directly interface with computers for process control.

COMBINATION AND MODULAR INSTRUMENTS

Where a large range of pressures are to be measured, a combination of sensors are required to cover the range. The vacuum instruments in this case must accommodate these multiple-sensor types. Combination instruments are then used. More recently these combination instruments are of a modular design that accepts printed circuit board "cards" for each type of sensor desired.

Fig. 22a.

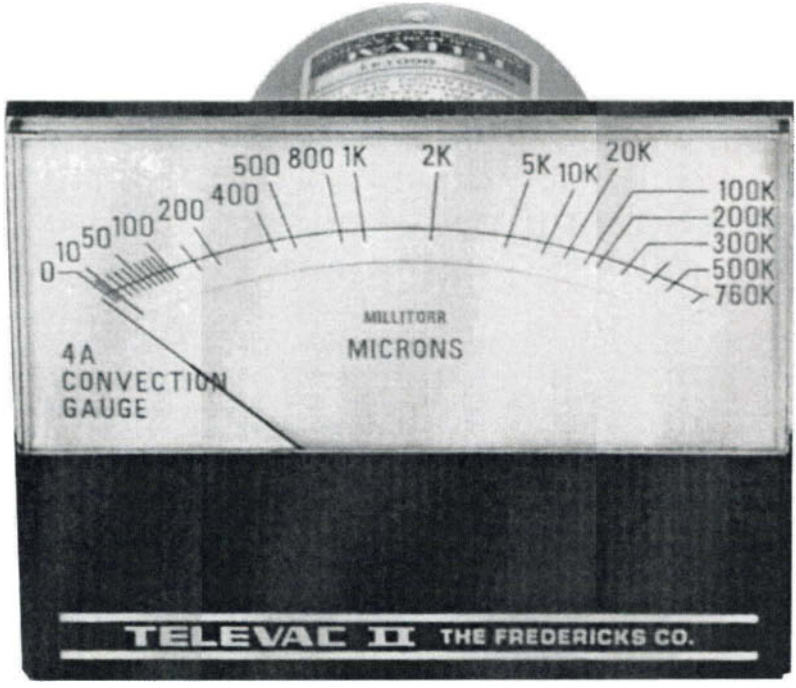
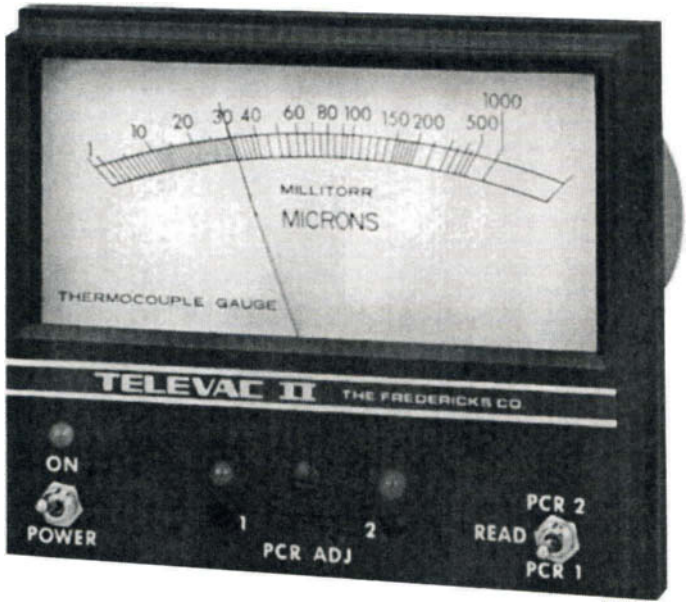


Fig. 22b.



The various types of sensors that can be accommodated by one such modular gauge are as follows:

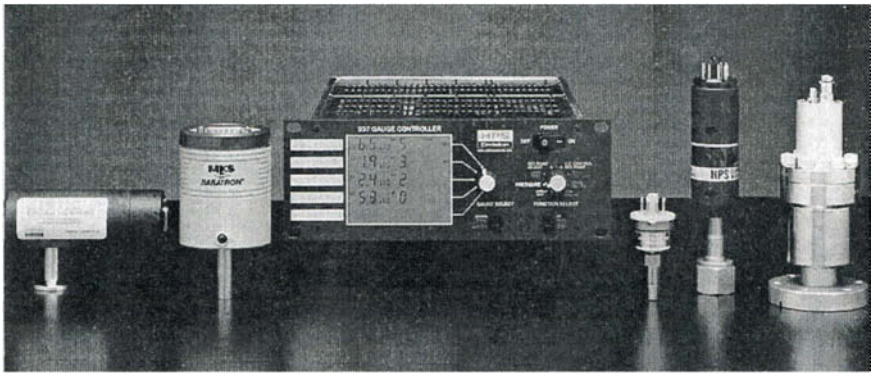
1 to 10 bar	Diaphragm sensor (positive pressure)
1 to 1000 torr	Diaphragm sensor
1, 10, 100, 1000, 10,000 torr	Capacitance diaphragm sensors
1 millitorr to 1000 torr	Convection sensor
1 to 20,000 millitorr	Thermocouple sensors
10^{-7} to 10^{-3} torr	Cold-cathode sensors
10^{-11} to 10^{-2} torr	Cold-cathode sensors
10^{-11} to 10^{-3} torr	Hot-cathode sensors

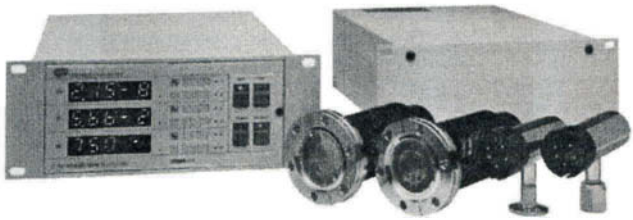
Examples of these combination instruments as sold by several manufacturers are shown in Figure 23a–e.

Fig. 23a–b.

Source: Photo a, courtesy of Varian Vacuum Products; photo b, courtesy of MKS Instruments, Inc.

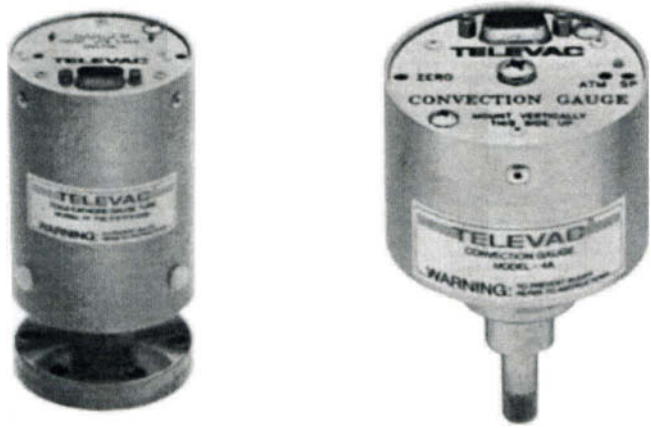
Fig. 23c–e.





Source: Photo c, courtesy of MKS Instruments, HPS Division; photo d, courtesy of Granville-Phillips Company; and photo e, courtesy of Televac.

Fig. 24a, b.



Intelligent Gauges

INTELLIGENT GAUGE TUBES

For some applications, an instrument with a display is not required. In this case an "intelligent" (or "smart" or "active") sensor is used. The sensor contains its own electronics. It is powered from an external source and provides an output signal (usually 0–10 volts or a digital communications interface). This signal then is provided to the user's control system or computer (see Figure 24).

3.1.4

CALIBRATION OF VACUUM GAUGES

The calibration of any vacuum gauge is not a trivial undertaking, but increasingly necessary as vacuum processes become more intricate. Certain processes performed on military end use items require compliance with military standards that include calibration of vacuum gauges with traceability to national standards labs. The capability of obtaining gauges calibrated by national standards labs or labs having traceability to national labs is a relatively recent occurrence.

In contemplating the calibration of a vacuum sensor, it is useful to first evaluate the pressure range that is important to the procedure being performed. The next and perhaps more difficult task is to determine (if possible) the sensitivity of the process to pressure change. These data should allow an estimate of the accuracy necessary to ensure operation within this pressure range. Certain operations such as vacuum melting and simple evaporation may only require knowledge that the operation is taking place at a pressure below a set limit. For that purpose, an ac-

curacy of 10% or even 20% would be sufficient. Other operations involving backfilling into essentially totally exhausted chambers may require a much higher degree of accuracy of approximately 2%. Perhaps the most important feature of any gauge is the repeatability.

One can calibrate a vacuum gauge by measuring it relative to a direct reading standard gauge, or by measuring the gauge at a fixed known pressure.

In the region from atmospheric to about 0.5 torr, the use of U-tube manometers of one sort or another, if corrected for local atmospheric pressure and temperature, can provide a standard that can be used to calibrate other indirect gauges.

In the region from 0.5 torr to 10^{-5} torr, McLeod gauges with calibrated volumes can be used as standards, but care must be taken to eliminate condensable vapors and the possibility of mercury contamination of sensor and system. For years it had been the practice to simply state that the sensor had been measured relative to a trapped McLeod gauge. This practice is no longer considered practical due to the toxic nature of mercury and the difficulty in obtaining accurate reproducible readings of the mercury meniscus. The MDG (spinning rotor gauge) can also be used as a calibration standard, as mentioned earlier.

More recently various national standards laboratories have instituted programs leading to the establishment of pressure standards and calibration services whereby gauges can be certified by reference to these pressure standards [72]. A few companies can now offer NIST-traceable calibration services. Other calibration services are also available [73]. The devices most often used by NIST (National Institute for Standards and Technology) as a secondary calibration standard in the region 10^{-3} torr and lower are the spinning rotor gauge (MDG) and the tungsten filament Bayard-Alpert gauge [41]. Other national labs have used triode gauges of their own design [74,75].

The basis of the standardization process is the production of an accurately known pressure. The static method is the simplest in that a known small volume of a known gas (usually nitrogen or argon) of known pressure measured by a direct reading gauge such as a U-tube is allowed to expand into a large volume. Then all, or in most cases a small volume, of the expanded gas is allowed to expand into a larger volume, and so on, until the desired pressure is obtained. With care, one can obtain known pressures as low as 10^{-7} torr, but more realistically, 10^{-5} to 10^{-6} torr. Gauges can be calibrated at the final pressure or at several points in the expansion process.

The dynamic process is the method preferred and used by NIST [71,72], other national standards labs and some companies. This method is somewhat more complex in that a gas is introduced into a chamber at a constant throughput Q. At the same time gas is pumped out at a constant pumping speed S. At equilibrium $P = Q/S$. The complicated part is keeping Q and S known and constant. The manufacture of such a system is an expensive and time-consuming enterprise and, given the availability of calibrated instruments, largely unnecessary.

REFERENCES

1. J. F. O'Hanlon. A User's Guide to Vacuum Technology (Wiley, New York, 1980), p. 47.
2. H. McLeod. *Phil. Mag.*, **47** (1874) 110.
3. C. Engleman. The Fredericks Co., private communication.
4. Wallace and Tiernan Div., Pennwalt Corp., Bellville, NJ.
5. R. W. Hyland and R. L.Shaffer. Recommended Practices of Calibration and Use of Capacitance Diaphragm Gauge for a Transfer Standard. *J. Vac. Sci. Tech.*, **A9**(6)(1991) 2843.
6. K. R. Spangenberg. *Vacuum Tubes* (McGraw-Hill, New York, 1948) p. 766.
7. S. Dushman. *Scientific Foundations of Vacuum Technique* (2nd ed.), edited by J. M. Lafferty (Wiley, New York, 1962), p. 220.
8. W. Steckelmacher and B. Fletcher. *J. Physics E.*, **5** (1972) 405.
9. W. Steckelmacher. *Vacuum*, **23** (1973) 307.
10. A. Beiman. *Total Pressure Measurement in Vacuum Technology* (Academic Press, Orlando, FL, 1985).
11. Granville-Phillips, 5675 Arapahoe Ave., Boulder, CO 80303.
12. Televac Div., The Fredericks Co., 2337 Philmont Ave., Huntingdon Valley, PA 19006.
13. Granville-Phillips Data Sheet 360127, March 1995.
14. Televac US Patent No. 5351551.
15. O. E. Buckley. *Proc. Nat. Acad. Sci.*, **2** (1916) 683.
16. M. D. Sarbey. *Electronics*, **2** (1931) 594.
17. R. Champeix. *Physics & Techniques of Electron Tubes*, Vol. 1 (Pergamon Press, New York, 1961), pp. 154–156.
18. N. Morgulis. *Physik Z. Sowjetunion*, **5** (1934) 407.
19. N. B. Reynolds. *Physics*, **1** (1931) 182.
20. J. H. Leck. *Pressure Measurement in Vacuum Systems* (Chapman & Hall, London, 1957), pp. 70–74.
21. S. Dushman and C. G. Found. *Phys. Rev.*, **17** (1921) 7.
22. E. K. Jaycock and H. W. Weinhart. *Rev. Sci. Instr.*, **2** (1931) 401.
23. G. J. Schulz and A. V. Phelps. *Rev. Sci. Instr.*, **28** (1957) 1051.
24. Japanese Industrial Standard (JIS-Z-8570), *Method of Calibration for Vacuum Gauges*. Electrotechnical Laboratory, Sakura-mura, Niihari, Ibaraki 305, Japan.
25. J. W. Leck, p. 69.
26. W. B. Nottingham. *Proc. 7th Ann. Conf. on Physical Electronics* (Massachusetts Institute of Technology, Cambridge, MA, 1947).
27. H. A. Steinhertz and P. A. Redhead. *Scientific American*, **206**(3) (March 1962), p. 78.
28. R. T. Bayard and D. Alpert. *Rev. Sci. Instr.*, **21** (1950) 571.
29. O. A. Weinreich. *Phys. Rev.*, **82** (1951) 573.
30. O. A. Weinreich and H. Bleecher. *Rev. Sci. Instr.*, **23** (1952) 56.
31. H. C. Hseuh and C. Lanni. *J. Vac. Sci. Technol.*, **A5** (September–October 1987) 3244.
32. T. S. Chou and Z. Q. Tang. *J. Vac. Sci. Technol.*, **A4** (September–October 1986) 2280.
33. P. A. Redhead. *Rev. Sci. Instr.*, **31** (1960) 343.
34. G. H. Metson. *Br. J. Appl. Phys.*, **2** (1951) 46.
35. J. J. Lander. *Rev. Sci. Inst.*, **21** (1950) 672.
36. P. A. Redhead. *J. Vac. Sci. Technol.*, **3** (1966) 173.
37. J. Groszkowski. *Le Vide*, **136** (1968) 240.
38. L. G. Pittaway. *Philips Res. Rept.*, **29** (1974) 283.
39. D. Blechshmidt. *J. Vac. Sci. Technol.*, **10** (1973) 376.
40. P. A. Redhead. *J. Vac. Sci. Technol.*, **6** (1969) 848.
41. S. D. Wood and C. R. Tilford. *J. Vac. Sci. Technol.*, **A3** (May–June 1985) 542.

42. C. R. Tilford. *J. Vac. Sci. Technol.*, **A3** (May–June 1985) 546.
43. P. C. Arnold and D. G. Bills. *J. Vac. Sci. Technol.*, **A2** (April–June 1984) 159.
44. P. C. Arnold, and S. C. Borichevsky. *J. Vac. Sci. Technol.*, **A12** (March–April 1994) 568.
45. D. G. Bills. *J. Vac. Sci. Technol.*, **A12** (March–April 1994) 574.
46. C. R. Tilford, A. R. Filippelli, et al. *J. Vac. Sci. Technol.*, **A13** (March–April 1995) 485.
47. P. C. Arnold, D. G. Bills, et al. *J. Vac. Sci.Technol.*, **A12** (March–April 1994) 580.
48. ETI Division of the Fredericks Co. Gauge Type 8184,
49. T. A. Flaim and P. D. Owenby. *J. Vac. Sci. Technol.*, **8** (1971) 661.
50. J. F. O'Hanlon, p. 65.
51. F. M. Penning. *Physica*, **4** (1937) 71.
52. F. M. Penning and K. Nienhauis. *Philips Tech. Rev.*, **11** (1949) 116.
53. P. A. Redhead. *Can. J. Phys.*, **36** (1958) 255.
54. J. P. Hobson and P. A. Readhead. *Can. J. Phys.*, **33** (1958) 271.
55. NRC type 552 data sheet. Varian Associates, Lexington, Mass., February 1973.
56. N. Ohsako. *J. Vac. Sci. Technol.*, **20**(4) (April 1982) 1153.
57. D. Pelz and G. Newton. *J. Vac. Sci. Technol.*, **4** (1967) 239.
58. R. N. Peacock, N. T. Peacock, and D. S. Hauschulz, D.S. *J. Vac. Sci. Technol.*, **A9** (May–June 1991) 1977.
59. E. Drubetsky, D. R. Taylor, and W. H. Bayles, Jr. *Amer. Vac. Soc. New Eng. Chapter*, 1993 Symposium. Unpublished paper.
60. B. R. Kendall and E. J. Drubetsky. *J. Vac. Sci. Technol.*, **A14** (May–Jun 1996), p 1292.
61. M. Ono, K. Hirata, et al. Quartz Friction Vacuum Gauge for Pressure Range from 0.001 to 1000 Torr, *J. Vac. Sci. Technol.*, **A4**(3) (May–June 1986) 1728.
62. O. E. Meyer. *Pogg. Ann.*, **125** (1865) 177.
63. J. C. Maxwell. *Phil. Trans. R. Soc.*, **157** (1866) 249.
64. F. T. Holmes. *Rev. Sci. Instrum.*, **8** (1937) 444.
65. J. W. Beams, J. L. Young, and J. W. Moore. *J. Appl. Phys.*, **17** (1946) 886.
66. J. K. Fremery. *Vacuum*, **32** (1946) 685.
67. NIST, Vacuum Calibrations Using the Molecular Drag Gauge, Course Notes, April 15–17, 1996. National Institute of Standards and Technology.
68. S. Dittman, B. E. Lindenau, and C. R. Tilford. *J. Vac. Sci. Technol.*, **A7** (November–December 1989) 3356.
69. K. E. McCulloh, S. D. Wood, and C. R. Tilford. *J. Vac. Sci. Technol.*, **A3** (1985) 1738.
70. G. Cosma, J. K. Fremerey, B. Lindenau, G. Messer, and P. Rohl. *J. Vac. Soc. Technol.*, **17** (1980) 642.
71. C. R. Tilford, S. Dittman, and K. E. McCulloh. *J. Vac. Sci. & Technol.*, **A6** (1988) 2855.
72. S. Dittman. NIST Special Publication 250–34 (1989). National Institute for Standards and Technology, Gaithersburg, MD 20899.
73. National Conference of Standards Laboratories, Boulder, CO. National Conference of Standards Laboratories, 1800 30th St., Suite 305B, Boulder.
74. M. Hirata, M. Ono, H. Hojo, and K. Nakayama. *J. Vac. Sci. Technol.*, **20**(4) (April 1982) 1159.
75. H. Gantsch, J. Tewes, and G. Messer. *Vacuum*, **35**(3) (1985) 137.
76. S. Dushman and A. Young. *Phys. Rev.*, **68** (1945) 278.
77. S. Wegener and C. Johnson. *J. Sci. Instr.*, **28** (1951) 278.
78. L. Riddiford. *J. Sci. Instr.*, **28** (1951) 375.
79. G. J. Schulz. *J. Appl. Phys.*, **28** (1957) 1149.
80. Electron Technology Inc. Div., The Fredericks Co. Private communication.

CHAPTER 3.2

Mass Analysis and Partial Pressure Measurements

Laszlo V. Lieszkovszky
GE Lighting

3.2.1 OVERVIEW AND APPLICATIONS

There are numbers of cases where more information is required than a total pressure gauge can provide. Often the real question is the composition, the change in composition of the sample or the partial pressures of some constituents.

When the sample pressure is about atmospheric, one can choose from several available techniques. Mass analysis is only one of them. At low pressures, or when the sample size is small, however, partial pressure analyzers (PPAs) face little competition. This section provides a general description of PPAs, identifies some areas of application and defines some terms used for characterizing the instrument. Following sections describe the main parts of a PPA: the inlet system, the ion source, the analyzer, and the detector. The inlet system converts the sample pressure to a vacuum level acceptable for ionization and directs the sample toward the ion source. The ion source creates ions from neutral species so that they can be separated by electric or magnetic fields. The analyzer separates the ions according to mass-to-charge ratio (M/Q; also called "mass number"). In the case of spatial separation ions created at the same time are made to travel on various trajectories according to mass-to-charge ratio. Ions with trajectories leading to the detector provide an ion signal, others are lost. Separation in time is another

ISBN 0-12-325065-7
$25.00

method: ions created at the same time are traveling at different speeds according to mass-to-charge ratio, reaching the detector at different times. Finally, the detector converts the ion signal into easily measurable electric current. Advantages and limitations of the choices are given and some basic principles discussed.

Mass analysis — like the most sensitive total pressure gauge, the ion gauge — uses ions (in the vast majority of cases, positive ions) formed from the sample atoms or molecules to provide a signal. Whereas the ion gauge measures all ions without distinguishing them from each other, the analyzer section in a mass spectrometer will separate the species based on their mass-to-charge ratio (M/Q); e.g., a double-charged ion of the most common argon isotope $^{40}Ar^{2+}$ has a nominal mass number $M/Q = 20$. The mass-to-charge ratio as the primary coordinate of the measurement represents an important challenge of mass spectrometry: a signal at a given mass-to-charge ratio may not uniquely identify a species. (M/Q = 20 can be $^{20}Ne^+$; $M/Q = 28$ can be CO^+, N_2^+, Si^+, or $C_2H_6^+$; $M/Q = 44$ can be CO_2^+ or N_2O^+, etc.)

The issue can be even more complex: Ion species from one constituent appear at various mass numbers not only because of isotopes, but also because various species are produced by the ionization process itself. (Referred to as *fragmentation*: e.g. CO_2 may produce ions such as C^+, $M/Q = 12$; CO^+, $M/Q = 28^+$; O^+, $M/Q = 16$; O_2^+, $M/Q = 32$; and CO_2^+, $M/Q = 44$ "parent ion".)

Tables 1–5 contain normalized spectra of different substances (magnitude of most abundant peak = 100). Atomic or molecular weight (M. Wt), relative sensitivity of the most abundant peak compared to $^{40}Ar^+$ (Rel. Sens.), and isotope ratios are also listed. Spectra and relative sensitivity values are given for illustrative purposes only — they depend on instrument design and settings. (Values given measured at 70 eV electron energy, $\Delta m = 1$ constant peak width mode of a quadrupole mass spectrometer.)

Nonetheless, some experience, isotope tables, and standard spectra usually help. Despite (or in some cases because of) these intricacies, mass analyses have been successfully employed in a number of applications ranging from air to vacuum.

3.2.1.1 Applications

Air pollutant measurement, when the composition of pollutants is unknown, may present a problem for gas chromatography, that would require different columns, detectors, etc. A freezeout–thermal analysis mass spectrometric technique is capable of identifying impurities such as various freons and hydrocarbons at sub-ppm levels [1]. In searching for abnormal or possibly toxic pollutants, the use of cluster analysis program and standard library spectra may be of help in their identification or at least classification [2]. The calibration for measuring unstable at-

mospheric constituents such as ozone requires special procedures [3], and so do the mass spectrometers flown in balloons to measure altitude profiles of standard constituents [4].

Environmental characterization and monitoring of hazardous constituents in the field typically involves a gas chromatograph–mass spectrometer combination (GC-MS), where the GC is used for coarse separation, the MS for fine separation and detection. Portable instruments are available for site characterization using quadrupole analyzers or the quadrupole ion trap [5]. (See end of Section 3.2.4.1 about the quadrupole ion trap.)

Biochemical analysis includes fermentation, environment monitoring and control, enzyme kinetics for determining velocity constants and reaction mechanisms, plant physiology–gas metabolism measurement, etc. Membrane inlet mass spectrometry has become a well-established technique capable of on-line monitoring of gases and small-molecular-weight organic compounds dissolved in water. If organic molecules are leaked into the vacuum chamber through a membrane, the signal is the superposition of two transients, one reflecting the membrane response and the other reflecting the adsorption/desorption processes on the vacuum walls. Therefore the design of the inlet system is critical [6]. Multicomponent analysis with multichannel sampling from the gas and the liquid phase is capable of monitoring or controlling fermentation processes [7,8,9]. In the research of plant physiology a suitable sampling unit enables *in vivo* measurements without influencing gas metabolism. Further example is the analysis of the effects of agricultural chemicals on crop plants [10].

Isotope ratio analysis is a classical field of magnetic sector mass spectrometry but quadrupole type is also used [11]. It is still constantly refined to measure isotopic effects in noble gases for study of natural or laboratory-induced nuclear processes [12], or the study of isotopes of light elements such as hydrogen, carbon, oxygen, sulfur, and nitrogen in terrestrial or extraterrestrial geologic substances [13]. Quadrupole-type instruments have been tested for possible nuclear safeguard applications requiring mobile uranium isotope measurement. With proper settings, a resolving power of 233 and an abundance sensitivity of 3500 were achieved [14]. Clearly, isotope ratio analysis is an area where the above-mentioned complications regarding spectrum interpretation may arise (e.g., oxygen isotopes ^{16}O, ^{17}O, ^{18}O measured at $M/Q = 32, 33, 34, 35, 36$ results in a set of linear equations—the molecules $^{16}O^{18}O$ and $^{17}O^{17}O$ both contribute to the peak at $M/Q = 34$) [15].

Gas composition and purity analysis of bulk gas samples usually use a batch inlet to determine gas composition or trace impurities. Gas composition analysis usually requires careful calibration procedures to assure quantitative results [16] primarily due to the fractionating versus nonfractionating gas flow in the inlet or the vacuum system [18]. Contaminants down to about 10 parts-per-million (ppm) range can be routinely measured. Analysis of electronic gases presents spe-

cial challenges due to the corrosive, toxic and reactive nature of the gases involved. The sample-related problems are loss of minor constituents and addition of contaminants as a result of adsorption/desorption processes, reactions with instrument components possibly leading to performance degradation, ion molecule reactions and release of gases by electron, ion, or neutral molecule bombardment of surfaces [19]. Techniques to minimize these effects include fast gas flow in the sample-handling system, differential pumping of ionizer and mass filter to maintain high pressure in the ionizer while keeping the analyzer pressure low and beam chopping with lock-in detection to be able to differentiate between sample signal and background [20]. Below the 10 ppm level, special procedures are required, such as the preconcentration techniques either using freezeout (He in N_2, and H_2 in Ar) [21] or chemical methods (measurement of trace impurities in oxygen by chemically removing the oxygen using NaK alloy) [22] or in some case increase of the high-pressure operating limit of the spectrometer (CO_2 in N_2) [23]. When trying to achieve parts-per-billion level, a combined gas chromatograph–mass spectrometer system has been used successfully for trace gas detection in high-purity gases used for semiconductor fabrication (H_2, N_2, CO, Ar, SiH_4, N_2O, etc.) [24]. The most sensitive method available uses the preferential ionization of impurities by charge transfer mechanism in the high-pressure chamber of the atmospheric pressure ionization mass spectrometer (APIMS) allowing contaminant measurement of ultrapure gases at the parts-per-trillion levels — representing the state of the art in trace gas analysis today [25].

Leak detection is a built-in feature of many contemporary partial pressure analyzer systems. Finding leaks or verification of vacuum tightness is an omnipresent problem of vacuum work and there are cases where traditional helium leak detectors prove to be an inferior solution to mass analyzers. In large and complex vacuum systems such as the Gas Centrifuge Enrichment Plant/Ohio consisting of 1000 miles of piping and 10^8 L vacuum volume, the difficulty of hooding and permeation of helium through O-rings makes traditional leak testing a very inefficient method. Instead, a special air leak test cart has been built that can be connected to large number of ports of the system and enables isolating leaky sections by manipulating valves in the area [26]. Leak testing of hermetically sealed products can also be realized for quality checking at a rate of 6000 pieces/day (leak testing of capsules filled with freon) [27].

Medical applications are exemplified by diagnosing certain diseases by observing the response of the respiratory system to gases of various solubilities in the bloodstream or monitoring the concentration of anesthetic gases (such as halothane, penthrane, nitrous oxide) for the safety of the patients and the operating staff [28,29]. Gas probes for blood gas monitoring can be of the membrane type — an *in vivo* catheter or skin transducer measuring the gas diffusing through the skin. Requirements are fast response time to sample changes, stability in performance [30,31], and thorough understanding of the inlet system behavior [32].

Metallurgical applications have to compete with infrared absorption instruments for the determination of carbon monoxide and carbon dioxide. The advantages offered by the mass analysis technique for effluent gas analysis, vacuum metallurgical or steelmaking process control are the speed and complete and continuous analysis of multicomponent mixtures by one instrument (including direct determination of nitrogen) [33]. The difficulty of the interpretation of overlapping peaks of two important gases, nitrogen and carbon monoxide ($M/Q = 28$) appears again, but the dependence of the covariant matrix on electrical settings of the spectrometer can help to resolve specific problems and achieve maximum accuracy [34]. Mass analysis has been used to analyze the finished product as well, for example, gas contents of steel or other metal products are measured by melting the samples and simultaneously determining carbon monoxide, nitrogen, and hydrogen in the few-parts-per-million range [35].

Outgassing and permeation rate measurements routinely include a mass analyzer to determine the principal gas species contributing to the measured outgassing rate of various metals and ceramics [36]. In the measurement of gas emissions from plastics (Plexi, polystyrol, Teflon) [37] and elastomers (Viton, Perbunan, Silastomer, and other organics sometimes used in vacuum work such as Araldite, Nylon, polyethylene, PTFE) [38], knowledge of the mass numbers of the outgassing species is at least as important as the rate of outgassing — although with the exception of polyethylene and PTFE the main component in most cases is water.

Plasma-related applications of mass analysis range from identification of ion species for the study of gaseous plasmas (NeAr and NeKr) [39], sampling of glow discharges, investigation of the gaseous electronics aspects of sealed CO_2 laser operation [40] to partial pressure or flux analysis of fusion devices (Tokamak Fusion Test Reactor/New Jersey) [41], (Axially Symmetric Divertor Experiment/ Germany) [42] (Joint European Torus) [43]. Mass spectrometers are used here as a diagnostic tool for optimizing and process-monitoring glow discharge wall-conditioning procedures in the various cleaning scenarios such as pulsed discharges or DC-glow discharges. (Fusion experiments themselves apply strong magnetic fields, making it impossible to use standard analyzers close to the plasma vessel.) The two basic categories of such investigations are (1) partial pressure analysis, sampling of volatile reaction products representative of the integrated effects of plasma–wall interactions and (2) plasma flux analysis, sampling neutral plasma particles streaming directly into the spectrometer [44]. Although the latter requires more sophisticated vacuum connections and coupling between the plasma and the spectrometer, it allows energy analysis of the plasma species. Since ion sources are operated usually near the high-pressure limit (10^{-4} torr, $\sim 10^{-2}$ Pa), hydrogen conditioning can be employed to reduce residual gas levels [45]. Long-term changes in sensitivity were reported to be reduced by the use of tungsten filament (as opposed to rhenium) and operation at a constant tempera-

ture [46]. Peak overlapping in hydrogen, deuterium, tritium plasmas (H_2, D_2, T_2, HD, HT, DT, ^{3}He, ^{4}He all appearing in the M/Q range 1–6) can be resolved by using the dependence of appearance of species on instrument settings (electron energy setting offering separation of helium and hydrogen — and in case of quadrupoles — field axis potential setting offering separation of molecular ions from fragment ions) [47].

Process control may be the single largest area of application. Quality of the seven main components of natural gas have been monitored to control mixing stations using mass analysis much faster or more thoroughly than alternative techniques (gas chromatography, infrared). The general features that make mass analyzers attractive as a process technique is the speed of analysis, versatility — capable of being adapted to a wide range of analyses with little development work involved (as opposed to gas chromatography), ruggedness (of specially built instruments) and reliability. Regarding cost, mass spectrometric technique is generally more expensive than existing specific instrumentation and there are some limitations in its specificity (e.g., ability to distinguish between isomers) [48].

There is one area, however, where its widespread use is unquestionable — semiconductor processing and thin-film deposition in general. The applications of mass analysis include monitoring or closed-loop control of sputtering and chemical vapor deposition (CVD) processes. End point detection in plasma-etching processes is one of the most important features. In sputter gas atmospheres, the typical pressure approaches 10^{-2} torr ($\sim$1 Pa) commonly requiring a pressure reduction step (orifice) and choking of the vacuum pump to minimize effects of preferential pumping. *In situ* calibration of the sensitivity factor of the mass spectrometer enables identification of the contaminants that may have considerable effect on semiconductor properties [49]. Oxidizing constituents of the sputter atmosphere such as water and oxygen generally must be monitored at the ppm level. The benefit of connecting the sputter gas atmosphere without pressure reduction step is the avoidance of large errors in the measurement of background gases [50]. Recent approaches expose the ion formation region directly to the high process pressure while the filament and the mass analyzer kept at high vacuum are connected through small orifices to the ion formation region [51]. Removing hot filaments from the ionizing region reduces process gas–filament interaction, and reduces the chance of instrument performance degradation (e.g., metal deposit on insulators when using WF_6). Also, stability is improved, since pumping variations have very little effect on sample pressure, while variations in sample pressure have little effect on electron production or ion separation efficacy. The contamination resistance of the analyzer section and its abundance sensitivity can be further improved by the use of a triple-filter arrangement where the "pre" and "post" filters are driven by RF (radio frequency voltage) only [52]. Real-time monitoring of reaction products provides an opportunity to dynamically optimize process parameters. The basis for end point detection is the rate of evolution of

reaction products from the reaction between the substrate and etch gas mixture (such as SiF_4 evolution while etching SiO_2 with perfluorocarbons, or as the formation of CO_2 that provides a signal for end point detection of photoresist strip with freon-free oxygen plasma) [53].

Mass spectrometers of the quadrupole type are also used as rate sensor in the feedback loop in UHV deposition processes (for example, electron gun evaporation in molecular beam epitaxy (MBE) systems). Noted advantages are the large sensitivity, large bandwidth, and multiplexing capability (monitoring several sources simultaneously by switching between different mass numbers), large dynamic range for rate measurement, and low maintenance needs. Limitations are the possible instability of this complex instrument, the fact that a separate flange is required where the instrument has to be inserted relatively deep in the chamber. In addition, the viewing angle may not be wide enough, the point of deposition may be farther away from the point of measurement, and the stabilization time of the mass spectrometer at the new mass number may be long in multiplexer mode [54,55]. For identification of vapor species present and particle density measurement in an MBE apparatus, a time-of-flight technique has been developed [56].

Mass spectrometric contamination control of process gases used in electronic device fabrication requires the solution of several issues related to reactive and corrosive nature of these gases. The memory effects, surface interactions, and reactions in the ion source or analyzer region and pumping system must be minimized along with the possible ion–molecule reactions. A proposed system consists of a molecular beam inlet system with a mechanical chopper and lock-in detection with a combination of turbomolecular and two-stage He cryopump — the 14 degree K surface enclosing the ion source is needed to assure low background [57].

A technique called *electron impact emission spectroscopy* must be mentioned here. It allows selective monitoring of partial pressures of species by detecting their optical emission lines as a result of excitation by electron bombardment. This approach is used in some partial pressure controllers as well as in certain vacuum deposition monitors. The simple calibration procedure, high stability, and its detection system's insensitivity to electrical noise or charged particles (such as in electron beam evaporation sources) offer advantages over mass spectrometric approaches [58,59,60].

Sealed gas sample analysis by breaking, piercing, or melting the sample are common in various industries such as microelectronics, glass, or lighting products industry as a quality check or research/development tool. In semiconductor devices moisture is generally the biggest concern, which may provide surface layers, unwanted electrical conductivity, and undesirable intermetallic growth, eventually leading to corrosion and failures. Even though nondestructive methods are available (surface conductivity sensor, capacitance ratio test), and destructive optical methods (such as infrared spectroscopy or more recently tunable diode laser

absorption and intracavity laser spectroscopy) exist [61], mass spectrometry with electron impact or chemical ionization is preferred as the tool for package technology development or failure analysis [62]. The reasons are primarily its sensitivity and ability to provide information on species beside moisture that may be present due to poorly cured resins, use of cleaning solvents, desiccants, etc. Some laboratories routinely monitor 32 mass numbers [63]. The principal problem concerning the accuracy of the measurement arises from wall effects — adsorption/desorption of moisture and other nonvolatile components. Of primary importance, therefore, is proper "burst" calibration, which imitates the puncturing of the package and the subsequent release of its gaseous content either by moisture standard packages [64] or using controlled moisture delivery apparatus with properly operated valve arrangements [65,66]. Another approach tries to break away from the wall effect problem by converting the moisture chemically and quantitatively to generate simple nonadsorbing gas products. Reagents used are sodium or calcium hydride giving hydrogen, magnesium nitride or sodium amide giving ammonia, and calcium carbide giving acetylene [67].

Measurement of gases entrapped in glass is of interest for the glass industry. The analytical procedure is further complicated by the effect of mechanical shock when breaking open the glass (desorption) and the surface effects related to the freshly broken glass [68]. Although little adsorption of either nitrogen or oxygen was noted by freshly broken glass surface, the adsorption of SO_2, CO_2, and moisture can be a significant source of error. The observed sudden drop of CO_2 signal is accompanied by the appearance of CH_4 — a conversion reaction has been proposed as an explanation [69]. The two basic schemes used to analyze such samples are the so-called "static" when the vacuum pumps are valved off and the "dynamic" when the vacuum pumps are on [70].

Analysis of electric light sources raises similar concerns. To extend the sensitivity into the sub-ppm range, a cryogenic preconcentration technique can be used for incandescent lamps [71]. In this industry moisture measurement is also of great interest, being the cause of blackening in incandescent lamps, possibly short life in halogen lamps, and a cause of starting problems in discharge sources. Several sampling techniques have been described to determine the moisture content of the hygroscopic mixed iodide ampoules used in dosing of metal-halide discharge lamps [72]. In these products as well as in sealed electrical contacts (Reed relays) the sealing process itself will change the contaminants through various reactions at the high temperature but the measured impurities will provide a basis for troubleshooting (e.g., CO, CO_2 indicative of O_2, etc.) [73,74]. Measurements of gas fill of laser fusion target pellets is a case of extremely small gas quantity determination ($<10^{-9}$ L) where a closed volume mass spectrometric technique has been used [75]. Commercial and custom-made systems are available for analysis of trapped gas samples (IC packages, lamps, inclusions in glass) featuring detection limits in the low-ppm range, and sample volumes down to 10^{-9} liter [76].

In case of glass, an alternative technique to mechanical breakers and piercing devices is to melt the sample in a furnace unit. In case of metals, this is practically the only way to measure the amount of impurity gases trapped. Measuring the decomposition products recorded with heating time provide information on glass types suitable for certain applications [77] or aid in the development of metals [78,33].

Vacuum applications are broadly defined; the largest portion being what is often referred to as residual gas analysis. Indeed, mass analyzers of limited mass range (1 to a few hundred amu) are also called *residual gas analyzers* (RGAs), implying the typical application. Such instruments can be found on a majority of vacuum systems, since they significantly enhance the ability to improve basic pumpdown conditions or allow understanding of phenomena in vacuum. As soon as 10^{-4} torr ($\sim 10^{-2}$ Pa) has been reached, an RGA can indicate the possible existence of leaks, enables monitoring the composition and limiting factors of the ultimate vacuum reached, and optimization or automation of processes. He-tuned RGAs may prove to be a simpler and more cost-effective way to check vacuum tightness than dedicated leak detectors [79]. Mass analyzers are indispensable tools in analyzing the behavior of vacuum chambers and small systems [80,81] or examining pump performance and effects of operational conditions of turbomolecular pumps [82]. Unfortunately, since the analyzer is part of the system its presence influences the residual gases that would be present without it and this has to be taken into account [83]. These applications commonly use low-resolution analyzers and interpretations must be based on assumptions especially in case of overlapping peaks. Only high-resolution residual gas analysis can reveal, for example, that the common $M/Q = 28$ peak on a typical liquid-nitrogen-trapped oil diffusion system is in reality a multiplet (CO, N_2, C_2H_4) and their ratio is markedly different on oil versus mercury diffusion pumped systems [84]. Large vacuum systems such as the Synchrotron Radiation Source [85] or space simulation facilities [86] also rely on RGAs for analysis or monitoring of facility conditions and identification of contamination products.

Various other applications are reported in basic or applied research such as the study of chemical reactions by means of coupling a microreactor to the mass spectrometer [87] or the study of reaction kinetics in cases where practical reasons prohibit the use of absorption spectroscopy (e.g., the case of high-melting-point solids that are insoluble in most solvents such as tungsten halides) [88]. Coupling the mass spectrometer to a thermoanalyzer will provide information on the gases evolving during the heating process—coupling combinations have been proposed for temperatures up to 1400°C in vacuum and in gas atmospheres as well as ways to reduce ion source contamination using cryosurfaces [89]. Pyrolysis study of materials—among others, study of degradation of PTFE (polytetrafluoro-ethylene) [90]—is another example of controlled heating/mass analyzer combination. Investigation of flames, understanding the processes in

combustion engines often uses modulated molecular beam technique to allow the detection of highly reactive particles, radicals or condensable species [91]. Study of neutral and charged particles from a transient event where the exact timing and position are not known requires a fast-responding technique with easy triggering. Time-of-flight mass spectrometry has been used for the electrical breakdown of insulating surfaces (thin polymer films) [92]. Investigation of the lifetime, laser output, and gas composition changes of gas lasers has been using various techniques such as spectroscopic and chromatographic analysis, but the mass spectrometric technique has the advantage of analyzing many components simultaneously in real time with the disadvantage of depletion of gas from the laser tube. To minimize gas losses, a special inlet system such as an extremely-low-conductance on/off valve with zero dead volume may be necessary [93].

Mass analysis using MS–MS (mass spectrometer–mass spectrometer), triple quadrupole, or ion traps are widely used in research and routine organic analysis (such as the ion trap as a detector of chromatographs) but these are beyond the scope of partial pressure measurement related to vacuum technology.

3.2.1.2 System Performance

The performance of the mass analyzer system is evaluated based on a number of characteristics. Which is important, depends largely on the application. For quantitative measurement, a proper calibration procedure is necessary. Recommended practice and methods of calibration as well as definition of the terminology can be found in a recent update of the AVS Standard 2.3-1972 (1991) [94]. Some important considerations are listed here:

Mass range: The range between the minimum and maximum mass numbers, over which an instrument operates.

Sensitivity: The ratio of change in ion current, or peak height due to the corresponding change in partial pressure of the particular species (A/torr or A/Pa). The main factors influencing sensitivity are the emission current, the resolution setting of the analyzer, and the gain of the multiplier if used (see Sections 3.2.3–3.2.5).

Abundance sensitivity: A term frequently used in isotope measurements; referring to the suitability of the instrument to measure rare isotopes adjacent to abundant ones. It is the ratio of the peak height of the abundant isotope at M/Q and the smallest detectable ion current of the rare isotope at the subsequent mass number, $M/Q + 1$.

Values in the range of 10^3–10^6 are typical. This figure of merit is related to the resolving power and the sensitivity of the instrument. *Tailing contribution* is another term to describe the residual signal from a major peak at the preceding and following mass number from the center of the major peak.

Linearity: Sensitivity may be independent of pressure only in a limited range. Linearity measures the extent to which the change in output signal is proportional to the corresponding change in partial pressure. Typically PPAs exhibit sensitivity independent from pressure within ±15% up to 10^{-5} torr ($\sim 10^{-3}$ Pa), above which the sensitivity decreases with pressure. This pressure is referred to as the upper operating limit. Above this pressure ion molecule collisions result in increasing ion losses and the sensitivity will gradually approach zero. The linear range of a quadrupole instrument depends on its construction but also the settings of ion source and analyzer parameters—ion space charge effects can cause significant non-linearity even at pressures orders of magnitude lower than the upper operating limit. In certain cases the sensitivity may increase with pressure before the sharp drop at the upper operating limit [95].

Resolving power: is the measure of the ability of the instrument to separate ions according to their M/Q ratio. The ratio of the width of a peak (Δm, at 10% of its height in amu units) and the mass number (m, amu) is the conventional definition of resolution. ($R = m/\Delta m$; $R = 200$ means that the peak width reaches 1 amu at mass number 200.) Resolving power of magnetic instruments tends to be constant along the mass range, while for quadrupoles it often varies proportionally with m (depending on electrical control settings).

Another way to measure the extent of peak separation is simply to use the peak width, Δm (at 10% of its height), which is especially well suited for quadrupoles. (Quadrupoles are often used in a way that peak width, Δm, is constant over the entire mass range.) It is to be noted, that as resolution increases, peak width decreases—a smaller number for Δm corresponds to higher resolution.

Scan rate: The speed at which the information on mass analysis is available (amu/s). There is a tradeoff between the scan rate and the noise level (random fluctuation in the output signal unrelated to the partial pressure change). Use of multipliers enable much faster scan rates at the same signal–noise ratio.

Stability: The change in sensitivity over a certain period of time. Contamination in the source, the analyzer and the electron multiplier, or shifts in electric supply voltages are the major causes of degradation of sensitivity (and also resolving power) with time and use [95].

The following sections describe the components of a mass spectrometer system. These are the inlet system, the ion source, the analyzer, and the detector.

3.2.2 INLET SYSTEMS

The mass analyzer section typically requires vacuum level below 10^{-5} torr ($\sim 10^{-3}$ Pa). The ion source region may operate at a pressure of one or two orders

of magnitude higher than that using special designs, but still in a large number of cases some means of sample pressure reduction is necessary. One way to categorize the different inlet systems is according to the sample pressure range, as follows.

If the sample pressure is less than 10^{-5} torr ($\sim 10^{-3}$ Pa), such as the residual gas analysis of UHV systems, or UHV thin-film deposition processes, the analyzer may be simply immersed into the vacuum chamber or connected to a large-diameter flange, possibly with an isolating valve. Immersion into the chamber assures that especially with an open-type ion source the atmosphere of the chamber, instead of an isolated pocket in a tube where filament-related or wall effects may be more pronounced, is monitored. The possibility to valve the analyzer off can assure that the analyzer is maintained at vacuum in case of temporary operations at higher pressures or the cleanliness of the analyzer can be preserved during short use of noninert gases.

If the pressure is within the range of 10^{-5}–10 torr ($\sim 10^{-3}$–10^{3} Pa) and its min:max ratio is 1:10 such as the case of certain process-monitoring applications, a single, appropriately dimensioned orifice can be used. This is basically a fixed aperture with an optional isolation valve. The flow through a small orifice typically used (such as 0.05 mm) will be molecular below 1 torr ($\sim$100 Pa) [94], therefore the flow into the analyzer will be of the same type as the outflow toward the vacuum pump. (The pumping speed for gases of differing molecular weights is assumed to follow ideal molecular flow laws.) Fractionation [96,97] of the sample will not alter the composition unless the sample gas quantity is too small and lighter components are depleted from the reservoir before the orifice. (In the molecular flow regime the throughput is inversely proportional to the square root of the molecular weight.)

If the max. pressure is below 10 torr ($\sim 10^{3}$ Pa), but the min:max ratio is higher than 1:10, then a parallel arrangement of several orifices of varying size with individual shutoff valves can cover a pressure range wide enough on the expense of complexity of the inlet system. It is very common to find a large-diameter high-vacuum shutoff valve parallel with an orifice to perform residual gas analysis, leak detection, and process control using the same instrument. Another alternative is to use variable leak valves that will cover a pressure range several orders of magnitude wide with a single valve. Even though this is a simpler solution, there are cases where the benefits of the high stability of a built-in fixed-size orifice, which provides reproducible pressure ratio, outweigh the benefit of freely adjustable pressure in the analyzer. (Some leak valves exhibit large hysteresis and creep; their leak rate depends on the direction of adjustment (upward or downward) with a step when changing direction (hysteresis) or the leak rate slowly changes with time at a given setting (creep).)

If the sample pressure exceeds 10 torr ($\sim 10^{3}$ Pa), the simplest solution is again the application of an adjustable leak valve or a fixed orifice. However, at such high sample pressures the flow in the leak valve or orifice will be certainly in the

viscous range while the flow out of the ionizer will be in the molecular range. Since the viscous flow will carry each component uniformly without altering the gas composition, while the molecular flow will preferentially carry out the lighter components, fractionation will occur. This must be accounted for in calculating the gas composition, or calibration must be performed with a mixture, whose composition is close to the sample composition, at similar sample and analyzer pressures. In that case the effect of fractionation will be "buried" in the calibration factor together with the various ionization and detection probabilities of the different species.

In place of a simple orifice or a leak valve, a fritted leak or a capillary may be used to reduce the sample pressure. The end of the capillary may lead directly into the source, minimizing contamination, but its slow response (usually above 10 seconds, maybe minutes) makes it unpractical in many cases. Fractionation-related problems must be dealt with as described earlier (calibration or calculations), but capillaries do not present clogging problems to the extent associated with orifices. (An orifice is an extremely "short tube," therefore its diameter must be much smaller than the diameter of a long capillary of comparable conductance.) In case of fritted leaks, where many parallel leaks exist, clogging is usually a slow process, but often observed. Therefore filtering of particulates is highly recommended in any case.

Another way to reduce pressure is to expand a small quantity of the high-pressure sample (e.g., 1 cc STP) into evacuated larger-volume chambers (a few liters). To minimize surface effects and losing trace species several expansion and pumpout cycles may be repeated. Such a system is referred to as a *batch inlet.* The need for repeated expansion and pumpout cycles makes it rather unpractical in case of environmentally dangerous gases. Correction for fractionation may be necessary depending on orifice size and expansion chamber pressure [16].

A widely used method to avoid fractionation is the use of a capillary, whereas the sample flows from the high-pressure side of the capillary toward a "Tee," whereas its pressure is reduced by two to three orders of magnitude. The other side of the "Tee" is connected through an orifice to the ionizer, while the third end is going through a restriction toward a vacuum pump. This continuous renewal of the gas sample along with the orifice operating in the molecular flow range makes it attractive, since through this arrangement much higher sample flow can be attained than if the capillary would lead directly into the ionizer. Response times are in the order of seconds or less. Several capillaries can be connected to a selector valve, forming a multicapillary inlet system saving evacuation and pumping time and enabling fast switching between various samples. An alternative way to multisampling is the use of zero dead-volume continuous-flow rotary valves (like column-switching valves used in chromatography) in the standard piping before the pressure-reducing element. A "Tee" element at the inlet port of the capillary transfer line can assure continuous venting, while the capillary leads directly into

the analyzer. This provides the highest venting flow, since the capillary is not restricting it, but at the expense of slower response time. (For attaining the same pressure reduction through a single capillary, a smaller conductance capillary must be used than in the case of the two-step pressure reduction — where a larger capillary and an orifice are used.) Again, since the capillary operates in the transient flow regime fractionation will take place and calibration with gas of similar composition to the sample is necessary. In case of adsorbing or easily condensing sample, heated tube and heated capillary are recommended.

A *membrane inlet* is often used when sampling fluids or gases, especially in medical or biological studies [26–29]. Its dynamic behavior has to be thorougly tested for correct interpretation of the results. Since substances have different permeabilities, a slow response of varying time constants are to be expected for different components and the depletion of the liquid film on the membrane must be avoided by agitation. Other problems are the long-term stability and mutual interference of the permeability of various components. The only remedy is testing the system over and over again.

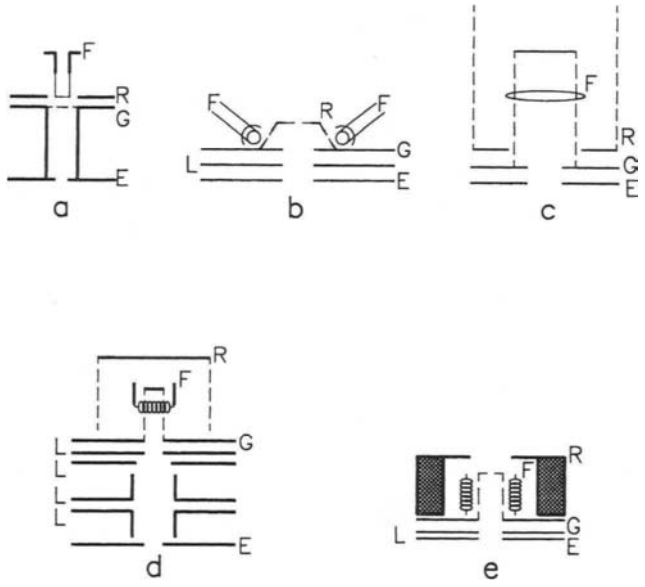
Modulated inlet such as pulsed gas sampling [98] can largely eliminate instability of the background levels resulting from interactions with the sample gas [19]. Modulated molecular beam coupled with phase sensitive detection improves the signal-to-noise ratio and also used in the study of chemical reactions.

3.2.3 ION GENERATION AND ION SOURCES

The sample arriving into the ion source faces a fundamental change. Not only its electrical state is changed, resulting in single- or multiple-charged ions, but often fragmentation processes take place, i.e., chemical bonds may be broken depending on their strength. The vast majority of analyzers use the positive ions for separation because they are formed more easily. The existence of many ions from a single component of a complex mixture may be confusing during their identification; on the other hand, this "cracking pattern" contains useful structural information on a molecule. One way to express the efficacy of this process of converting neutrals into ions is the ratio of the ion current and pressure (source sensitivity [A/torr or A/Pa], which is often close to a constant up to a high-pressure limit). There are several means to generate ions, as follows.

Electron impact is the most widely used method because of its simplicity, ruggeddness, and high sensitivity. In this source a filament is heated to a temperature high enough for thermionic emission (Figure 1). A grid, held at a potential higher than the filament, accelerates the electrons (50–150 eV range). The accelerated electrons then enter the "ionization chamber," or ion cage, a region where

Fig. 1.



Electron impact ion source configurations. F = filament, R = electron repeller, G = grid, L = ion-focusing lens, E = ionizer exit lens.

ion formation takes place, since the electrons have enough energy to form positive (and to some extent negative) ions from the neutral species present. (The ionization potential of most species falls within the range of 1–25 eV.) There is an exit aperture (focus plate) held at a lower potential to draw out the ions and prevent electrons from entering the analyzer. Beside these basic features, several more electrodes may be present depending on design, partly to help to form the electron beam (such as a repeller plate, to direct the electrons toward the ionization region, or Wehnelt electrode to focus the beam), partly to form and focus the ion beam to prepare it for entry into the analyzer.

Filament materials [99] are selected based on the need to obtain high emission current at possibly low operating temperature and long filament life.

Tungsten, as the highest melting point metal, has been used extensively, since it provides high emission current per unit surface area (~0.4 A/cm^2) at a temperature (~2500°K) where average filament life is several thousand hours (6 months in continuous use is typical). Filament life at constant filament temperature is proportional to the wire diameter used and for a given wire diameter life is a very strong function of filament temperature (life ~ (fil. temp.)$^{-35}$; 10% decrease in filament temperature (°K) increases life 40 times while 10% increase in filament temperature shortens life to about 4% of its original value). Unfortunately, tungsten is sensitive to the Langmuir cycle, whereas residual water molecules react

with the hot filament, forming volatile tungsten oxides and hydrogen. The oxides deposited on cooler surrounding surfaces will eventually be reduced by the hydrogen, resulting in a water molecule available for the next cycle and leaving a metallic deposit behind. This may create unwanted current path on isolators or alter potentials in the source. Another unwanted reaction of tungsten filament is its reaction with oxygen, partly because of the deposit of tungsten oxide — which may form insulating layers on metal electrodes. The charge-up of these layers can result in change in the potentials in the source leading to inefficient source operation, unsatisfactory peak shapes, etc. Also, the carbon impurities in the tungsten wire reacts with the oxygen on the surface, while the carbon is continuously diffusing to the surface. The resulting CO and CO_2 produce an annoyingly high background level. Molecular hydrogen dissociates at the filament, forming atomic hydrogen that is readily adsorbed by most surfaces. As a result, hydrogen signal is less — a local "pumping" of hydrogen occurs.

Rhenium is more resistant to water cycle, but its evaporation rate is about 150 times higher than that of tungsten for the same emission current density, therefore operation at a lower temperature is recommended. It does not form stable carbides, therefore the high CO background experienced with W filament can be lessened. Its resistance to other compounds is also known to be higher, although halogen compounds will form oxyhalides [100].

(Tantalum — another high-melting-point material — is useless as a filament material because of its tendency to absorb large amounts of hydrogen, which at high temperatures will alter its crystal structure causing extreme brittleness and short life.)

Thoriated tungsten is a 0.5–1% ThO_2 doped tungsten wire. At operating temperatures the thoria will diffuse to the previously carburized surface, providing free thorium. This monolayer will reduce the work function of tungsten by almost a factor of two, providing high emission current per unit surface area ($\sim 4 A/cm^2$) at lower temperatures, therefore extending life.

Thoria-coated cathodes are made by a different process, wherein the base metal is coated with a thick layer of ThO_2, which is activated by operation in vacuum. The lower work function again enables operation at lower temperatures [101], therefore the base metal does not have to be a refractory metal. The choice is therefore often iridium or rhodium to obtain a much higher level of oxidation resistance. (Accidental exposure to air will not result in sudden burnout.) Even after repeated exposure to corrosive gases constant emission characteristics were observed. This filament provides advantages when operating at high ionizer pressure because sample–filament reactions are suppressed, but also at extremely low pressures where the lower filament temperature keeps outgassing of surfaces at a lower level.

Lanthane hexaboride-coated base metals can (and have to) be operated at lower temperatures ($<1400°K$) also, to ensure long life and resistance to active gases. At higher temperatures, the boron tends to diffuse into the base metal, deteriorat-

ing its properties and the LaB_6 evaporation rate will become high. Rhenium is the most resistant to this boron diffusion, and so a preferred choice as base metal for lanthane hexaboride cathodes. As opposed to pure metals, coated cathodes can be poisoned by aggressive gases, resulting in elevated work functions.

Ion source configurations are as diverse as the special requirements they have to meet. Still, it is possible to categorize at least the typical arrangements used in commercial quadrupole instruments.

Axial (*radial*) sources are usually offered as the standard source. The name *axial* refers to either that the sample enters in the axis of the instrument (quadrupole rods) or that the electron beam is in-axis (and the sample enters through all the openings in the source). The name *radial* refers to the radial symmetry of the arrangement, when ring-type filaments are used. Electrons are generated by a filament either on axis, or by filament(s) in a plane perpendicular to the axis (ring-type filament, or several filaments forming a polygon). These are intended for general gas analysis or work at medium pressure (10^{-7}–10^{-5} torr, $\sim 10^{-5}$–10^{-3} Pa). A version called *closed, gas-tight,* or enclosed design is available to improve trace sensitivity. In this version the ionization chamber is separated from the filament and the rest of the analyzer by small apertures, enabling it to operate it at a higher pressure. This will increase sample–background ratio also limiting sample–filament interactions. The sample is led directly into this chamber through a small pipe.

Cross-beam sources will lead the sample into the source in a line perpendicular to the axis of the analyzer. This is especially suited for molecular beams or poorly collimated beams of condensable material or aggressive gases. Liquid-nitrogen-cooled beam stops or built-in beam shields then permit the continuous operation on vacuum deposition systems, molecular beam epitaxy, etc., without forming disturbing deposits in the analytical instrument. Again, an "open" and a "closed" version or versions integrated with ion extraction optics may be available.

Grid, open, RGA, or *UHV source* are all names covering an extremely open design for high pumping speed, low outgassing rate, and low electron-stimulated desorption. These sources are used at extremely low pressures, residual gas analysis, and UHV work. Instead of metal sheets, mesh are used for electron repulsion, cage, etc. If the instrument is intended as a residual gas analyzer, this may be the standard ion source.

Figure 1 shows some typical configurations found in partial pressure analyzers. (Sources a, d, e are variations of an axial-type source, c is a grid source, and b can be used as a cross-beam source.) The electrons emitted by the filament (F) are accelerated by the positive voltage applied to the grid (G) and directed toward the ionization chamber in the center by the negative voltage applied to the repeller (R) (both compared to the filament potential). Ions are extracted from the ionization region by the negative voltage applied to the lens (L) or the exit aperture (E) (voltage compared to the grid potential).

Special construction ion sources have been developed for particular applications, particular analyzers or trying to deal with the imperfections of standard sources. The most common ion source used in *magnetic sector* instruments is some variant of the traditional construction described by Nier [102]. An improved version of the source provides additional focusing to improve transmission through the analyzer and can be operated at higher accelerating voltages without arcing [103]. Further improvements replaced the rectangular lens elements with ones with circular symmetry, constructing the ion beam along the originally long side of the rectangular element. This resulted in better precision, better transmission and sharper focusing [104]. A hemispherical cavity in the focusing (ion exit) electrode acting as a Pierce lens can keep the potentials penetrating into the ion formation chamber nearly constant. Thereby the ions are formed along an equipotential surface, keeping their energy spread low [105]. A source providing wide electron energy range (19–70 eV), the incorporation of a spare filament and about a 10-fold differential pumping factor have been described for space application [106]. Several computer trajectory calculations were performed, aiming at producing an ion beam of higher current density, with mixed success [107,108,109], or aiming at understanding space charge and mass discrimination effects.

For ion sources operated with *quadrupole* instruments, trajectory tracing has also been used to improve ion yield and sensitivity. In an effort to eliminate ions from electron stimulated desorption [110], the electron-accelerating voltage is modulated and the ion current is measured in a phase-lock loop [111]. To increase the high operating limit of the ion source, auxiliary electrodes were installed at the inlet into the ionization chamber providing short, parallel electron path [112]. Linear source response has been reported with such a source up to 10^{-3} torr ($\sim 10^{-1}$ Pa) [113].

Photo ionization is a "soft" ionization relative to electron impact. Since chemical bonds are generally not broken, the amount of fragmentation or the strong signal from the matrix can be suppressed. On the other hand, the sensitivities of these sources are only a portion of the sensitivity of an electron impact source. Compared to low-energy electron ionization (<15 eV), it is much easier to produce photons in a narrow energy range than electrons. Also, just above the ionization potential, the efficacy of electron impact ionization is low compared to the impact with photons. Trace contaminants in air have been measured using the selective laser ionization — choosing the photon source so that the energy of the light is above the threshold ionization potential for contaminants but below that of the main components [114]. Laser ionization opens up the way of extending partial pressure measurement in the extreme-high-vacuum range (less than 10^{-9} torr, $\sim 10^{-7}$ Pa) by avoiding filament- and electron-induced desorption related problems [115].

Alpha ionization source is a way to avoid hot-cathode–sample reactions. The alpha radiation from a ^{210}Po source provides a "cold" way to produce ions. It

is extremely stable, being independent from external power supplies. However, to prevent sublimation of polonium and the resulting increase in noise level of the multiplier, it is necessary to place the source in a separate HV chamber with a 2–4 μ Ni window. The sensitivity of the source was only about 10^{-9} A/torr ($\sim 10^{-11}$ A/Pa) [116].

Field ionization is another method to avoid thermal effects. However, in its general form it produces relatively energetic ions (keV range) and if coupled to a quadrupole analyzer, deceleration is required, to slow down the ions into the 5–50 eV range. Complicated ion optics are necessary to prevent beam losses. A single pointlike source, where the sample is flowing axially through a volcano-shaped hole (~ 20 μ) of the field emitter, with a grid counterelectrode held at 1–3 keV potential is capable to achieve sensitivity around 10^{-8} A/torr ($\sim 10^{-10}$ A/Pa) [117]. Field emitter arrays, having a huge number of emitters on a small area can achieve emission currents around 1 mA. Using this semiconductor-type device in place of a filament, low-energy ions can be generated since the high voltage (~ 100 V) is between the base and the gate of the semiconductor device and does not affect the potential of the ionization chamber. Sensitivity approaching that of a standard hot-cathode electron impact ionizer can be achieved (10^{-2} A/torr; $\sim 10^{-4}$ A/Pa), but fluorine outgassing from the semiconductor is a problem yet to be solved [118].

Chemical ionization is a soft ionization method that found its way to partial pressure analysis by the name of *gas secondary ion mass spectrometry*. In this method a reagent gas (such as Kr, Xe, or CH_4) is ionized in a chamber by electron impact ionization. The resulting ions are extracted by electric fields and driven to the charge exchange chamber where, through ion molecule collisions with the sample, charge transfer takes place. Those constituents of the sample that have ionization energies below the noble gas ions are ionized without significant fragmentation (no electron bombardment). This method, beside being "soft," also opens the way to distinguish between different ion species of the same M/Q based on differences in their ionization energies (CO vs. N_2 or CH_3OH vs. O_2).

3.2.4 ION SEPARATION AND ANALYZERS

Analyzers are traditionally divided into categories such as "static" or "dynamic" depending on whether the force field in the right-hand side of the equation of ion motion is a function of time or not.

$$\frac{M}{Q}\frac{d^2\mathbf{r}}{dt^2} = \mathbf{E}(\mathbf{r}, t) + \mathbf{v}(\mathbf{r}, t) \times \mathbf{B}(\mathbf{r}, t) \tag{1}$$

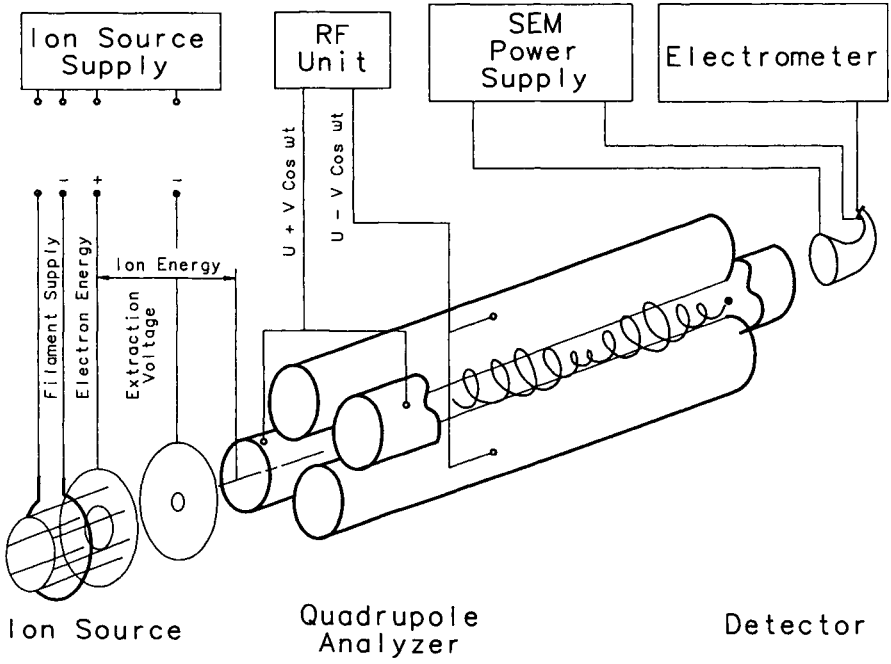
(M = mass of the ion, Q = electric charge of the ion, $\mathbf{E}(\mathbf{r}, t)$ = electric field intensity, $\mathbf{v}(\mathbf{r}, t)$ = velocity of the ion, $\mathbf{B}(\mathbf{r}, t)$ = magnetic induction, $\mathbf{r}$ = position vector of ion, t = time)

Since the force field used for separation is created by magnetic and/or electric fields, another conventional way of categorization is referring to "magnetic" or "electric" instruments. Whereas magnetic instruments usually apply a combination of magnetic and electric field, instruments that apply electric field only are mostly of the "dynamic" type such as the electric quadrupole. Dynamic instruments can be further divided into "path stability" or "time stability" instruments depending on whether the ions of the selected mass-to-charge ratio are separated from the others based on their path (electric quadrupole, magnetic sector), or based on the time during which they travel a certain distance (time-of-flight).

3.2.4.1 Electric Quadrupole Analyzers

Quadrupole analyzers composed of four parallel metal rods driven by an RF generator are the most widely used "path stability" dynamic instruments (Figure 2).

Fig. 2.



Schematic of an electric quadrupole instrument.

The equations of motions are determined by the constant U DC potential and $V\cos\omega t$ RF potentials applied with 180-degree phase difference to the two opposing pairs of electrically connected rods (of hyperbolic cross section ideally, but often round in practice).

$$\frac{M}{Q}\frac{d^2x}{dt^2} = -\frac{2x}{r_0{}^2}(U + V\cos\omega t) \tag{2}$$

$$\frac{M}{Q}\frac{d^2y}{dt^2} = \frac{2y}{r_0{}^2}(U + V\cos\omega t) \tag{3}$$

$$\frac{M}{Q}\frac{d^2z}{dt^2} = 0 \tag{4}$$

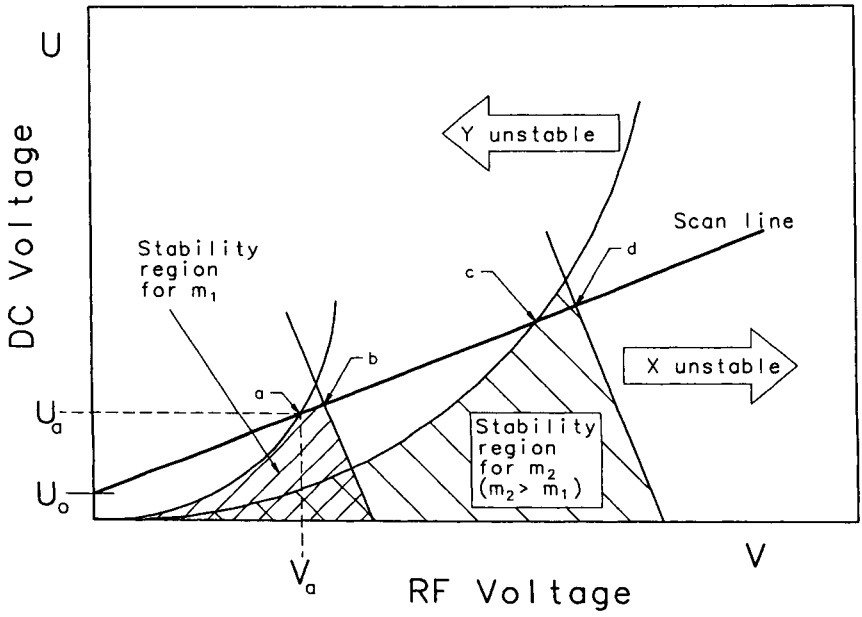
These equations describe oscillatory motions under the influence of $V\cos\omega t$ superimposed on the focusing effect of U in x direction and defocusing effect in y direction. (r_0 is the pitch circle radius of the rods, ω is the circular frequency of the RF voltage, and z axis is parallel with the axis of the cylinders.) The relationship between M/Q and the scanning parameter (voltage) is linear, resulting in a linear mass axis. (The space between peaks is independent from M/Q.)

The motion of ions is independent in the x and y direction. In x direction Equation (2) offers the simple explanation that the DC component in itself would generate a simple harmonic motion, while the RF component in itself would generate a forced oscillation. The superposition of the two forces will result in trajectories similar to a resonance system excited "off resonance"—oscillation with increasing, then decreasing amplitudes [119]. In y direction the DC component defocusing action is counterbalanced only by the focusing effect of the RF field. Since, as in x direction, the RF component in itself would generate simple harmonic motion where the displacement will be 180 degrees out of phase of the force, the addition of the DC component will cause oscillations around a path that will be increasingly farther apart from the z axis. However, the more the mean of the oscillation deviates from zero, the larger is the difference between the force at maximum displacement (being out of phase—pointing toward the z axis) and at minimum displacement (pointing away from the z axis). This difference, which results in a net force toward the z axis, creates the focusing effect of RF for the motion in y direction.

By substitution, the equations of motions in the x and y direction can be brought to a single form—to a Mathieu-type equation [120]. Solutions to the Mathieu equation give ion trajectories as a function of operating parameters, U, V, ω, and mass-to-charge ratio. Those combinations that result in stable trajectories define the "stability region" (can be expressed in terms of two dimensionless variables $a = 8U/mr_0{}^2\omega^2$ and $q = 4V/mr_0{}^2\omega^2$ where $m = M/Q$). In the a-q plane an area resembling a triangle contains all those a, q pairs that correspond to stable trajectories. (The three corner points are: a,q (0,0); a,q (0.237, 0.706) apex; and a,q (0, 0.91).) The curve connecting the first two points marks the border of the instability in the y direction, all points to the left being unstable in y direction. The line

connecting the last two points marks the border of the x instability region, all points to the right being unstable in x direction. At the apex of the triangle only one mass number (m) have stable trajectories (infinite resolution), while on a line segment within the triangle a range of mass numbers are "stable." The high end of this mass range is called "high pass critical mass," ions with higher mass number will have unstable trajectories in the y direction. Ions with lower mass number than the "low pass critical mass" will have unstable trajectories in the x direction [121]. Examples of ion trajectories and the a-q stability chart has been published in an earlier volume [122]. Instead of the a-q stability chart, we will discuss stability criteria and scanning in the U-V plane (Figure 3). As seen above, three corner points mark the triangle-like area of a-q pairs which assure stable trajectories. In a given instrument (given r_0) operating at a certain frequency (given ω) a and q are proportional to U/m and V/m respectively. Therefore in the U-V plane each mass has its own stability triangle. Scanning is realized by simultaneously increasing the DC and RF voltage from an initial DC value (U_0) along the "scan line." Ions with higher mass numbers (m) than the "high pass critical mass" will have unstable trajectories in y direction, while ions with lower mass numbers than the "low pass critical mass" have unstable trajectories in the x direction [121]. Examples of ion trajectories can be found in *Methods of Experimental Physics*, Volume 14 [122], and a sketch of the stability diagram is shown in Figure 3.

Fig. 3.



Sketch of ion path stability diagrams (quadrupole instruments). X and Y directions perpendicular to the axis of the analyzer.

Point a in Figure 3 on the scan line corresponds to the U_a and V_a pair where ions of mass to charge ratio equal to m_1 become stable not only in x direction but also in y direction. Ions of higher mass number are unstable in y direction, as shown by the "Y unstable" curve of m_2 ($m_2 > m_1$), which crosses the scan line at point c. Ions of much lower mass number than m_1 are unstable in x direction. The "x unstable line" for mass m_1 crosses the scan line at point b, while for m_2 it crosses at point d. For the U_a, V_a voltage pair the "x unstable" line would cross the scan line for a somewhat lower mass than m_1 at point a. (This mass number will be called $m_1 - \Delta m$.) It means that ions with mass number higher than $m_1 - \Delta m$ but lower than m_1 would exhibit stable trajectories both in x direction and y direction when the voltage pair U_a, V_a is selected. The width of this window, Δm, is clearly related to the slope of the scan line — at close to the tip of the stability regions the bandwidth of the filter will be narrow.

If U and V are increased simultaneously along the scan line, eventually point b is reached. It is the point where the x trajectories of ions of mass m_1 become unstable. The y unstable line for a somewhat higher mass number ($m_1 + \Delta m$) would cross the scan line at the same point b. Therefore, as U and V are increased simultaneously along the scan line, a mass window of width Δm is moved from lower mass numbers to higher mass numbers (scanning).

The overall result is that the RF field without the DC component will create a "high pass" filter that allows ions above a critical M/Q ratio to pass while the addition of the DC component will create a "low pass" filter. The magnitude of the DC voltage determines the "low pass" critical mass — if it is lower than the "high pass" critical mass, no ions can reach the detector; if it is higher than the "high pass" critical mass, a window is created.

Mass scan is accomplished by moving the "window" (Δm) by changing the amplitude of the RF (and the magnitude of the DC) voltage. When DC/RF ratio is low (<0.1), the width of the window is large, resulting in low resolution. When increasing the DC/RF ratio, the resolution increases on the expense of the transmission efficacy. If DC/RF ratio approaches about 0.168, the apex of the stability region is reached where the transmission drops to zero at infinitely high resolving power. Experimentally this critical DC/RF ratio has been found to depend weakly on the excitation frequency [123]. The excitation frequency, ω, also affects the ultimate obtainable resolving power ($M/\Delta M$ 2–3 times higher at 1.66 MHz than at 0.87 MHz).

While the resolving power will change as the mass number is varied in a typical instrument, two special cases can be identified. If the DC/RF ratio is kept constant along the mass sweep (U/V independent from M/Q), then this represents a "constant resolving power mode" of operation (constant $M/\Delta M$). In this case the peak width will be proprtional to the mass number — resulting in narrow peaks at low masses and wider peaks at high masses. Another special case is when the DC/RF ratio is varied as the mass number is increased in such a way that the width of the peak (ΔM) is kept constant. This is referred to as the "constant peak width

mode" of operation. The resolving power ($M/\Delta M$) will be consequently proportional to the mass number.

In practice the magnitude of an "offset" DC voltage (U_0, "resolution low") and the magnitude of the slope of the DC/RF function (approximately U/V at high masses, "resolution high") can be varied separately to adjust resolution at low and high M/Q. The magnitude of the RF and DC voltage is increased to shift the "window" along the mass scale from low M/Q toward high M/Q. A convenient and often attempted setup is to keep the peak width (ΔM) more or less constant along the entire mass range. A so-called integral spectrum is obtained easily by switching off the DC component and allowing all ions above M/Q to reach the detector as RF (V) is swept and the spectrum is recorded. As Equation (4) shows, the initial axial velocity remains unchanged in the quadrupole field.

Besides the so-called first-stability regions shown on Figure 3, there are several more predicted by theory as the Y-unstable and X-unstable curves cross each other again. These are of limited practical value due to decreased transmission and usable mass range, but some applications may benefit from the much improved abundance sensitivity (such as separating He from D_2) [124,125].

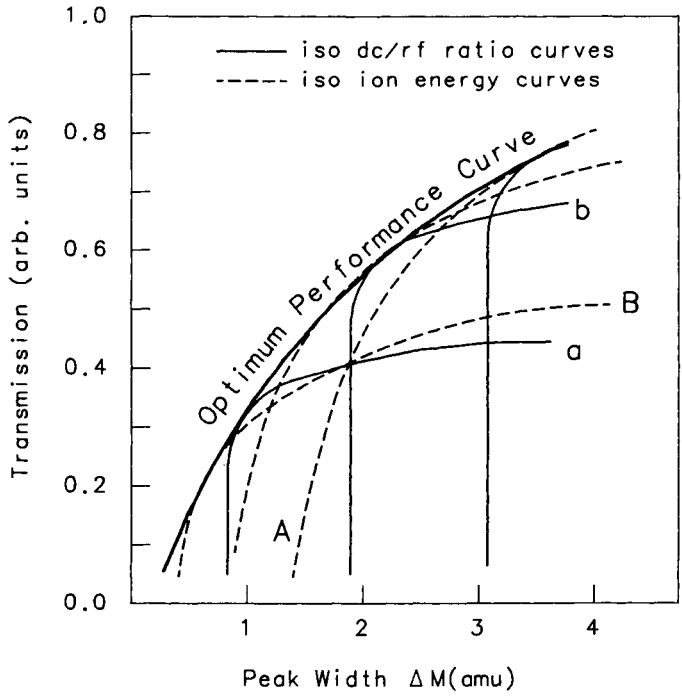
Even though commercial instruments have been marketed using the ideal hyperbolic cross section electrodes (Pt-coated ceramics) [126], the majority of practical analyzers employ metal rods of circular cross section. The benefit of hyperbolical, coated ceramic electrodes is not only the more perfect quadrupole field but rather the ability to produce partial pressure information at elevated temperatures up to 400°C (750°F). Operating the analyzer at elevated temperatures is helpful when quantitative measurement of adsorbing components such as moisture is attempted [127]. The optimum radius for cylindrical electrodes has been determined theoretically ($r = 1.148\ r_0$) [128]. The effect of replacing the hyperbolic cross section with a circular cross section has been studied extensively. The equation of motion includes terms reflecting coupling of the x and y fields. This results in a "blunted tip" and a slight shift to the right of the stability diagram, limiting the theoretically achieved maximum resolution [129,130]. The effect of the radius of the electrode relative to the optimum is important because of the increase in the required driving power to achieve the same resolving power and sensitivity (a factor of two, if r/r_0 is 1% higher than optimum at $M/\Delta M$=400). At constant driving power, the result is lower resolving power and sensitivity [129]. Although hyperbolic rods exceed round rods in performance (maximum achievable resolving power), increased driving frequency can make up for the difference (round rod quadrupole driven at 1.7 MHz performs similarly to a same-size hyperbolic driven at 0.7 MHz) [131]. The sharp decrease in sensitivity of quadrupoles with circular rods at a resolving power lower than a quadrupole with hyperbolic rods is primarily due to the decrease in the effective instrument radius. The perturbations in the electric fields caused by the circular shapes limit the maximum excursion of the oscillating ions from the z axis to a radius significantly smaller than r_0.

The transmission efficacy—which is the ratio of current collected to that en-

tering the quadrupole can be found by measuring the step corresponding to the selected mass number in integral mode (DC off) and standard mode (DC on) when using a conventional ion source. Plotting the transmission against the DC/RF ratio (which determines the resolving power) one will typically find transmission exceeding 80% up to about 0.9 times the DC/RF ratio corresponding to the apex of the stability diagram (about DC/RF=0.168) [123]. (Transmission was measured as a ratio of the ion current collected vs. the current entering the quadrupole at a given *M/Q*. The step in the "AC only" or total pressure mode was used as a measure of the current entering the filter. Overall transmission of the system, i.e., the ratio of ions collected to ions generated in the source (for a given *M/Q* when the filter is tuned to this *M/Q*), is typically 10^{-1}–10^{-3}.) Using a narrow ion beam, well-defined energy source movable under vacuum (such as an alkali ion source Na^+, K^+, etc.) enables one to measure the transmission efficacy as a function of initial displacement, direction and velocity [132]. From these measurements empirical transmission laws [133] and the acceptance ellipses of a particular quadrupole configuration can be derived.

Examining the transmission as a function of resolving power (Figure 4) at a given ion energy, three different regions can be distinguished. Since quadrupoles

Fig. 4.



Dependence of quadrupole analyzer transmission on resolution measured as the width of a peak (at 10% height).

in practice most often operate in constant peak width rather than in constant resolution mode, in the following discussion we use peak width as a measure of resolution. (Note that peak width increases as resolving power decreases.)

If at a given ion energy setting the resolution is low (peak width high), the transmission will be found to be practically a slow function of resolution (such as the quasi-horizontal part of line B on Figure 4). This is the "source-limited" mode of operation; the acceptance of the analyzer approaches or exceeds the emittance of the source. Stability and reproducibility are determined primarily by changes in the source. Transmission typically increases with initial velocity of ions (ion energy) due to the fringing fields effects (see section on "coupling of the ionizer to the analyzer," later). The peaks are "flat-topped"—trapezoidal.

Increasing the resolution (decrease in peak width) will take us to the region where the acceptance of the analyzer will be considerably smaller than the emittance of the source. This is the "analyzer limited" mode of operation, where the transmission efficiency will be strongly dependent on the entrance conditions. The latter is primarily due to the transient momentum impulse the ions receive at the entrance of the analyzer characterized by imperfect and highly complex quadrupole fields. It is well established that fringing fields introduce a dependence of the acceptance on axial ion velocity. In the fringing fields, ions are stable in the xz plane but unstable in the yz plane [134,135]. Combined acceptances (xz and yz planes) are at the maximum at axial velocities where the ions spend about two cycles in the fringing fields, and it falls off sharply for slower ions. (Combined acceptance decreases by a factor of 10 if the time spent in the fringing fields increases to four cycles [135].) This effect shows up as a decrease in sensitivity with decreasing ion energy at a given resolution for a given mass or as mass discrimination for ions injected at constant ion energy. (Heavier ions are slower at constant ion energy, resulting in lower transmission as they spend more time in the fringing fields.) Transmission in the "analyzer-limited" mode of operation will be inversely proportional to the square of resolving power as a general rule. Peak shape is triangular. (Between 0.7 and 1.5 Δm on line B.)

Further increase in resolution at the same ion energy will result in very sharp decrease in transmission, indicating that the "length" limited mode of operation is reached. Here the analyzer is increasingly unable to distinguish an ion with stable trajectory from an ion with unstable trajectory, which is related to the number of RF cycles the ion takes to pass through [135]. Peaks are sharp, triangular but their height decreases much more rapidly than their width as resolution is increased.

If these transmission–resolution (or peak width) curves are taken at all possible ion energies, the resulting envelope of all these curves defines the optimum performance curve. This curve connects the points corresponding to the highest possible transmission (or sensitivity) at each resolution (or peak width). Note that a different pair of DC/RF ratio and ion energy setting will realize maximum sensitivity at each resolution. Even though operating on this curve provides the maximum performance, in some cases it may be justifiable to operate below this curve

(such as in a source-limited mode to avoid that change in DC/RF ratio affects peak height.) The falloff of this optimum performance curve for a particular analyzer represents the limitations of the instrument itself. Factors that determine the maximum attainable resolution (besides finite length) are imperfections such as circular rods instead of hyperbolic, deviations from the perfect symmetrical geometry (displacement of rods, mechanical distortions due to baking), contamination of surfaces, surface charges [136]. Alignment of the ion source, imbalance of electric fields (DC/RF ratio) and spurious ripple superimposed on the RF can all be further limiting factors.

A further aspect of looking at transmission as a function of resolution is that only the points along the optimum performance curve are defined by a unique pair of ion energy and DC/RF ratio. All other transmission–resolution combinations below that can be realized by two pairs of ion energy DC/RF setting. One pair will be closer to the "source limited mode" (such as the intersection of lines b and B on Figure 4), while the other will be closer to the "length-limited" mode (such as the intersection of lines a and A). This is the result of the fact that both transmission and resolution are functions of the two variables: ion energy and DC/RF setting, thus defining two surfaces. The intersection of these two surfaces is the curve called the "optimum performance curve."

The optimum performance curve is not independent of mass-to-charge ratio. Since ions of higher masses spend longer time in the analyzer (at the same ion energy), the length-limited region will shift toward the left (if peak width is used as a measure of resolution — higher resolving power, i.e., lower Δm). On the other hand, high mass ions will spend more time in the fringing fields also, resulting in a downward shift of the source-limited region. The overall effect is that the optimum performance curve shifts to the left and downward as mass-to-charge ratio increased.

Coupling of the ionizer to the analyzer is an important topic especially in the case of *quadrupole mass filters*. If the substance to be analyzed fills the ionization chamber volume uniformly and is in thermal equilibrium with the walls, ions are created in the space determined by the ionizing beam. Consider first the much simpler one-dimensional case, where ions are generated along a line between two points determined by the width of the ionizing beam (Figure 5). Plotting the transverse initial velocity of the ions exiting the source against their location results in an ellipse that contains all the possible combinations. This is the *emittance* of the source (phase space) at the exit. On the other hand, the analyzer will accept ions depending on the position as well as transverse initial velocity of the ions. The ellipse that shows ions of what initial conditions will pass the analyzer is called the *acceptance* ellipse. (This ellipse represents the upper boundary for initial velocity/initial displacement combinations in the phase space that enables an ion to pass the analyzer tuned to its particular M/Q ratio.) The overlap of these two ellipses is the extraction. Depending on which ellipse is smaller, and so which one deter-

Fig. 5.

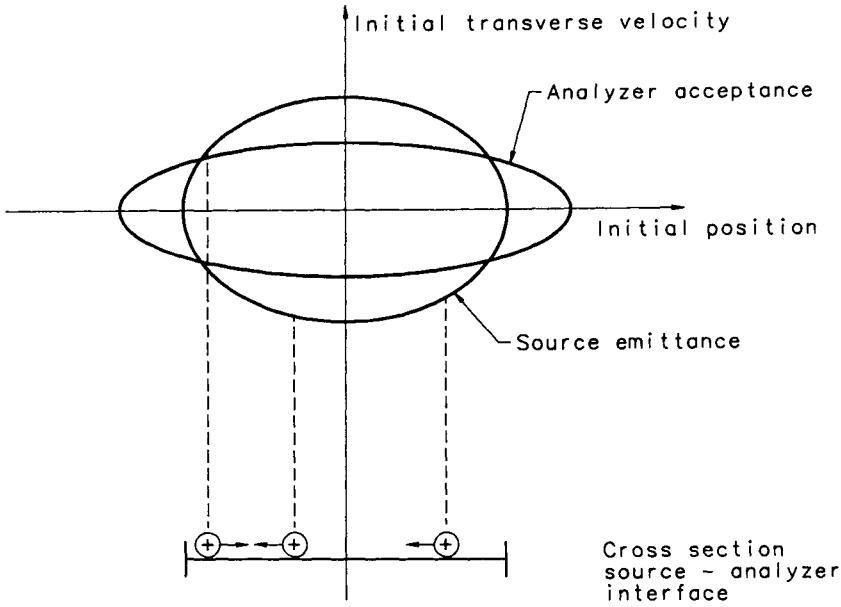


Illustration of the ion source emittance and quadrupole analyzer acceptance ellipses.

mines extraction, there are source-limited and acceptance-limited modes of operation [137]. In a practical case, a source can be characterized by a five-dimensional phase space—two directions perpendicular to the axis of symmetry along with two velocities in these directions and the velocity along the axis. (The acceptance and emittance can be characterized in more detail by a series of ellipses corresponding to various percentages of transmission and emission probabilities.) The electric fields between the source and the analyzer play a crucial role in determining transmission by altering the emittance of the source through the stray, so-called fringing fields [138,139]. Fringing fields are also responsible for mass discrimination; slow ions (at a constant ion energy, high-mass ions) have less chance of passing this region than faster ones.

As seen from the preceding discussion, the entrance and exit regions of the analyzer, where the DC field gradually approaches its final value (fringing fields), play a major role in analyzer overall transmittance. In the stability diagram, the fringing fields corresponds to the section of the scan line before it reaches the stability region (e.g., on Figure 3 the U_a, V_a voltage corresponds to the state when the filter is tuned to about mass number m_1. However, as an ion (mass number m_1) is traveling from the source to the area between the quadrupole rods, it is experiencing a gradual "buildup" of the electric quadrupole field. This corresponds to being exposed to the portion of the scan line between U_0 and point a that is beyond

the stability region for an ion of mass number m_1.) As a result, the ion trajectories are unstable leading to high losses. A way to avoid this is to "bend" the scan line so that it lies entirely within the stability region. This means that RF has to be increased first, DC following only later on. This is called the "delayed DC ramp" [140], which can be realized by using a set of four additional electrodes between the source and analyzer (and possibly also between the analyzer and detector) driven by RF only. Several commercial partial pressure analyzers apply this technique.

Variations of the quadrupole analyzer from its standard version as described earlier can be grouped as variations in the *driving voltage* and variations in the geometry of the filter.

A possible way to avoid losses and mass discrimination of high masses associated with fringing fields is not to use DC voltage at all. This so-called *AC-only* mode of a quadrupole is essentially a high-pass filter, providing an integral spectrum. To obtain a regular spectrum, a "retarding field" or energy filter is applied at the exit of the analyzer, which works based on the observation that ions that are close to the border of the stability region (corresponding to the step in integral mode) tend to be defocused, filling the aperture at the exit. Since the fringing fields (from the RF) increase in field strength as the displacement from the z axis increases, these ions gain more energy from the fields than others [141]. A further clear advantage is the elimination of the need to keep a constant DC/RF ratio. Indeed, instruments with low mechanical tolerances behave in a superior way in this mode of operation, but there is little difference if the instruments have been manufactured accurately. Disadvantages are the difficulties at low M/Q (below 20) and high-ion energy (above 4eV) related to the increased number of RF cycles an ion must spend in the filter to obtain the same resolution as compared with the case where the focusing/defocusing action of DC voltage is also present [142,143,144].

The removal of stringent control of DC/RF ratio can alternatively be achieved by driving the rods with a rectangularly time-varying voltage waveform [145], with the promise of better stability, but this variation has not yet been adapted in commercial instruments.

Variations in the *filter geometry* have been investigated, such as concave shaped [146], flat electrodes [147], coated ceramics [148], and electrodes made of (4×3) wires to approximate quadrupole fields or to eliminate higher-order terms [149]. All these are still using functionally four electrodes.

It has been recognized very early (1963) [150] that it is not necessary to use four electrodes to employ the mass analyzer properties of quadrupole fields and Equations (2) to (4). The combination of one circular (or hyperbolic) and a V-shaped electrode (the circular rod being in the 90-degree opening of the V) can simulate a quarter of the quadrupole field. This arrangement is the "*monopole*" mass spectrometer that has also been manufactured commercially. Other proposals — which have not been commercialized — are the "*dipole*" mass spectrometer composed of two circular rods on one side of a plate electrode creating "half" a

quadrupole field [151]. A theoretical investigation of "*multipole*" devices (hexapole, octapole) as a generalization of the geometry of the quadrupole has already been carried out [152]; and some of the geometries and means to realize it (such as shaped glass) are patented.

Shorter but higher number of rods open up the way to increase the maximum allowable operating pressure. Since the mean free path is inversely proportional to the pressure, ion molecule collisions will affect linearity above 10^{-5} torr ($\sim 10^{-3}$ Pa) (see Section 3.2.1.2). Shorter rods reduce ion-molecule collisions, but in order to maintain resolution the frequency of the RF voltage has to be increased. (Resolution is related to the number of cycles the ions spend in the quadrupole field.) Using the same range of U and V (limited by arcouts) the pitch circle radius of the rods (r_0) has to be decreased proportionally to the increase in ω. (See Section 3.2.4.1 The variables defining the stability region — a and q — are function of the product of r_0 and ω). The smaller r_0 in turn will reduce the acceptance area of the analyzer which can be offset by using an array of quadrupoles acting parallel. A practical device using a 3×3 *quadrupole array* (16 rods forming 9 quadrupole analyzers) has been described [153]. While the smaller size of the analyzer head and the higher operating pressure limit are clear advantages, resolution overall and power required to drive the analyzer are less favorable than in the case of regular quadrupoles.

If the cross section of the quadrupole (four circles or hyperbolae) is rotated around an axis passing through the center of a pair of opposing rods (e.g., the x axis) a ring and two end-caps are formed. This structure is a "three dimensional quadrupole" or the *quadrupole ion trap*. The electrons may enter via a center hole in one of the end-caps while the other may provide an exit for the ions formed (and stored) in the space confined by the ring electrode and the two end caps. Ions formed by electron impact are analyzed by increasing the amplitude of the RF voltage on the ring electrode. This results in destabilization of the ion trajectories for higher and higher masses in axial direction. The ejected ions are detected by an external detector. Trapped ions can be measured also by power absorption of a supplementary excitation signal or by measuring the differential charge on the end-caps induced by ion motion [154]. Quadrupole ion traps found application as detectors of field apparatus used for environmental characterization or monitoring [155].

3.2.4.2 Magnetic Sector Analyzers

As Equation (1) shows, ions with a velocity v in a perpendicular magnetic field B will accelerate in a direction perpendicular to both v and B, thus describing a circular path of radius r:

$$\frac{M}{Q}\frac{v^2}{r} = vB \tag{5}$$

Ions obtain their initial velocity, v, by acceleration through an electric field, E, thus:

$$\frac{Mv^2}{2} = QE \tag{6}$$

Combining the two equations results the "magnetic mass spectrometer equation":

$$\frac{M}{Q} = \frac{B^2 r^2}{2E} \tag{7}$$

Since the source and the collector are usually at fixed positions, the two options to "scan"—i.e., change the radius of the ion beam according to M/Q is to either vary the magnetic field strength, B (using an electromagnet) or to use a fixed magnet and vary the strength of the accelerating electric field E (voltage scanning). In either case the relationship between M/Q and the scanning variable is nonlinear, resulting in a nonlinear mass axis. (The space between the peaks is smaller (the spectrum is "compressed") at the high end as compared with the lower M/Q end.)

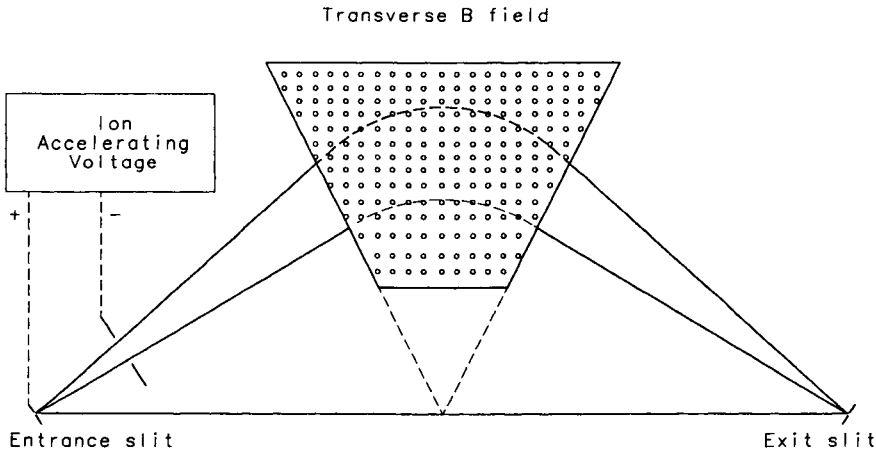
To prevent ions with widely different angles of velocity (and energy) to enter the spectrometer, an entrance slit is used that reduces the yield of the ion source, defining the beam size. For proper resolution, an exit slit is used that allows only those ions to enter the detector that are focused. (Magnetic sector field instruments of this type provide single focusing—ions of equal M/Q and energy but divergent direction of velocity at the entrance are focused on the exit slit.) For the greatest sensitivity, the width of the entrance slit (S_1) should be as great as possible. The width of the exit slit (S_2) has to be even larger by ΔS to allow for a wider beam caused by the aberration due to variation in initial directions ($S_2 = S_1 + S$). In this case the resolving power will be [156]

$$\frac{M}{\Delta M} = \frac{r}{2S_2} \tag{8}$$

Directional focusing can be achieved with various beam deflection angles. In the Dempster-type instrument the ion beam is deflected in a 180-degree sector field, thus the ion beam is parallel through the entrance and the exit slits, but moving in opposite directions. The ion source and the detector are within the magnetic field in this configuration.

There may be a need to remove the source and the detector from the magnetic field for easier access and to eliminate interferences with multipliers. This is possible using 30-, 60-, or 90-degree sector fields. To ensure focusing, the entrance and exit slits must be on the line perpendicular to the symmetry axis of the magnet, the line being placed at the point where the field boundary lines intersect (Figure 6). This arrangement will provide for single (directional) focusing effect.

Fig. 6.



Schematic of a magnetic single-focusing analyzer (60 degrees).

Alignment of the magnet is typically more critical in this arrangement and therefore recalibration may be necessary after removing it (e.g., for bakeout).

These single-focusing spectrometers separate ions according to M/Q from a monoenergetic ion beam. Since there is an unavoidable spread in energy, only an additional electrostatic deflection can provide directional and energy focusing. Such a double-focusing instrument for partial pressure analysis has been described by Nier [157].

In all cases mechanical parts (entrance and exits slits) must be changed to affect sensitivity and resolution. Compared to quadrupoles (change in electronic setting), this may be a limitation if frequent changes are needed, but may be an advantage in stability and repeatability depending on the application. Resolving power $M/\Delta M$ is constant throughout the mass range, and typical scanning speeds are of the order of a few amu/s.

The coupling of a magnetic sector analyzer (90-degree) to a Nier-type electron source has been studied by Werner [110]. At low emission currents ($< 50\mu$A) and low pressures (less than 10^{-6} torr or about 10^{-4} Pa) the source magnetic field, the initial thermal energy and negative as well as positive space charge effects could be neglected. At higher emission currents and low pressures, negative space charge; at high pressures, positive space charge may influence the potentials, leading to variation in ion energy. Mass discrimination during voltage scanning can be explained by a decrease in the size of the ion formation region and by an increase in the angular divergence due to the reduction in accelerating voltage. At low accelerating voltage (corresponding to higher mass numbers), the ion trajectories have shown a larger divergence and concentration on the edges of the

exit slit. As opposed to this, at higher accelerating voltage (2 kV) the ions are concentrated along the axis (higher transmission at lower mass numbers).

3.2.4.3 Time-of-Flight Analyzers

Ions are generated by a short pulse of transverse electron beam in the source, which is followed by another voltage pulse (extraction) to provide ions with a certain kinetic energy. The residence time of the ions in this accelerating region, τ will depend on their M/Q ratio ($\tau(M/Q)$). Based on Equation (1),

$$\frac{M}{Q}\frac{d^2\mathbf{r}}{dt^2}\tau(M/Q) = E\tau(M/Q) \tag{9}$$

From energy considerations—or integrating Equation (9) along the path of the ions in the constant field E—the velocity of the ions will be proportional to the square root of their mass:

$$v = \sqrt{\frac{2EQ}{M}} \tag{10}$$

(The kinetic energy of an ion equals the energy gained from the electric field.) If these ions are let drift along a field-free tube (drift tube of a length l), the time required to reach the detector will depend on their M/Q ratio, enabling mass analysis:

$$t = l * \sqrt{\frac{M}{2EQ}} \tag{11}$$

Whereas sensitivity will be proportional to the duration of the ionizing pulses, resolving power will be proportional to the ratio of ion flight time to the duration of the ion extracting pulse (typically μs). Therefore the length of the drift tube is critical in determining resolution (drift tube length of 1 meter or more is typical). Although this is the simplest analyzer structure of all PPAs, rather complex driving circuitry is required to provide the pulses and detection of ions. Reaction studies and study of fast-changing processes use the extremely short response time of this instrument.

Some instruments employ techniques to improve resolving power, which is limited by—beside the length of the tube—the energy spread of ions formed at different regions in the source (space focusing, time lag focusing techniques) [158].

Time-of-flight instruments can be coupled relatively easily to electron impact ionizers. Ions generated by electron bombardment are held in the potential well

formed by the electron beam traveling at right angle to the axis of the instrument. A pulse applied to the extraction grid perpendicular to the axis will deflect the beam, and the ions are extracted through this grid. Several other grids can be applied to reduce field penetration of the drift tube (0–3 kV) in the source [159]. For XHV measurements ($<10^{-12}$ torr, $\sim 10^{-10}$ Pa) segregation of gas-phase ions from electron stimulated desorption related ions is a major challenge. A time-of-flight quadrupole has been constructed for such low pressure measurements [160].

3.2.5 DETECTION OF IONS

The ions leaving the analyzer have been separated from each other qualitatively (according to M/Q), but their abundances need to be measured for quantitative analysis or just simple identification purposes. Therefore the ion signal is converted into an electric signal by either simply converting the ion current into an electron current in a metal electrode (Faraday collector) or in a way that also provides for amplification (multipliers).

3.2.5.1 Faraday Collectors

Ions striking a metal surface are neutralized by electron transfer — the related current when the metal is held at constant potential can be measured through a large ohmic resistor. A current to voltage converter then can produce a signal proportional to the ion current, however, double-charged ions will produce a signal twice as high as the corresponding ion stream (ion/s) compared to single-charged ions.

In its simplest form the Faraday collector is only a metal plate following the exit of the analyzer. A major drawback of this simple device is that the incident ions may eject secondary electrons when striking into the metal surface, causing a higher measured current than the corresponding ion current. Gold plating is a way to minimize secondary electron emission.

Secondary electron effects can also be minimized by using a closed structure except for the aperture for ions to enter (Faraday cup). Secondary electrons then hit a surface inside the cup and be recaptured, and so their effect is eliminated. Another solution is to use an additional grid before the Faraday cup to act as an electron suppressor (held at a potential well below that of the Faraday cup).

A limitation of the Faraday collector is its speed of measurement. This is related to the capacitance of input cables and the necessity to use high-input resistors. At 10^{-16} A, a time constant is typically about 1 s. Measurement at such low current

levels are limited by noise, the main sources of which are the cables and connectors and ultimately the thermal noise produced by the resistor [161].

3.2.5.2 Discrete Dynode Multipliers

To overcome limitations such as speed of measurement, sensitivity, and dynamic range, electron multipliers are used. These devices make use of the secondary electron emission — a phenomenon being suppressed in Faraday collectors. In a discrete dynode multiplier the ions are accelerated by voltages of the order of kilovolts, and directed to the front end (first dynode) of the multiplier. Secondary electrons in a number greater than the number of impinging ions are emitted from this surface and accelerated to the second dynode. On impact, even more secondary electrons are emitted. Repeating this process through several stages (14–16 are typical) amplification by a factor of 10^6 can be achieved. The electron current from the last dynode is measured by a Faraday collector. The overall gain of a given multiplier depends on the voltage applied in a stronger-than-linear fashion. The principal secondary emission material used is most often beryllium-copper (beryllium-oxide layer as the secondary emitting surface), but silver-magnesium and aluminum have also been used [162].

The requirements for the multipliers are (beside high secondary emitting ratio), the stability of the gain over time and possibly following exposure to samples. As a first approximation, it is assumed that the gain varies as $M^{-0.5}$ as a result of the ion-electron conversion efficacy being proportional to the velocity of the ion (which at constant energy varies as $M^{-0.5}$). Unfortunately the gain also depends on the type of the ion (M/Q) [161], and on the state of the surfaces that emit secondary electrons. Exposure to samples or contamination may cause severe deterioration of gain, making frequent checks necessary. Gain measurement can be accomplished by comparing the signal obtained using the multiplier with the signal collected when the first dynode is used as a Faraday collector. Rejuvenation of multipliers after a decrease in their gain have been reported to be successful (such as baking in 1 atm of O_2 at 300°C) [162,163].

If further increase in sensitivity is needed, the multiplier can be operated as an "ion counter," extending the detection limit to a single ion.

3.2.5.3 Channel Electron Multipliers

Instead of a complex structure of many dynodes, a compact single piece of special (high-lead) glass tube can be used that exhibits high secondary emission characteristics. High voltage (kV) applied to contacts on both ends provides for replenishing the charge lost and also accelerates the secondary electrons. In prin-

ciple there is no difference between this device and the discrete dynode multiplier, but the small size and lower cost make it appealing. One ion will produce typically a pulse of about 10 ns width (at half height).

To eliminate ion feedback (ions created by secondary electrons from residual gases, hitting the walls and creating a further burst of secondary electrons) the tube is curved limiting the distance an ion can travel and the amount of energy they can acquire. Some channel electron multipliers (CEMs) [164] have been designed especially for pulse-counting mode. These devices allow the use of discriminators to distinguish between signal and noise simply based on pulse amplitude. The lifetime depends on exposure time to high input levels (ion bombardment of front surfaces), and also the residual gas environment, which may cause degradation of the secondary emissive surface [165], but in general CEMs are less susceptible to gain change resulting from exposure to air than discrete dynode multipliers.

A method for measuring CEM gain based on the ratio of DC current and noise current eliminates the need to measure gain by switching cables and particularly suitable for CEMs [166]. Besides multiplier voltage, the gain depends on the ion mass, incident energy, and chemical nature of the ion (typically decreases with increasing mass number) requiring the use of correction factors [167].

3.2.5.4 Microchannel Plate Multipliers

An extension and integration of the CEM with the fiber optic technology resulted in the appearance of microchannel plate multipliers (MCPs). Microchannel plates are thin lead glass plates with millions of 10 μm diameter holes acting as parallel short CEMs. The walls of the holes are treated to create a high-resistance coating with higher than 1 coefficient of secondary emission. High voltage is applied on the top and the bottom surfaces of the plate, coated with a conductive layer. The channels are set at an angle to ensure that electrons traveling normally to the plate will strike the wall. To eliminate ion feedback two of these plates are placed on each other ("Chevron configuration") [168]. MCPs are especially well suited to quadrupole analyzers, providing a large area for detection, but magnetic sector spectrometers have also applied it successfully [169]. As with other multipliers, strong ion signals can cause saturation due to space charge or field distortion in the output. Contamination of the surfaces can cause field emission, which in a severe case can lead to a permanently saturated state. MCPs are extremely fragile, and if they must be removed from the vacuum environment they should be stored in desiccated containers to avoid breakage due to moisture takeup. Short exposure to atmosphere (few hours) will generally not lead to performance losses; however, exposure to hydrocarbon vapors or operation in vacuum but at elevated temperatures (above 50°C) may cause permanent damage.

REFERENCES

1. R. Schubert, *Anal. Chem.*, **44** (1972) 2084–2087.
2. J. Hardy, *Advances in Mass Spec.*, **6** (1974) 1061–107.
3. K. Anderson and K. Mauersberger, *Rev. Sci. Instrum.*, **52** (1981) 1025–1028.
4. K. Mauersberger and R. Finstad, *Rev. Sci. Instrum.*, **50** (1979) 1612–1617.
5. Leybold Inficon Inc., *American Environmental Laboratory*, (1995) Nov.–Dec., 34–35.
6. F. R. Lauritsen, *Int. J. Mass Spec. Ion Proc.*, **95** (1990) 259–268.
7. S. Bohatka, I. Futo, and I. Gal, *Vacuum*, **44** (1993) 669–671.
8. S. Bohatka, *Vacuum*, **46** (1995) 767–768.
9. S. Bohatka and I. Berecz, *Vakuum Technik.*, **35** (1986) 79–88.
10. G. Langer, *Vacuum*, **34** (1984) 757–784.
11. S. Hiroki et al., *J. Vac. Sci. Technol.*, **A12** (1994) 2711–2715.
12. C. M. Hohenberg, *Rev. Sci. Instrum.*, **51** (1980) 1075–1082.
13. I. P. Wright, I. A. Boyd, and S. R. Franchi, *J. Phys. E. Sci. Instrum.*, **21** (1988) 865–875.
14. F. E. Jones and F. E. Cartan, *Report # ICP-1085*, 1976 (Available from the Nat'l Technical Information Service, 5285 Port Royal Road, Springfield VA 22161.)
15. J. Dattner and J. Fischler, *Brit. J. Appl. Phys.*, **14** (1963) 728–729.
16. R. E. Ellefson and D. Cain, *J. Vac. Sci. Technol.*, **A5** (1987) 134–139.
17. D. J. Santeler, *J. Vac. Sci. Technol.*, **A5** (1987) 129–133.
18. E. Dancs, R. Dobrozemsky, and G. Schwarzinger, *Vacuum*, **43** (1992) 537–539.
19. L. Lieszkovsky, and J. Borossay, *J. Vac. Sci. Technol.*, **A5** (1987) 2819–2822.
20. J. Koprio, and G. Rettinghaus, *Advances in Mass Spec.*, **10** (1985) 891–892.
21. R. T. Parkinson and L. Toft, *Analyst*, **90** (1965) 220–227.
22. J. Roboz, *Anal. Chem.*, **39** (1967) 175–178.
23. E. E. Hughes and W. Dorko, *Anal.Chem.*, **40** (1968) 750–755.
24. I. Morisako, I. Kato, and Y. Ino, *Int. J. Mass Spec. Ion Phys.*, **48** (1983) 19–22.
25. K. Siefering, H. Berger, and W. Whitlock, *J. Vac. Sci. Technol.*, **A11** (1993) 1593–1597.
26. G. M. Solomon, *J. Vac. Sci. Technol.*, **A2** (1984) 1157–1161.
27. H. Selhofer and R.Wagner, *Proc. 7th Intern. Vac. Congr., Vienna 1977*, (1977) 211–214.
28. I. Berecz and S. Bohatka, *Proc. 8th Intern. Vac. Congr., Cannes 1980*, (1980) 251–254.
29. A. Holme, J. D. Buckingham, and K. T. Fowler, *Dynamic Mass Spectrometry*, **5** (1978) 129–136.
30. B. Goodwin, A. E. Holme, and J. H. Leck, *Proc. 7th Intern. Vac. Congr. Vienna 1977*, (1977) 177–180.
31. A. E. Holme, M.Parris, and S. Purdy, *J. Vac. Sci. Technol.*, **A6** (1988) 1156–1157.
32. I. Niishi, S. Sugai, and K. Kimura, *Proc. 6th Intern. Vac. Congr. Kyoto 1974*, (1974) 163–166.
33. A. Smith and M. J. Pettifor, *Vacuum*, **32** (1982) 175–181.
34. J. M. Welter, *Proc. 6th Intern. Vac. Congr. Kyoto 1974*, (1974) 175–178.
35. O. G. Koch and R. Reinhard, *Fres'Z. Anal. Chem.*, **231** (1967) 420–429.
36. W. G. Perkins, *J. Vac. Sci. Technol.*, **10** (1973) 543–556.
37. G. Thieme, *Vacuum*, **13** (1963) 137–143.
38. R. S. Barton and R. P. Govier, *J. Vac. Sci. Technol.*, **2** (1965) 113–122.
39. M. Mosharrafa and H. J. Oskam, *Physica.*, **32** (1966) 1759–1765.
40. J. J. Sullivan and R. G. Buser, *J. Vac. Sci. Technol.*, **6** (1969) 103–108.
41. H. F. Dylla, W. R. Blanchard, and R. B. Krawchuk, *J. Vac. Sci. Technol.*, **A2** (1984) 1188–1192.
42. W. Poschenrieder, *J. Vac. Sci. Technol.*, **A5** (1987) 2265–2272.
43. C. G. Pearce, *Advances in Mass Spec.*, **8** (1980) 1771–1777.
44. H. F. Dylla, *Proc. 9th Intern. Vac. Congr. Madrid 1983*, (1983) 578–588.
45. H. F. Dylla, *J. Vac. Sci. Technol.*, **A1** (1983) 1297–1301.

46. W. R. Blanchard, H. J. McCarthy, and H. F. Dylla, *J. Vac. Sci. Technol.*, **A4** (1986) 1715–1719.
47. R. Dobrozemsky and G. Schwarzinger, *J. Vac. Sci. Technol.*, **A10** (1992) 2661–2664.
48. J. H. Scrivens, *Vacuum*, **32** (1982) 169–174.
49. P. L. Swart and H. Aharoni, *J. Vac. Sci. Technol.*, **A3** (1985) 1935–1945.
50. J. Visser, *J. Vac. Sci. Technol.*, **10** (1973) 464–471.
51. N. Muller, G. Rettinghaus, and G. Strasser, *J. Vac. Sci. Technol.*, **A8** (1990) 2822–2825.
52. M. E. Buckley, *Vacuum*, **44** (1993) 665–668.
53. B. Raby, *J. Vac. Sci. Technol.*, **15** (1978) 205–208.
54. A. J. G. Schellingerhout, T. M. Janocko, and M. A. Klapwijk, *Rev. Sci. Instrum.*, **60** (1989) 1177–1183.
55. J. A. Koprio, G. Peter, and H. Fischer, *Vacuum*, **38** (1988) 783–786.
56. M. Benyezzar, D. King, and P. L. Jones, *Vacuum*, **43** (1992) 21–24.
57. E. Hasler and G. Rettinghaus, *Vacuum*, **38** (1988) 1–5.
58. C. Lu, M. J. Lightner, and C. A. Gogol, *J. Vac. Sci. Technol.*, **14** (1977) 103–107.
59. C. A. Gogol and S. H. Reagan, *J. Vac. Sci. Technol.*, **A1** (1983) 252–256.
60. R. Kubiak, *J. Vac. Sci. Technol.*, **A1** (1991) 2423–2427.
61. P. Looney and S. Tison (eds.), *Water: Its Measurement and Control in Vacuum* (Gaithersburg 1994).
62. R. K. Lowry, NBS IR 84-2852, Proc. Moist. Meas. Workshop (National Bureau of Standards, Gaithersburg, 1984), p. 59.
63. L. J. Rigby, *Int. J. Mass Spec. Ion Proc.*, **60** (1984) 251–264.
64. K. L. Perkins, NBS SP 400-69, Proc. Moist. Meas. Workshop (National Bureau of Standards, Gaithersburg, 1981), p. 58.
65. K. Lin and J. D. Burden, *J. Vac. Sci. Technol.*, **15** (1978) 373–376.
66. T. J. Rossiter, F. Gollob et al. NBS IR 84-2852, Proc. Moist. Meas. Workshop (National Bureau of Standards, Gaithersburg, 1984), 8–23.
67. W. M. Hickam, NBS SP 400-69, Proc. Moist. Meas. Workshop (National Bureau of Standards, Gaithersburg, 1984), p. 153.
68. J. E. Fenstermacher, *J. Vac. Sci. Technol.*, **8** (1971) 380–384.
69. S. Baptist and F. Levy, *Vacuum*, **43** (1992) 213–214.
70. J. E. Fenstermacher, *J. Vac. Sci. Technol.*, **A5** (1987) 142–143.
71. A. Pebler and W. M. Hickam, *Anal. Chem.*, **45** (1973) 315–319.
72. S. K. Gupta, *Analyst*, **109** (1984) 793–794.
73. R. Schubert and J. A. Augis, *J. Vac. Sci. Technol.*, **A1** (1983) 248–251.
74. D. G. Mahoney, *J. Vac. Sci. Technol.*, **A12** (1994) 1740–1743.
75. C. M. Ward and L. E. Bergquist, *J. Nucl. Mat.*, **85** (1979) 117–120.
76. J. A. Koprio, H. Gaug, and H. Eppler, *Int. J. Mass Spec. Ion Phys.*, **48** (1983) 23–26.
77. L. J. Rigby, *Dynamic Mass Spectrometry*, **3** (1972) 237–242.
78. O. G.Koch, *Fres'Z. Anal. Chem.*, **284** (1977) 13–18.
79. Y. Rosenthal, G. Adam, and G. Kohn, *J. Vac. Sci. Technol.*, **A5** (1987) 389–390.
80. M. C. Paul, *Vacuum*, **15** (1965) 239–247.
81. S. Watanabe, M. Aono, and S. Kato, *J. Vac. Sci. Technol.*, **A14** (1996) 3261–3266.
82. W. Nesseldreher, *Vacuum*, **26** (1976) 281–286.
83. J. R. J. Bennett and R. J. Elsey, *Vacuum*, **44** (1993) 647–651.
84. W. L. Fite and P. Irwing, *J.Vac. Sci. Technol.*, **11** (1974) 351–356.
85. R. J. Reid, *Vacuum*, **28** (1978) 499–505.
86. C. Wursching, *Vacuum*, **43** (1992) 137–141.
87. B. Podor, I. Bertoti, and G. Mink, *Vacuum*, **33** (1983) 71–72.
88. R. D. Pilkington and D. Price, *Dynamic Mass Spectrometry*, **6** (1981) 284–290.
89. H. Eppler and H. Selhofer, *Thermochimica Acta*, **20** (1977) 45–52.
90. R. D. Collins, P. Fiveash, and L. Holland, *Vacuum*, **19** (1969) 113–116.

91. K. Hoefler, A. Hoeglund, and L. G. Rosengren, *Int. J. Mass Spec. Ion Phys.*, **46** (1983) 127–130.
92. B. R. F. Kendall, V. S. Rohrer, and V. J. Bojan, *J. Vac. Sci. Technol.*, **A4** (1986) 598–601.
93. Y. Wang and Z. Yang, *J. Phys. E: Sci. Instrum.*, **20** (1987) 1483–1486.
94. J. Basford, M. Boeckmann, R. Ellefson, A. Filippelli, D. Holkeboer, L. Lieszkovszky, and C. Stupak, *J. Vac. Sci. Technol.*, **A11** (1993) A22–A40.
95. L. V. Lieszkovszky, A. R. Filippelli, and C. R. Tilford, *J. Vac. Sci. Technol.*, **A8** (1990) 3838–3854.
96. J. H. Batey, *Vacuum*, **44** (1993) 639–642.
97. G. Bunyard, *Vacuum*, **44** (1993) 633–638.
98. I. E. Sodal and D. A. Hanna, *J. Vac. Sci. Technol.*, **A15** (1997) 176–181.
99. P. E. Gear, *Vacuum*, **26** (1976) 3–10.
100. B. Podor, *Vacuum*, **33** (1983) 67–68.
101. J. H. Batey, *Vacuum*, **43** (1991) 15–19.
102. A. Nier, *Rev. Sci. Instrum.*, **18** (1947) 398–411.
103. L. Dietz, *Rev. Sci. Instrum.*, **30** (1959) 235–241.
104. J. F. J. Todd and H. S. McKown, *Int. J. Mass Spec. Ion Phys.*, **42** (1982) 183–190.
105. S. Dworetsky, R. Novick, and W.Smith, *Rev. Sci. Instrum.*, **39** (1968) 1721–1723.
106. W. M. Brubaker, D. Burnham, and G. Perkins, *J. Vac. Sci. Technol.*, **8** (1971) 273–2741.
107. P. S. Naidu and K. O. Westphal, *Brit. J. Appl. Phys.*, **17** (1966) 645–651.
108. W. Fock, *Int. J. Mass Spec. Ion Phys.*, **3** (1969) 285–291.
109. H. W. Werner, *J. Phys. E: Sci. Instrum.*, **7** (1974) 115–121.
110. J. L. Segovia, *Vacuum*, **47** (1996) 333–340.
111. F. Watanabe and H. Ishimaru, *J. Vac. Sci. Technol.*, **A3** (1985) 2192–2195.
112. H. Miyake and M. Michijima, *Proc. 6th Intern. Vac. Congr. Kyoto, 1974* (1974) 171–174.
113. B. Hu and J. Qiu, *J. Vac. Sci. Technol.*, **A5** (1987) 2657–2660.
114. C. He-neng, A. J. Genuit, W. Boerboom, and J. Los, *Int. J. Mass Spec. Ion Phys.*, **46** (1983) 277–280.
115. K. Kokubun, S. Ichimura, and H. Shimizu, *Vacuum*, **44** (1993) 657–659.
116. F. Daublin, *Vakuum-Technik*, **20** (1971) 72–78.
117. W. Aberth and C. Spindt, *Int. J. Mass Spec. Ion Phys.*, **25** (1977) 183–198.
118. N. Ogiwara and M. Shiho, *Vacuum*, **44** (1993) 661–663.
119. W. Brubaker, *Proc. 5th Instrum. & Measur. Conf. Stockholm 1961* (1961) 305–315.
120. P. H. Dawson, *Int. J. Mass Spec. Ion Phys.*, **14** (1974) 317–337.
121. J. E. Campana, *Int. J. Mass Spec. Ion Phys.*, **33** (1980) 101–117.
122. W. M. Brubaker, in *Methods of Experimental Physics*, Vol. 14, edited by G. L. Weissler and R. W. Carlson (Academic Press, New York, 1979), pp. 81–100.
123. W. M. Brubaker and J.Tuul, *Rev. Sci. Instrum.*, **35** (1964) 1007–1010.
124. S. Hiroki, T. Abe, and Y. Murakami, *Rev. Sci. Instrum.*, **62** (1991) 2121–2124.
125. S. Hiroki, T. Abe, and Y. Murakami, *Rev. Sci. Instrum.*, **63** (1992) 3874–3876.
126. G. Reich, *Proc. 7th Intern. Vac. Congr. Vienna 1977* (1977) 197.
127. L. Frees, D. Holkeboer, and J. Farden, *30th Annu. Conf. Mass Spec. & Allied Topics Honolulu 1982* (1982) 656–656.
128. I. E. Dayton, F. C. Shoemaker, and R. F. Mozley, *Rev. Sci. Instrum.*, **25** (1954) 485–495.
129. D. R. Denison, *J. Vac. Sci. Technol.*, **8** (1971) 266–269.
130. J. Y. Tu, *J. Phys. E: Sci. Instrum.*, **22** (1989) 368–372.
131. W. M. Brubaker and W. S. Chamberlin, *NASA Re. NASW 1298* (1970) 98–103.
132. J. F. Hennequin and R. L. Inglebert, *Int. J. Mass Spec. Ion Phys.* **26** (1978) 131–135.
133. R. L. Inglebert and J. F. Hennequin, *Advances in Mass Spec.*, **8** (1980) 1764–1770.
134. P. H. Dawson, *Mass Spectrometry Reviews*, **5** (1986) 1–37.
135. P. H. Dawson, *J. Vac. Sci. Technol.*, **A4** (1986) 1709–1714.

136. A. E. Holme, W. J. Thatcher, and J. H. Leck, *J. Phys. E: Sci. Instrum.*, **5** (1972) 429–433.
137. P. H. Dawson, *Adv. Electron. & Electron Phys.*, **53** (1980) 153–208.
138. P. H. Dawson, *Int. J. Mass Spec. Ion Phys.*, **6** (1971) 33–44.
139. W. M. Brubaker, *Advances in Mass Spectrometry*, **4** (1967) 293–299.
140. C. Trajber, M. Simon, and S. Bohatka, *Vacuum*, **44** (1993) 653–656.
141. U. Brinkmann, *Int. J. Mass Spec. Ion Phys.*, **9** (1972) 161–166.
142. A. E. Holme, *Int. J. Mass Spec.Ion Phys.*, **22** (1976) 1–5.
143. A. E. Holme, S. Sayyid, and J. H. Leck, *Int. J. Mass Spec. Ion Phys.*, **26** (1978) 191–204.
144. P. H. Dawson, M. Meunier, and W. Tam, *Advances in Mass Spec.*, **8** (1979) 1629–1637.
145. J. A. Richards, R. M. Huey, and J. Hiller, *Int. J. Mass Spec. Ion Phys.*, **15** (1974) 417–428.
146. T. Hayashi and N. Sakudo, *Rev. Sci. Instrum.*, **39** (1968) 958–968.
147. C. G. Pearce and D. Halsall, *Int. J. Mass Spec. Ion Phys.*, **27** (1978) 31–41.
148. S. Hiroki et. al., *Vacuum*, **44** (1993) 71–74.
149. H. Matsuda, and T. Matsuo, *Int. J. Mass Spec. Ion Phys.*, **24** (1977) 107–118.
150. U. von Zahn, *Rev. Sci. Instrum.*, **34** (1963) 1–4.
151. P. H. Dawson, *Int. J. Mass Spec. Ion Phys.*, **12** (1973) 53–85.
152. I. Szabo, *Int. J. Mass Spec. Ion Proc.*, **73** (1986) 197–235.
153. R. J. Ferran and S. Boumsellek, *J. Vac. Sci. Technol.*, **A14** (1996) 1258–1265.
154. D. Goeringer, R. Crutcher, and S. McLuckey, *Anal. Chem.*, **67** (1995) 4164–4169.
155. W. C. McDonald, B. Abraham, and A. Robbat Jr., *Environ. Sci. Technol.*, **28** (1994) 336A–343A.
156. G. F. Weston, *Vacuum*, **30** (1980) 49–67.
157. A. Nier, *Rev. Sci.Instrum.*, **31** (1960) 1127.
158. W. C. Wiley and I. H. McLauren, *Rev. Sci. Instrum.*, **26** (1955) 1150–1157.
159. D. B. Wilson, *Vacuum*, **19** (1969) 323–325.
160. N. Ogiwara, Y. Miyo, and Y. Ueda, *Vacuum*, **47** (1996) 575–578.
161. R. D. Collins, *Vacuum*, **19** (1969) 105–111.
162. R. Young, *Rev. Sci. Instrum.*, **37** (1966) 1414–1417.
163. G. Salser, *Rev. Sci. Instrum.*, **37** (1966) 674–675.
164. E. Kurz, *American Laboratory*, **3** (1979) 67–82.
165. S. Kaszczyszyn and A. Grzeszczak, *Vacuum*, **45** (1994) 323–324.
166. W. J. Fies, *Int. J. Mass Spec. Ion Proc.*, **82** (1988) 111–129.
167. N. R. Reagan, L. C. Frees, and J. W. Gray, *J. Vac. Sci. Technol.*, **A5** (1987) 2389–2392.
168. G. F. Vanderschmidt, *J. Vac. Sci. Technol.*, **A9** (1991) 2001–2006.
169. D. Murphy and K. Mauersberger, *Rev. Sci. Instrum.*, **56** (1985) 220–226.

Table I

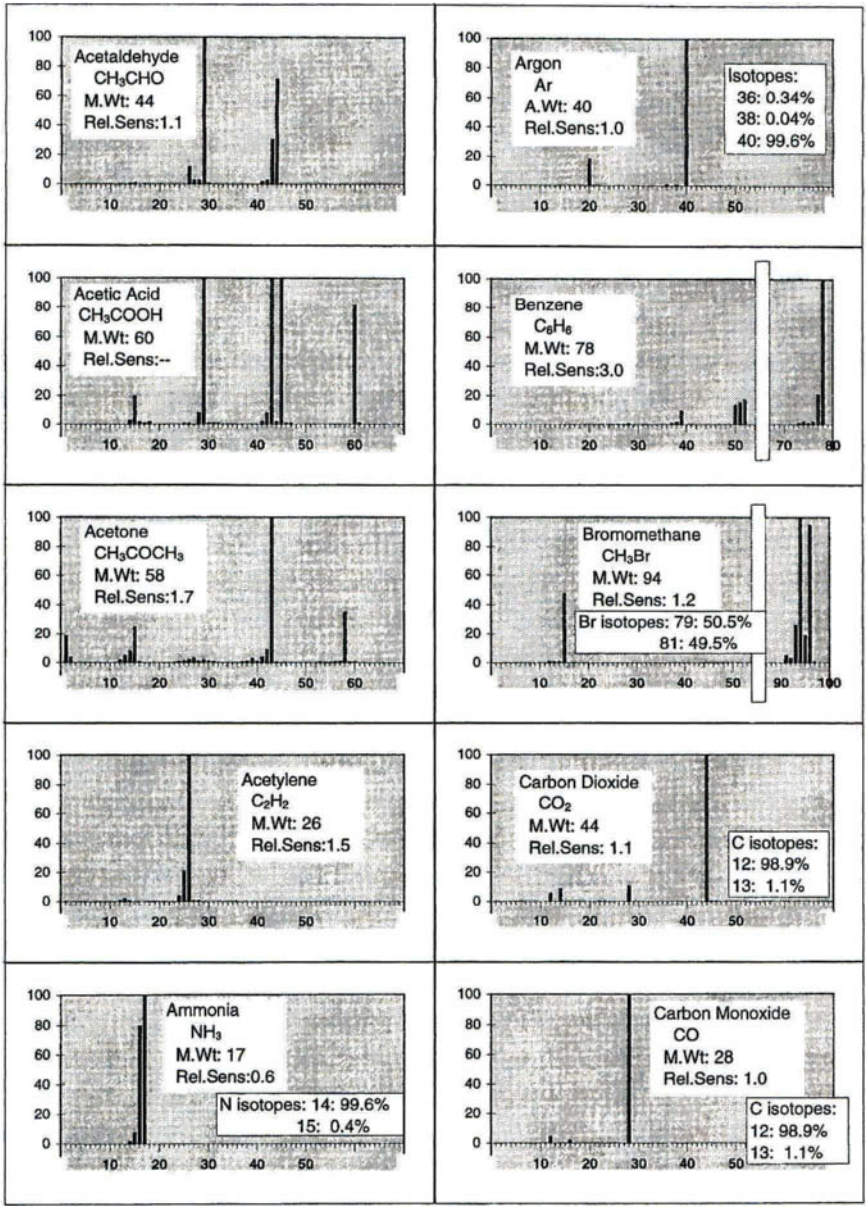

Acetaldehyde
CH3CHO
M.Wt: 44
Rel.Sens:1.1
Argon
Ar
A.Wt: 40
Rel.Sens:1.0
Isotopes:
36: 0.34%
38: 0.04%
40: 99.6%
Acetic Acid
CH3COOH
M.Wt: 60
Rel.Sens:--
Benzene
C6H6
M.Wt: 78
Rel.Sens:3.0
Acetone
CH3COCH3
M.Wt: 58
Rel.Sens:1.7
Bromomethane
CH3Br
M.Wt: 94
Rel.Sens: 1.2
Br isotopes: 79: 50.5%
81: 49.5%
Acetylene
C2H2
M.Wt: 26
Rel.Sens:1.5
Carbon Dioxide
CO2
M.Wt: 44
Rel.Sens: 1.1
C isotopes:
12: 98.9%
13: 1.1%
Ammonia
NH3
M.Wt: 17
Rel.Sens:0.6
N isotopes: 14: 99.6%
15: 0.4%
Carbon Monoxide
CO
M.Wt: 28
Rel.Sens: 1.0
C isotopes:
12: 98.9%
13: 1.1%

Table 2

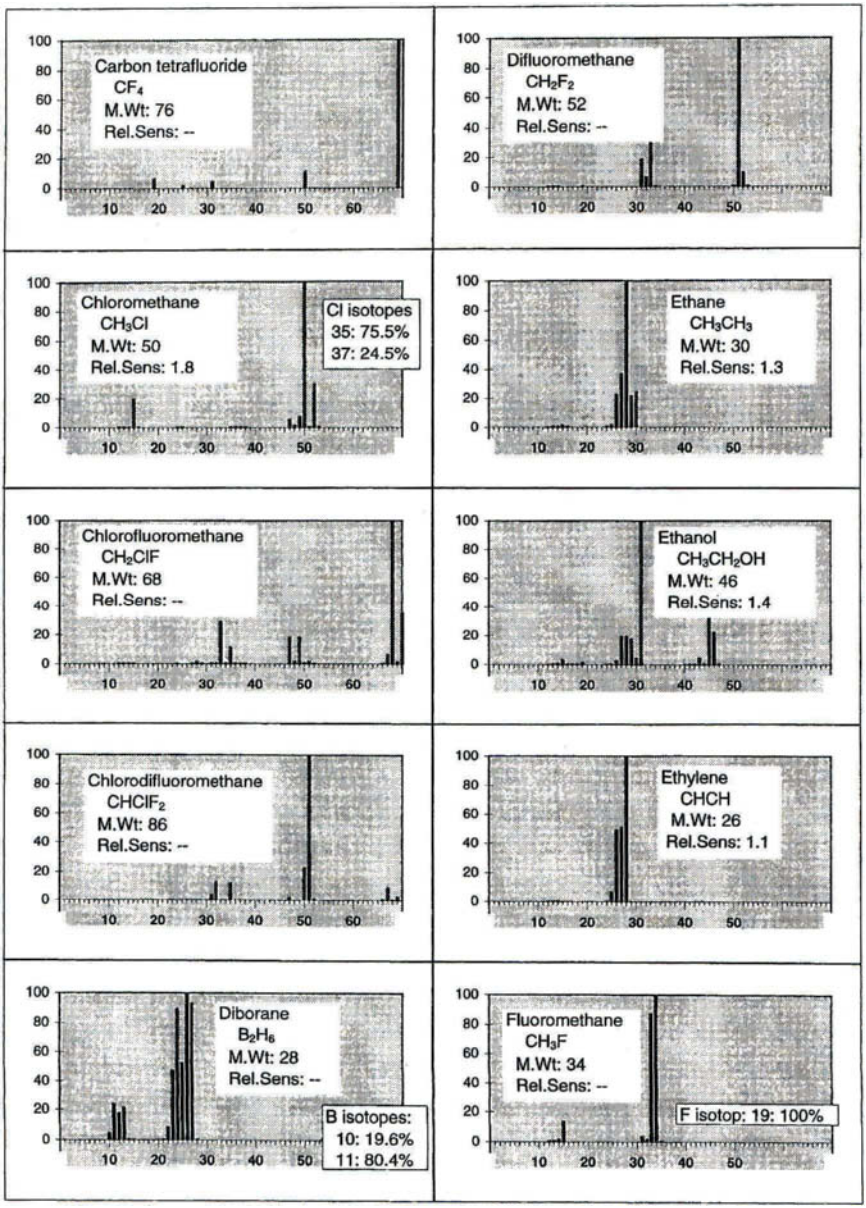

Carbon tetrafluoride
CF_4
M.Wt: 76
Rel.Sens: --
Difluoromethane
CH_2F_2
M.Wt: 52
Rel.Sens: --
Chloromethane
CH_3Cl
M.Wt: 50
Rel.Sens: 1.8
Cl isotopes
35: 75.5%
37: 24.5%
Ethane
CH_3CH_3
M.Wt: 30
Rel.Sens: 1.3
Chlorofluoromethane
CH_2ClF
M.Wt: 68
Rel.Sens: --
Ethanol
CH_3CH_2OH
M.Wt: 46
Rel.Sens: 1.4
Chlorodifluoromethane
$CHClF_2$
M.Wt: 86
Rel.Sens: --
Ethylene
CHCH
M.Wt: 26
Rel.Sens: 1.1
Diborane
B_2H_6
M.Wt: 28
Rel.Sens: --
B isotopes:
10: 19.6%
11: 80.4%
Fluoromethane
CH_3F
M.Wt: 34
Rel.Sens: --
F isotop: 19: 100%

Table 3

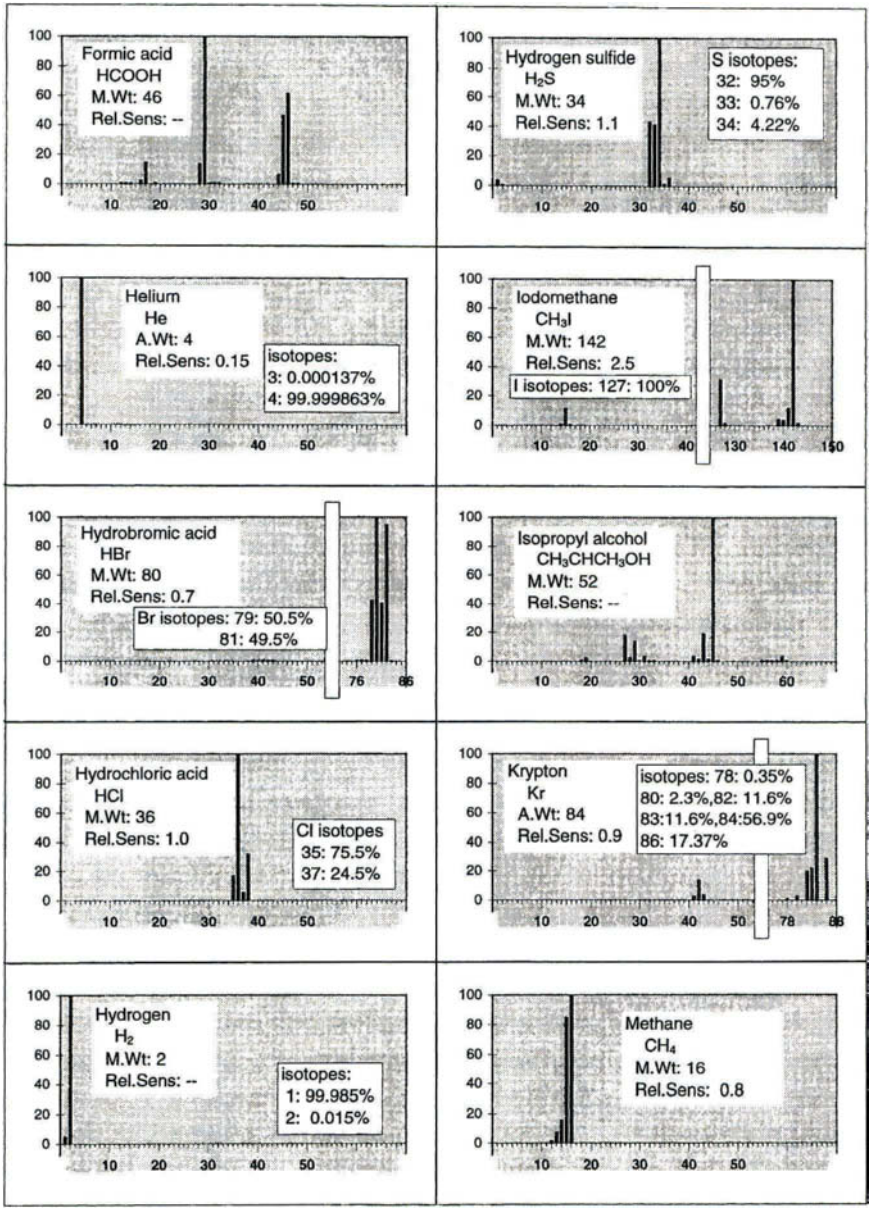

Formic acid
HCOOH
M.Wt: 46
Rel.Sens: --
Hydrogen sulfide
H2S
M.Wt: 34
Rel.Sens: 1.1
S isotopes:
32: 95%
33: 0.76%
34: 4.22%
Helium
He
A.Wt: 4
Rel.Sens: 0.15
isotopes:
3: 0.000137%
4: 99.999863%
Iodomethane
CH3I
M.Wt: 142
Rel.Sens: 2.5
I isotopes: 127: 100%
Hydrobromic acid
HBr
M.Wt: 80
Rel.Sens: 0.7
Br isotopes: 79: 50.5%
81: 49.5%
Isopropyl alcohol
CH3CHCH3OH
M.Wt: 52
Rel.Sens: --
Hydrochloric acid
HCl
M.Wt: 36
Rel.Sens: 1.0
Cl isotopes
35: 75.5%
37: 24.5%
Krypton
Kr
A.Wt: 84
Rel.Sens: 0.9
isotopes: 78: 0.35%
80: 2.3%,82: 11.6%
83:11.6%,84:56.9%
86: 17.37%
Hydrogen
H2
M.Wt: 2
Rel.Sens: --
isotopes:
1: 99.985%
2: 0.015%
Methane
CH4
M.Wt: 16
Rel.Sens: 0.8

Table 4

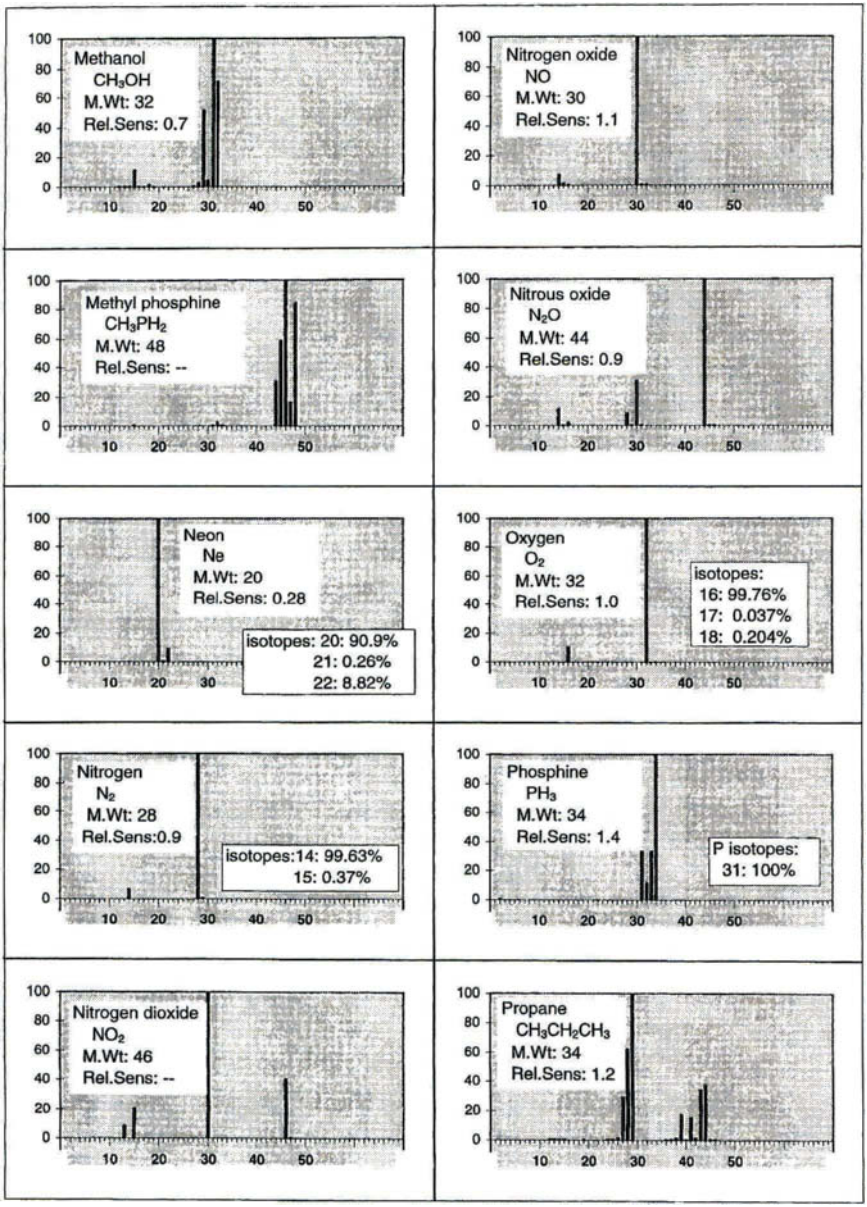
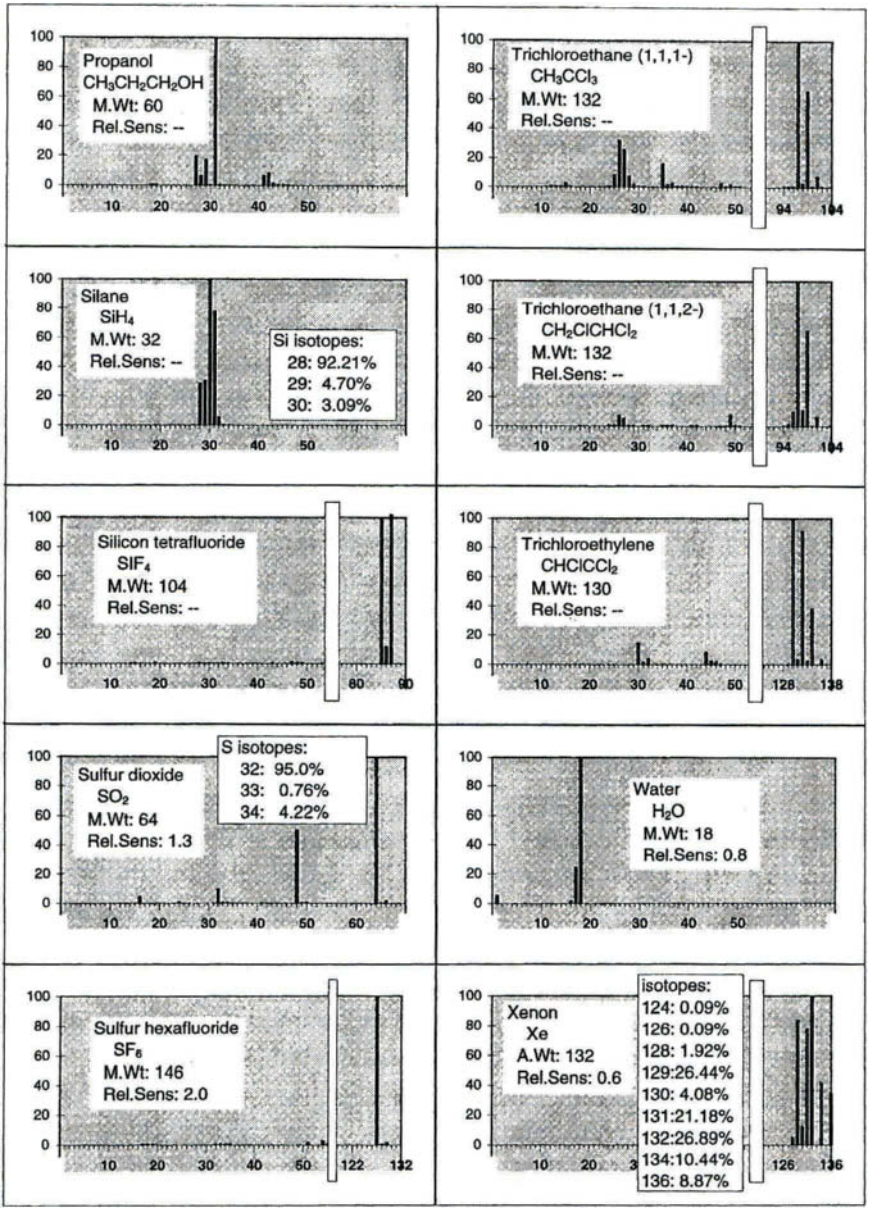

Methanol
CH3OH
M.Wt: 32
Rel.Sens: 0.7
Nitrogen oxide
NO
M.Wt: 30
Rel.Sens: 1.1
Methyl phosphine
CH3PH2
M.Wt: 48
Rel.Sens: --
Nitrous oxide
N2O
M.Wt: 44
Rel.Sens: 0.9
Neon
Ne
M.Wt: 20
Rel.Sens: 0.28
isotopes: 20: 90.9%
21: 0.26%
22: 8.82%
Oxygen
O2
M.Wt: 32
Rel.Sens: 1.0
isotopes:
16: 99.76%
17: 0.037%
18: 0.204%
Nitrogen
N2
M.Wt: 28
Rel.Sens:0.9
isotopes:14: 99.63%
15: 0.37%
Phosphine
PH3
M.Wt: 34
Rel.Sens: 1.4
P isotopes:
31: 100%
Nitrogen dioxide
NO2
M.Wt: 46
Rel.Sens: --
Propane
CH3CH2CH3
M.Wt: 34
Rel.Sens: 1.2

Table 5


Propanol
$CH_3CH_2CH_2OH$
M.Wt: 60
Rel.Sens: --
Trichloroethane (1,1,1-)
CH_3CCl_3
M.Wt: 132
Rel.Sens: --
Silane
SiH_4
M.Wt: 32
Rel.Sens: --
Si isotopes:
28: 92.21%
29: 4.70%
30: 3.09%
Trichloroethane (1,1,2-)
$CH_2ClCHCl_2$
M.Wt: 132
Rel.Sens: --
Silicon tetrafluoride
SiF_4
M.Wt: 104
Rel.Sens: --
Trichloroethylene
$CHClCCl_2$
M.Wt: 130
Rel.Sens: --
S isotopes:
32: 95.0%
33: 0.76%
34: 4.22%
Sulfur dioxide
SO_2
M.Wt: 64
Rel.Sens: 1.3
Water
H_2O
M.Wt: 18
Rel.Sens: 0.8
Sulfur hexafluoride
SF_6
M.Wt: 146
Rel.Sens: 2.0
Xenon
Xe
A.Wt: 132
Rel.Sens: 0.6
isotopes:
124: 0.09%
126: 0.09%
128: 1.92%
129:26.44%
130: 4.08%
131:21.18%
132:26.89%
134:10.44%
136: 8.87%

CHAPTER 3.3

Practical Aspects of Vacuum System Mass Spectrometers

R. A. Outlaw
Teledyne Brown Engineering: Hastings Instruments

The most severe deficiency of total pressure instruments is that they provide virtually no chemical composition information concerning the gas environment in a vacuum system. The lack of identification of the partial pressures that make up the total pressure is such an enormous disadvantage to system diagnostics that any good vacuum system should have a mass spectrometer (also known as residual gas analyzer, partial pressure analyzer, or simply mass spec). It provides, first of all, a built-in leak detector, which is necessary to confirm the vacuum integrity of a system. Second, it measures the state of cleanliness of a system by indicating the partial pressure and type of residual gas. Third, it is also a research tool that provides numerous diagnostic capabilities such as determining gas purity, gas desorption from elements of interest, and gas changes that occur during various processes.

3.3.1 HISTORICAL INSIGHT

Historical insight into the still developing utility of vacuum system mass spectrometers is of some interest [1–3]. Before 1950, a good vacuum or a low total pressure was the essential environment required for conducting various processes such as minimizing the slag in a vacuum furnace or minimizing the oxidation rate

ISBN 0-12-325065-7
$25.00

of the filament in an early incandescent light bulb or a vacuum tube. Other processes, such as charge particle spectroscopy, require a low total pressure to ensure a sufficiently long mean free path to allow the collection of charged species for analysis. Measurement of the composition and partial pressure of the gases that comprise the total pressure was initially of some interest, but not achievable. Eventually, because of progressing scientific objectives, engineers and scientists began to actively pursue methods to determine the identity of the gas species. One needed to know, for example, the actual composition of a total pressure to understand what the limiting factors were to achieving even lower pressures. Mass spectrometers developed in the beginning of this century had emphasis primarily on the physics of the mechanism. J. J. Thompson's positive ray analyzer (1910) and Dempster's 180-degree magnetic sector (1918) were the first of this kind [4]. However, in the latter part of the 1930s, A. O. C. Nier began applied studies in atmospheric research with the primary interest in understanding the variation in composition and magnitude of the gas species as a function of altitude [5]. Nier made many advances in the technology of mass spectrometry, especially with a unique ion source that provided high resolution. He used both magnetic sector and quadrupole-type instruments. Concurrent with the work of Nier, interest in understanding the chemical composition of complex organic molecules, the study of isotopes, and the pursuit of equipment to provide lower total pressures drove the technology. In the latter part of the 1950s, the well-funded space program gave another boost to the continuing development. The National Aeronautics and Space Administration's (NASA) research in examining the composition profile of the atmosphere with altitude, the residual gases in space, the determination of the atmospheres of planetary bodies, and space simulation propelled new and innovative laboratory and flight experiments [6]. The *Apollo* missions used mass spectrometers to measure the hydrogen atmosphere on the moon and the *Viking* mission to measure the CO_2 atmosphere of Mars. These are only a few examples of the wealth of information mass spectrometers have provided to the understanding of space. The funding pumped into the study of vacuum technology waned significantly after the moon landings, but mass spectrometers still continued to be slowly developed and the uses for these instruments rapidly increased. In the middle 1970s, the semiconductor industry became the primary driver for vacuum equipment and the development of mass spectrometers got another boost. This continues today—the sensitivity, resolution, and performance of these instruments as well as the software have all dramatically improved.

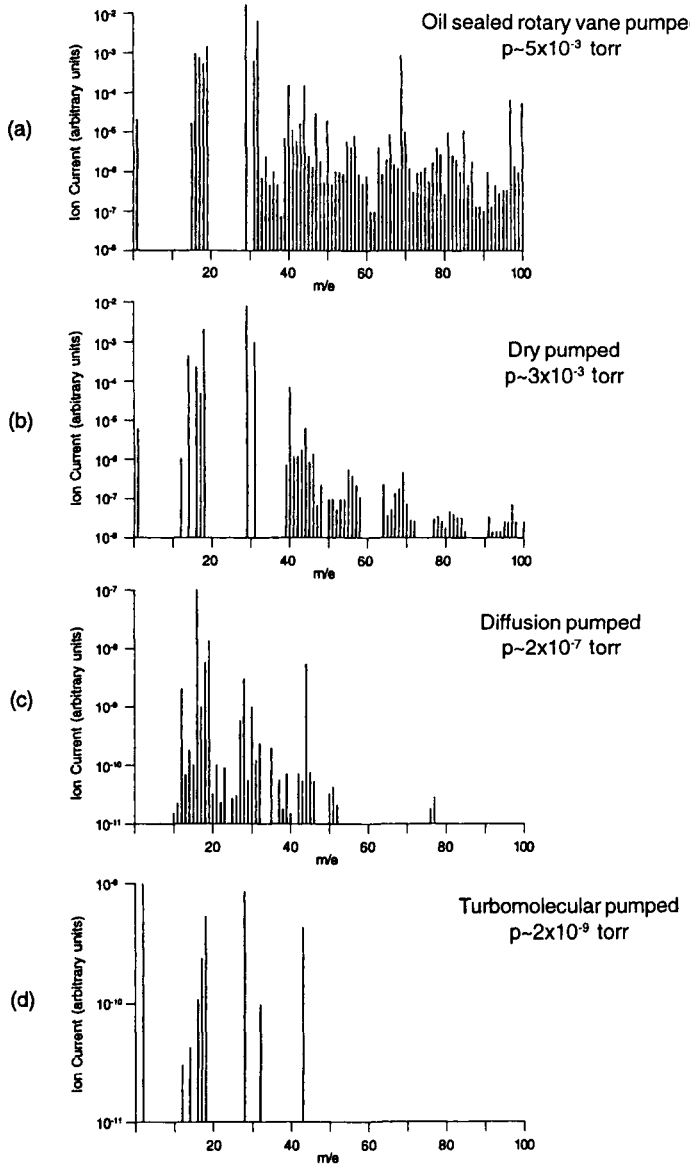
3.3.2 EXPECTED GASES IN A VACUUM SYSTEM

The operational history of a vacuum system determines what residual gases exist in that vacuum system. For example, of what sort of materials is the system made

and what materials have been added to the system? What treatment did the materials receive? Have the materials been chemically cleaned (what chemical process?), electropolished (what process?), vacuum fired (at what temperature, time, and vacuum level?), baked out (at what temperature, time, and vacuum level?), or given other processes that will leave a tell-tale spectrum? Exposure to various gases and environments in normal operation usually always leaves a characteristic adsorbed residue and associated vapor. It may not be an exact representation of the previous gas exposure because of surface chemistry, but may be some reaction product. The longer the exposure to a particular high pressure, the more the residual effect due to that exposure. Water vapor (m/e = 18 amu) is the dominant residual gas after exposure to an atmosphere of air because of the strong, multilayer, physical adsorption that occurs on the oxide surfaces of the system walls and components contained therein. If it is a rainy day and the laboratory air handling system does not substantially reduce the partial pressure of water vapor in the laboratory, one will see a much larger residual water peak in the vacuum system simply because of the longer time for accumulation. If a vacuum system is exposed to the vapors of someone having an aromatic lunch, or exposed to a strong aftershave, perfumes, or even cleaning solvents, the organic molecules will certainly end up adsorbed on the surfaces of the vacuum system and a residual organic partial pressure will be observed after pumpdown. A newly painted room is another obvious disaster for open vacuum chambers! The usual thumb rule is that if you can smell it, then the system is already very contaminated.

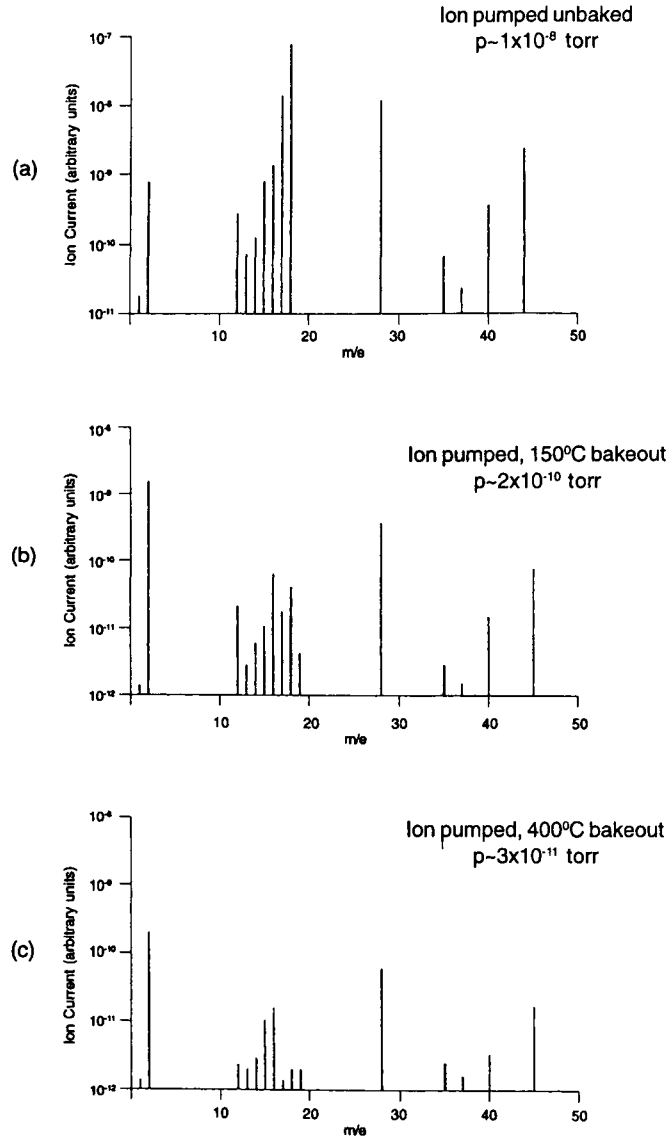
Another very important issue is the type of pumping employed. Consider, for example, the usual stainless steel system pumped down by different pumping mechanisms. If the resulting ultimate pressure is acquired by an untrapped rotary vane mechanical pump (the ultimate pressure $\sim 1 \times 10^{-3}$ torr is too high to be directly sampled by most mass spectrometers, but can be observed indirectly by differential pumping of the instrument), the spectrum will show a large organic signature of pump oils and water vapor. A most undesirable level of contaminant vapors is shown in Figure 1(a). The pump oils will generally have a characteristic series of fragment peaks at 43 and 57 amu. If the system is dry-pumped, a much lower level of residual organics will be observed, primarily because the signal is vapor from the bearing greases of the pump—see Figure 1(b). High vacuum ($< p \sim 1 \times 10^{-6}$ torr) obtained by using an untrapped diffusion pump shows a substantial reduction in the water vapor but a significant level of organic signature from the type of diffusion pump oil that is used (Figure 1c). If a turbo pump is used—Figure 1(d)—the organic peaks are much less prevalent. The dominant partial pressure is again water vapor, but the second most prevalent signal is now CO and CO_2 at mass 28 and 44 amu respectively. Of course, as the total pressure is reduced the sum of all the partial pressures is also reduced. The impact of system bakeout is shown in the next series of figures. Figure 2(a) is a spectrum of an unbaked ion pumped system with an ultimate pressure of 1×10^{-8} torr. A modest bakeout of 150°C will remove a large fraction of the water vapor and

Fig. 1.



(a) The mass spectrum represents the gas phase of a system evacuated by a rotary vane mechanical pump ($p \sim 5 \times 10^{-3}$ torr). (b) Mass spectrum of a system evacuated by a dry pump ($p \sim 3 \times 10^{-3}$ torr). (c) Mass spectrum of a system evacuated by a diffusion pump backed by a rotary vane mechanical pump ($p \sim 2 \times 10^{-7}$ torr). (d) Mass spectrum of system evacuated by a turbomolecular pump backed by a rotary vane mechanical pump ($p \sim 2 \times 10^{-9}$ torr). All systems are unbaked. (Reprinted with permission of Edward's High Vacuum.)

Fig. 2.



(**a**) The mass spectrum represents the gas phase of an unbaked system evacuated by an ion pump roughed by a molecular drag pump ($p \sim 1 \times 10^{-8}$ torr). (**b**) The same system baked at 150°C for several days ($p \sim 2 \times 10^{-10}$ torr). Note the reduction in water at $m/e = 18$. (**c**) The same system baked at 400°C for 4 days ($p \sim 3 \times 10^{-11}$ torr). Note that the water peak has virtually disappeared.

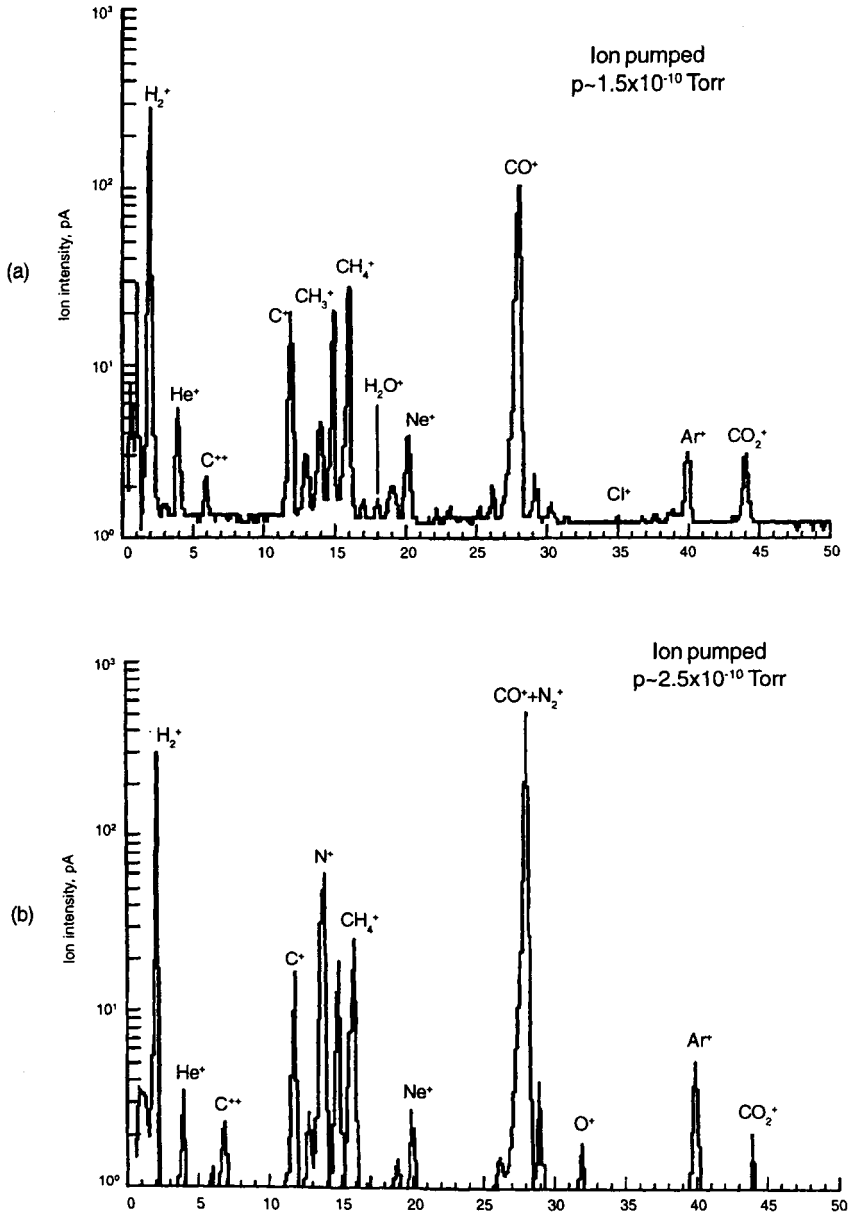
other condensables, but is not sufficient to remove all the water because surface long range forces—see Figure 2(b). The predominant peak is now hydrogen. Some small peaks like Cl (m/e = 35, 37), CH_4 (parent peak m/e = 16) and CO (m/e = 28) are also present. A high-temperature bakeout ($T > 400°C$) of a stainless steel system for four days is sufficient to desorb almost all the water and it is no longer significant at room temperature—see Figure 2(c). As long as one is using hot filaments for ion gauges, ion sources in mass spectrometers or for other instrumentation, the carbon and oxygen-bearing peaks will be clearly apparent. Hot filaments also serve as heating sources that keep the surrounding stainless steel envelope hot (~200°C for a 1.5-inch I.D. [inner diameter] tube) so the outgassing peaks will be significant as well. Systems that have elastomer seals will, in fact, have an additional concern in that the O-rings are permeable to helium, hydrogen, carbon dioxide, and other gases that are not permeated by all-metal seal systems [7,8].

Finally, some systems will have or acquire a leak from faulty materials, welds, seals, bakeouts, or some inadvertent blow to a system component like an electrical feedthrough or a flange seal. Figure 3(a) shows a UHV system that has a residual pressure of 1.5×10^{-10} torr that is leak free, and Figure 3(b) shows this same system with a small leak that increased the pressure to 2.5×10^{-10} torr. Note that the peak at m/e = 14 is much higher than that observed in Figure 3(a). Normally, when there is no leak, no admission of nitrogen into the system, and the application does not generate nitrogen, then no contributions to m/e = 28 or m/e = 14 peaks should occur (CO^{++} contributes slightly to m/e = 14)! When there is no leak, there is only the residual CH_2 at m/e = 14 due to the fragmentation of methane, which is only 15% of the parent peak at m/e = 16. Thus, for a small residual methane peak, the CH_2 is, in general, quite small. However, when there is a leak, the atomic nitrogen fragmentation peak at m/e = 14 becomes quite significant and is usually larger than the 16 series peaks due to methane. One cannot separate the N_2 and CO_2 signals at m/e = 28 with most vacuum system mass spectrometers so the existence of the nitrogen is not immediately clear. The fragmentation to atomic nitrogen is, however, tell-tale and serves as an excellent way of quickly examining the leak-tight integrity of the system.

3.3.3 THE ION GENERATION PROCESS

The generation of ions from solids, liquids, and gases for mass spectral analysis can proceed by a variety of different methods. For example, Knudsen cells raise the temperature of the solids and liquids to create a vapor pressure sufficiently high for electron impact ionization of the gas phase; laser or a spark sources di-

Fig. 3.



(a) Analog quadrupole spectrum of a UHV system that is leak free. Note the $m/e = 14$ peak, which is N and the methane fragment of CH_2. **(b)** The same system with a small leak. The unresolved $m/e = 28$ (N_2 + CO) and the $m/e = 14$ (N) peaks have increased, but the $m/e = 16$ (CH_4) peak has not.

rectly generate ions by localized vaporization at the point of photon/electron and solid impact; glow discharges create ions that bombard the surface and sputter the material of interest into the gas phase. However, for almost all gas-phase measurements in vacuum systems, the ions are created by electron impact from electrons that are generated by thermionic emission. The myriad of different filament emitters has been addressed in the preceding section (see Chapter 3.2, Lieszkovszky).

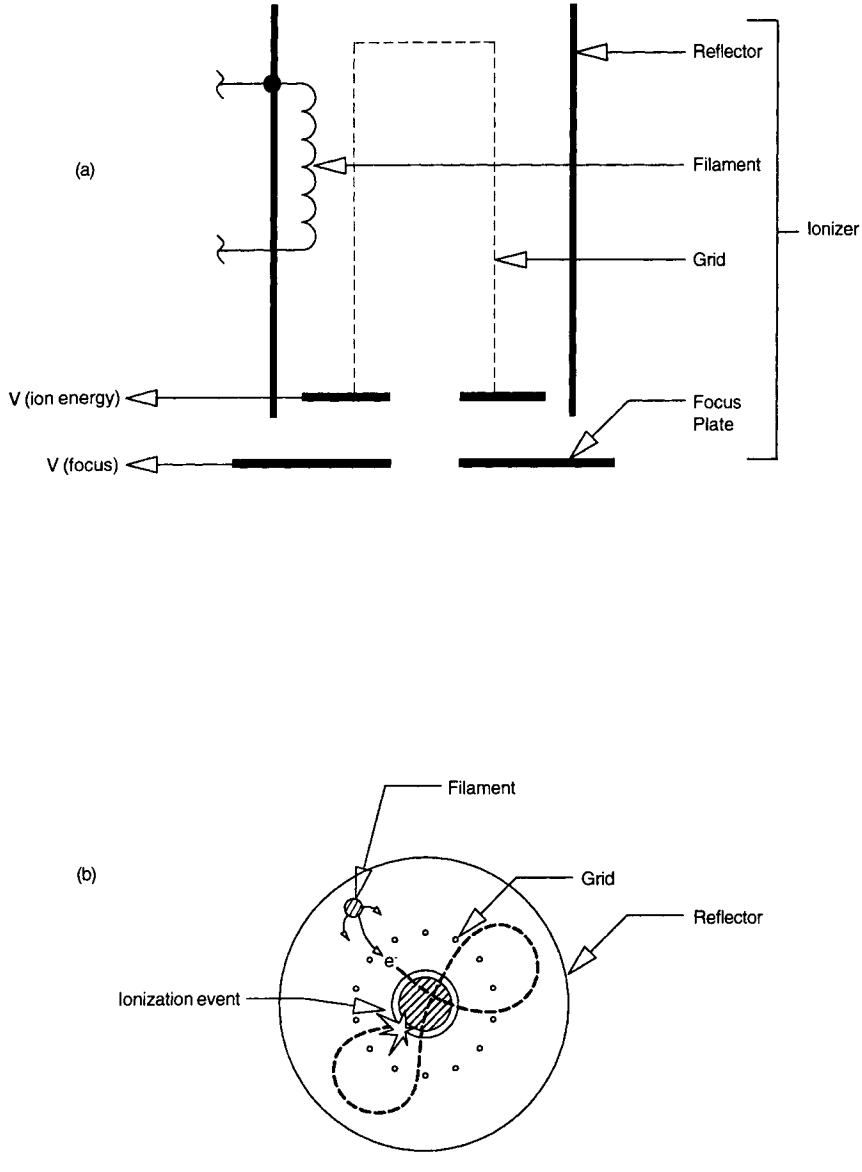
3.3.3.1 Ionization Efficiency

The efficiency of ionization is a strict function of the ion source design. An axial type of ion source (similar to an ionization gauge) is shown in Figure 4(a). The axial design is similar to a total pressure ionization gauge with the exception there is no axial collector and the ions are extracted through an aperture plate at the base of the grid. The electrons are accelerated toward the grid and normal to the axis of the grid where they make an orbital trajectory that passes through the grid about five times on the average — see Figure 4(b). The longer the path length, the higher the probability of an ionization event and therefore the higher the sensitivity. The yield of ions (J_0) exiting the aperture to the flux of atoms and molecules (J_i) entering the ion source can be roughly estimated from,

$$\frac{J_0}{J_i} = \frac{Sk_BT}{Aev_i} \tag{1}$$

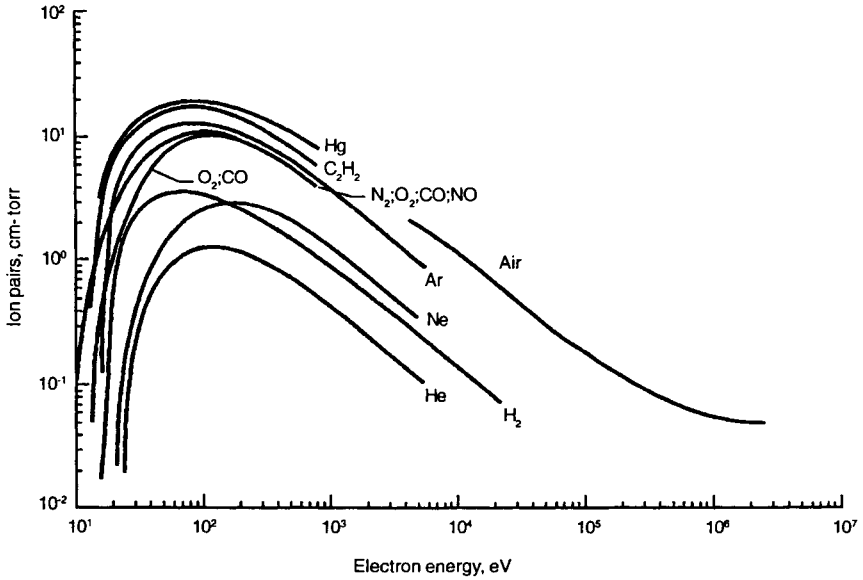
where S is the ion source sensitivity, k_B is Boltzmann's constant, T is the gas temperature, A is the extraction aperture, e is the electronic charge and v_i is the molecular velocity [9]. As an example, the number of ions formed per molecule for molecular nitrogen at room temperature and the ion source sensitivity for the source configuration in Figure 4 (S $\sim 1 \times 10^{-4}$ amperes torr^{-1}), is approximately 4×10^{-6}. Or said another way, one ion is formed for approximately 4×10^5 molecules that move through the ion source. Figure 5 shows the probability of ionization for positive ions of several gases as a function of electron energy. Negative ions are also formed, but in far less abundance. The probability of ionization for all the gases is a maximum between 50 and 200 eV and the usually followed custom is to operate ion sources at 70 eV. The ions that are formed within the grid shown in Figure 4 experience a negative extraction potential leaking through the aperture (along the axis), which causes the ions to accelerate through the aperture into the analyzer with typically about 10–30 eV kinetic energy (depending, of course, on the applied potentials). The angle of entry into the mass analyzers is, of course, critical to the resulting resolution and is a function of the type of analyzer.

Fig. 4.



(a) Typical axial ion source. (b) Electrons make numerous oscillations in and out of the grid region before encountering and ionizing a molecule.

Fig. 5.

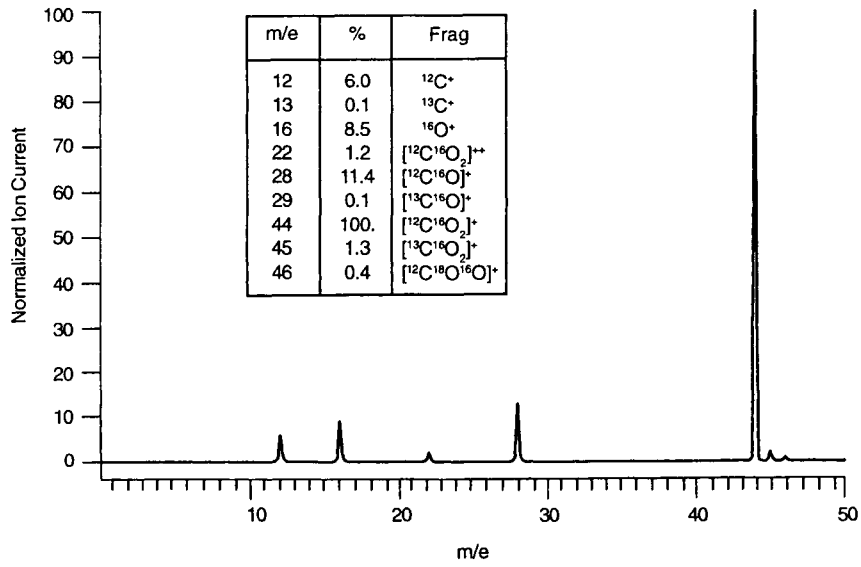


Ionization cross sections for several gases have a most probable level at between 65–100 eV. Prominent residual gases such as N_2, Ar, O_2, CO, CO_2, and H_2O all have very near the same most probable peak (within ~20%).

3.3.3.2 Fragmentation

Since the first ionization potential for most atoms and molecules extends from only a few eV to 24.6 eV for helium (second ionization potential extends to about 54 eV), there is sufficient electron energy to give a spectrum of multiply ionized species. The higher the ionization potential, however, the less probable the ionization, so the signals for multiply ionized species are substantially smaller (e.g., Ar^{++} is about 2% of Ar^{+}) [10]. Additionally, since the electron energy is substantially greater than the formation energy of most molecules, dissociation occurs as well, giving rise to fragments of the original molecule. Thus, for a constant electron energy and ion source configuration (electron trajectory length), a particular fragmentation pattern will result. These patterns are also called *fractionation patterns*, *cracking patterns*, or *spectrum fingerprints*. It is this characteristic mass spectrum of a gaseous environment that helps in identifying unknown species. As an example, $^{12}C^{16}O_2$ in an ion source will have quite detectable signals at CO_2^{+} (m/e = 44), CO_2^{++} (m/e = 22), CO^{+} (m/e = 28), O^{+} (m/e = 16) and C^{+} (m/e = 12). The gas fragmentation pattern showing the relative peak heights of CO_2 is shown in Figure 6. In addition to the major signals observed

Fig. 6.



The gas fragmentation of CO_2 including the significant isotopes and doubly ionized species.

here, the isotopes ^{13}C and ^{18}O, some doubly ionized species and some negative ions will contribute other fragments (observed only with an instrument of high sensitivity). Another source of fragmentation is the dissociation of the gas molecules that occur at the hot filament. Molecular collisions with the filament and with the ion source elements in the vicinity that are at elevated temperature will add to the fragmentation. The question of the relative fraction of fragmentation caused by temperature compared to that caused by electron impact is very much a function of filament and ion source element materials, surface areas, and temperatures, but will not be further discussed here. In the previous chapter, Lieszkovszky presents a number of fragmentation patterns for various gases [11], therefore, similar data will not be repeated here. A more extensive collection of characteristic spectra can be found in Mass Spectra Data or NIST Mass Spectral Library of Compounds [12].

3.3.3.3 Instrument Outgassing

Perhaps the most significant errors that occur in the ion source are from the surfaces of the elements that make up the ion source, i.e., electron emission source, ionization grid, insulators, extraction plates, focus plates, and the walls that surround the ion source. Most mass spectrometers have hot filaments as an electron

source, which may operate at a temperature up to 2000°C. The radiant heat to the surrounding elements can increase their temperature to greater than 200°C, which, in turn, substantially increases the outgassing from the aforementioned elements and more often than not contributes more outgassing to the system than from the cold surfaces from the rest of the system. Thus, at ultimate pressure, the mass spectrometer may be measuring its own generated pressure and not a true representation of the system pressure. Closed ion sources are much worse in providing nonrepresentative spectra in that they have less conductance available for efficient cleanup during bakeout or degassing. Although normally at room temperature, outgassing from the analyzer section and detector section can also be a significant contributor to the total instrument outgassing flux because most are closed designs (surrounded by shields), which tend to limit efficient degassing. Pumping of these volumes can only be done through the small conductance of the ion source extraction aperture back into the ionization volume and into the system. The obvious solution set to these problems is to use instrument materials (especially in the vicinity of that hot filament) that have been vacuum-fired so that the bulk hydrogen and other dissolved species are substantially reduced, to design all the instrument components with a very open structure, and to degas the surfaces by external bakeout and/or internal electron bombardment. Some mass spectrometers do not have internal degas capability so one often depends on external bakeouts. Usually, the bakeout temperature employed is so low (<250°C, limited by the construction materials in the instrument) and the bakeout time kept so short (<24h), that the instrument still contributes too much outgassing to the system pressure. For UHV applications, a bakeout of 400°C for four to five days followed by several days at 150°C at full operating emission current is desirable. A simple way to ensure that the bakeout time and temperature have been sufficient to properly clean the mass spectrometer is to turn off the filament and wait until the total pressure indication of the system ionization gauge has reached steady-state level. If the relative decrease is more than 10% of the original reading of that gauge, the mass spectrometer is contributing too much gas load to the system for a representative measurement and must be further cleaned. Of course, the use of cold-cathode sources would substantially reduce these problems and thus is highly desirable, but does not obviate the need for the aforementioned material processing and degassing.

3.3.3.4 Surface Chemistry

The operation of the hot filament has many unwanted effects from the generation of gases that are not representative of the true system pressure. The filament temperature may be as high as 2000°C for tungsten (depending on the emission cur-

rent) and, as previously mentioned, will heat the surrounding materials, including the walls of the chamber, to much higher than room temperature where, in addition to higher outgassing rates, a variety of surface reactions occur. In general, the filament will crack or dissociate most incident molecular gases, leaving a veritable plethora of energetic species to promote chemical reaction at the filament and vicinal surfaces. In addition to the species produced by neutral reaction at hot surfaces, the chemistry is further complicated by reactions from positive and negative ions and the range of excited states generated by electron impact. At these elevated temperatures, the hydrogen, carbon, sulfur, alkalis, and other interstitial impurities in the bulk of the filament and adjacent surfaces rapidly diffuse to the surface and are available for chemical combination with existing surface species or the incident gases. All the surfaces have oxides, so oxygen is readily available. Carbon has been shown to segregate to the surface of stainless steel at bakeout temperatures of only 250°C [13]. Ubiquitous hydrogen is continually present at the surface because it is the most prevalent atom dissolved in metals. Another primary source of oxygen and hydrogen can be from system gases such as H_2O. The end result is that the filament temperature creates atomic oxygen and hydrogen, which result in the generation of CO, CO_2, CH_4 and other species as well. Some probable reactions at a hot W- or ThO_2-coated W filament are [14,15]:

$$H_2 + W(s) \rightarrow H + H + W(s) \tag{2}$$

$$ThO_2/W(s) \rightarrow Th/W(s) + O/W(s) \tag{3}$$

$$O/W(s) + C/W(s) \rightarrow CO + W(s) \tag{4}$$

$$W(s) + 3H_2O \rightarrow WO_3\,(v) + 6H \tag{5}$$

$$WO_3 + M(s) \rightarrow WO_3/M(s) \tag{6}$$

$$O_2 + 2C/W(s) \rightarrow 2O/W(s) + 2C/W(s) \rightarrow 2CO + W(s) \tag{7}$$

$$O_2 + C/W(s) \rightarrow CO_2 + W(s) \tag{8}$$

$$O_2 + 4H/W(s) \rightarrow 2H_2O + W(s) \tag{9}$$

$$C_xH_y + W(s) \rightarrow C/W(s) + H/W(s) \tag{10}$$

Clearly, however, many other surface contaminants are available, so that in less clean systems, the generated species are far more complex and far more extensive. If lower work function filament materials (such as LaB_6) or coatings (such as ThO_2) are used, then the filament temperature (1200–1400°C) and the vicinal surface temperatures are reduced, which, in turn, results in reduced surface chemistry. Another problem is the admission of reactive gases to the system, which will further exacerbate the chemistry and the complexity of undesired foreign species.

3.3.3.5 Electron Stimulated Desorption

The ion sources and the electron multipliers of mass spectrometers experience electron impact collisions with the grid (or collector) and with multiplier surfaces, respectively. Since all surfaces have oxides, carbon, and other adsorbed species, there is a probability of desorption of these species into the gas phase. Electron-stimulated desorption (ESD) has been a major scientific area in surface science for many years since first studied by Redhead and by Menzel and Gomer [16,17]. Later, different mechanisms for this phenomena were described by Knotek and Feibleman and by Antoniewitz [18,19]. Several good reviews exist on this subject and give a comprehensive discussion of the area [20,21]. Essentially, the incident primary electrons from the filament impact the ion source grid or electron collector and excite the adsorbed atoms or molecules to anti-bonding states where they desorb as neutrals or, to a much lesser extent, as ions and become part of the gas phase—see Figure 7(a). The cross sections for the particular mechanism vary from $\sim 1 \times 10^{-16}$ to 1×10^{-22} cm^2 and the mean kinetic energy for the ions has been found to be often around 5 eV [20]. The ESD neutrals are ionized by the primary electron flux and, along with the ions, are analyzed as if they were part of the representative gas phase. The adsorbed state of the surface species controls what is the operative mechanism of ESD. The flux, J_0, of desorbed species has been observed to follow

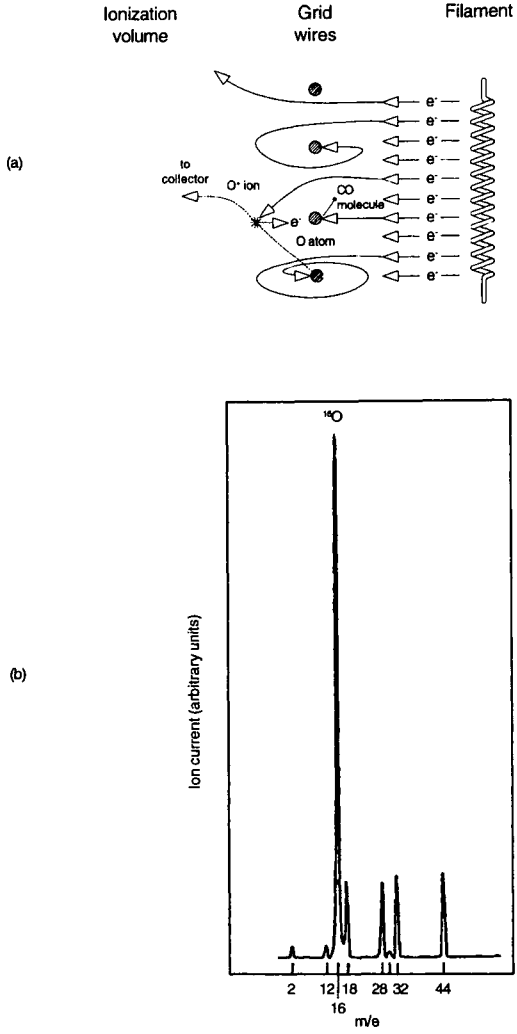
$$J_0 = Q\,\theta\sigma_0\,J_{e^-} \tag{11}$$

where Q is the ESD cross section for the adsorbed state, θ is the coverage, σ_0 is the monolayer concentration, and J_{e^-} is the electron flux. For a given cross section, the ESD flux is directly related to the electron flux and the surface concentration of adsorbed species. For example, if the adsorbed species on the grid of the ionizer has a coverage equal to 1, a conservative ESD cross section of 1×10^{-18} cm^2 is employed and the electron emission is 1 mA, then the flux of molecules into the ionizing region is $\sim 6 \times 10^{12}$ molecules cm^{-2}s^{-1} or 2×10^{-7} torr liters cm^{-2}s^{-1}. This flux is the equivalent of >10 times the outgassing of unbaked stainless steel, clearly indicating that substantial errors can result!

The origin of atomic H and O (and the recombined molecules H_2 and O_2), H_2O, CO and CO_2 detected by the mass spectrometer is quite difficult to establish. The question is simply what is due to ESD and what is truly representative of the gas phase. Another often observed ESD peak is at $m/e = 19$, which can be F^+ or H_3O^+. The source of the peak may less likely be fluorine than hydronium, which readily forms because there is an ample supply of oxygen and hydrogen.

A dramatic example of ESD-generated oxygen atom neutrals occurs when the grid of an ion source is covered with a thin film of Ag. Silver is well known to dissociatively adsorb oxygen at room temperature, so the grid surface is covered

Fig. 7.



(a) Schematic of the electron-stimulated desorption (ESD) mechanism in an ion source of mass spectrometer. (b) ESD-generated atomic oxygen signal from Ag-coated grid of the ion source.

with atomic oxygen. The spectrum detected in the mass spectrometer is shown in Figure 7(b) and gives the impression that the predominant gas-phase species is $m/e = 16$. Since there is no $m/e = 15$ peak to speak of, this cannot be confused with methane and since there is a much smaller $m/e = 32$ peak, this is not a fragmentation peak of molecular oxygen (11.4% of the parent peak). It is clearly from the ESD of atomic oxygen from the grid into the ionizing volume of the mass

spectrometer. When the grid was degassed above 400°C, the 16 peak substantially diminished, as predicted from studies of atomic oxygen adsorbed on Ag [22]. Replacement of the grid with a standard molybdenum grid returned the spectrum to the more conventional series shown in Figure 3. Minimizing this unwanted gas desorption can be accomplished by operating at low emission currents, minimizing the grid or collector area, and using materials that have a low propensity for adsorption.

3.3.3.6 Charge Exchange Mechanisms

In addition to electron impact ionization, dissociation, recombination, sputtering and ESD, charge transfer can also affect the actual ionization volume particle mix. At normal operational pressures for mass spectrometers ($p < 1 \times 10^{-4}$ torr), the mean free path (l) is sufficiently long that ion-molecular impact is not a significant process ($l < 50$ cm), but it does occur and can be a significant signal in some applications. Essentially, the collision of a positive ion with a gas atom or molecule results in an exchange of energy, momentum and direction, but also a change of charge can occur. An energetic ion can "steal" an electron from a neutral and become an energetic neutral, leaving behind a slow ion. Actually, the cross sections for charge exchange vary substantially with the difference in mass between particles. The transfer is more likely to occur in the proximity of resonance or the smaller the energy exchanged in a collision. Thus, if the incident ions are the same species as the target atoms or molecules, the exchange has a higher probability. The amount of energy transferred is equal to the sum of the kinetic and ionization energies plus the excitation energy of the target molecule (including vibration and rotation) less the neutralization energy. The charge transfer may also include other processes such as dissociation, excitation (with subsequent light emission), and ion formation. Some examples of exchange are [23,24]

$$He^+ + H_2O \rightarrow He + H_2O^+ \quad (12)$$

$$H_2^+ + H_2O \rightarrow H_2 + H_2O^+ \quad (13)$$

$$H_2O^+ + H \rightarrow H_3O^+ \quad (14)$$

$$H_2 + H^+ \rightarrow H_3^+ \quad (15)$$

$$Ar^+ + Ar \rightarrow Ar + Ar^+ \quad (16)$$

$$N_2^+ + Ar \leftrightarrow Ar^+ + N_2 \quad (17)$$

$$He^+ + Xe \rightarrow He + Xe^+ \quad (18)$$

Since hydrogen and water are so prevalent in vacuum systems reactions, (12) to (15) are often observed. The production of H_3^+ is frequently seen in ion sources especially with environments of hydrogen-containing species, such as hydrocar-

bons. Argon and other inert gases have been used for many years in charge neutralization cells for forming energetic neutral beams of atomic oxygen and other species.

3.3.4 TECHNIQUES FOR ANALYSIS

3.3.4.1 Partial Pressure Measurement

The accuracy and overall performance of the mass spectrometer is a function of each section of the mass spectrometer. As an example, consider a quadrupole-type instrument that is primarily comprised of an ion source, a set of quadrupole rods for mass filtering, and an electron multiplier for signal amplification (see Figure 8).

THE ION SOURCE

For a given ion source geometry, the ionization efficiency is primarily a function of the number of electrons in the atom or molecule. As previously indicated, the highest probability for ionization of most gases occurs at approximately 70 eV. The quantity of ions formed and extracted from the source is a function of the design of the ion source and is characterized by a source factor (ionization efficiency). Flaim and Ownby [25] have determined an equation that approximates the ionization efficiency (I).

$$I = 4.3 \times 10^{-2} \sum_{i} n_i Z_i + 0.4 \tag{19}$$

where Z_i is the number of electrons in an atom of molecule i and n_i is the number of atoms in molecule i. As an example, consider the water vapor molecule, H_2O; it has 2 atoms of H (1 electron) and 1 atom of O (8 electrons), thus

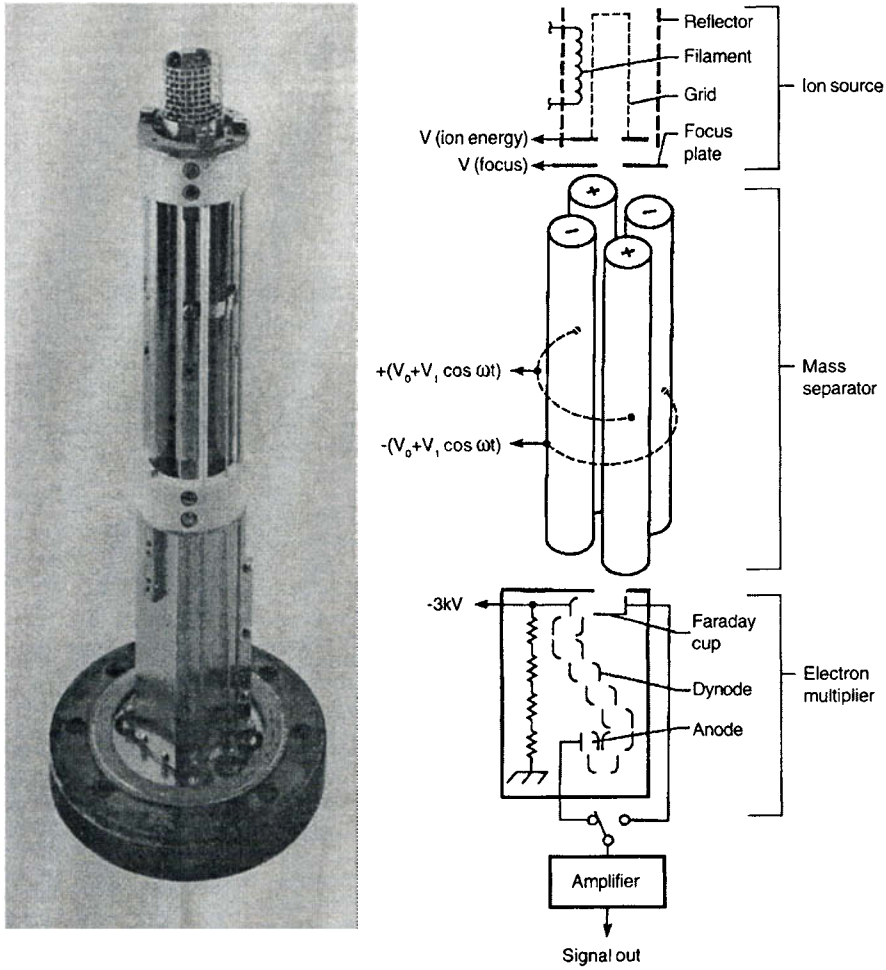
$$I[H_2O] = 4.3 \times 10^{-2} [2(1) + 1(8)] + 0.4 = 0.83$$

which is the ionization efficiency relative to nitrogen. Figure 9(a) shows the fit of the equation for several atoms and molecules. There is also a geometrical factor similar to a gauge constant that reflects the efficiency in which the ions are extracted from the source and directed to the mass filter.

THE MASS FILTER

When the ions enter the mass filter, those ions that match the selected filtering conditions—i.e., those which have the right mass to charge ratio, entrance vec-

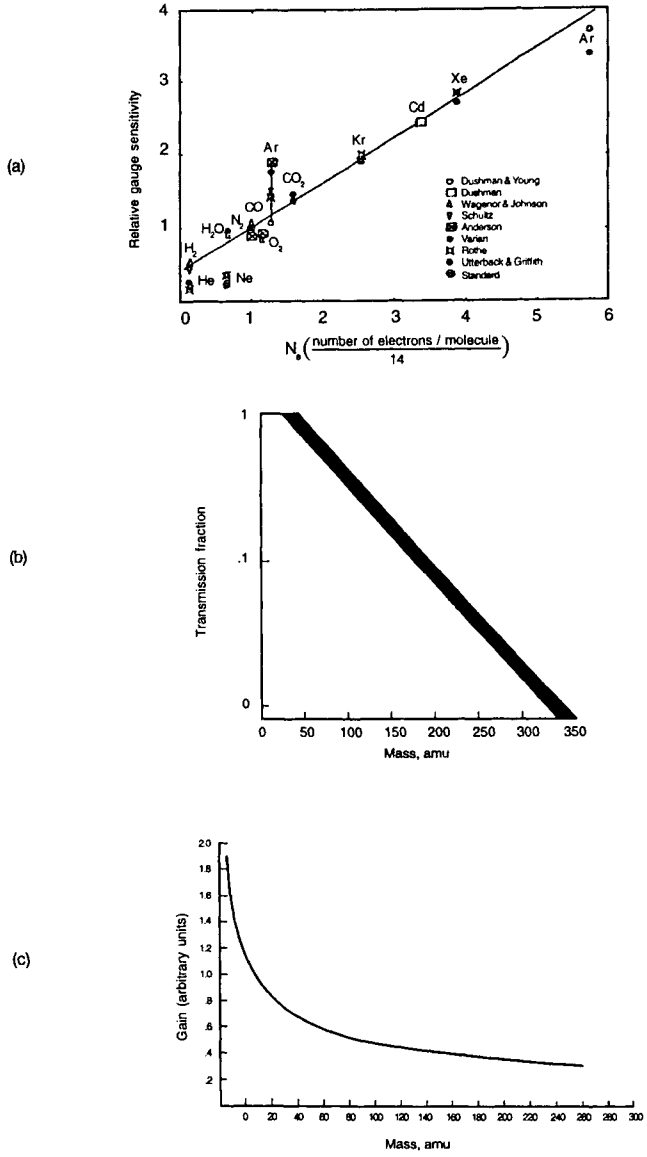
Fig. 8.



Actual and schematic sections of a commercial mass spectrometer. (Reprinted with permission of UTI division of MKS.)

tor, and kinetic energy—have a particular probability of exiting the filter [26]. In a quadrupole filter, this transmission ($T_{m/e}$) factor is essentially 1 for the light masses (1–20 amu) and then decreases logarithmically with mass. Figure 9(b) shows the typical behavior as a function of mass. Note there is a band for masses higher than 20 amu that takes into account variances in molecular degrees of free-

Fig. 9.



(a) General fit of ion source sensitivity versus number of electrons per molecule (after Flaim and Ownby [25]). **(b)** Representation of the transmission probability versus m/e. Shaded area represents error bar. **(c)** Relative gain variation of an electron multiplier as a function of mass.

dom, ion velocity (time in the filter), rf frequency, and potential uniformity in the radial and axial dimensions. As shown in the figure, the transmission decreases to about 0.5 at 75 amu.

THE ION DETECTOR

On exiting the mass filter, the ions are collected by a Faraday cup or amplified by an electron multiplier. A well-designed Faraday cup should completely collect all ions and not gain or lose any signal due to secondary emission of electrons. An electron multiplier (continuous dynode, box and grid, venetian blind, multichannel plate, etc.) can provide a gain (G) from 1×10^3 to 1×10^8 depending on the number of electron collision stages and on the secondary emission coefficient of the dynode or collision materials. Most often the gain is adjusted to a modest value, e.g., $G \sim 1 \times 10^5$ or less in order to ensure long-term stability. Further, the gain is very sensitive to ion mass, chemical reactivity, kinetic energy, and total charge. Some multipliers have been observed to have a very high gain for low masses and decrease as a function of $m^{-1/2}$ with increasing mass [27]—see Figure 9(c). In addition, the secondary emission coefficient from impacting ions varies as the $\cos^2 \theta$, where θ is the angle of incidence of the ion. A calculation of the partial pressure for a given species can be determined from

$$p = \frac{1}{I} \sum_i \left[\frac{i}{G_i T_i} \right] \tag{20}$$

Considering an actual case where the partial pressure of H_2O has been measured one can compare the calculated value. The primary peaks are $m/e = 18$ (H_2O^+) and the fragmentation peak is $m/e = 17$ (OH^+). If the peak height currents are measured to be 1×10^{-8} and 2.3×10^{-9} A, respectively, and the ion source sensitivity is 1.5×10^{-3} A/torr and the electron multiplier gain for N_2 is 1×10^5 then one can calculate a value for the partial pressure for H_2O. The G for H_2O is

$$G[H_2O] = G[N_2] \left[\frac{m(N_2)}{m(H_2O)} \right]^{1/2} \tag{21}$$

$$= 1 \times 10^5 \left(\frac{28}{18} \right)^{1/2} = 1.25 \times 10^5$$

and for the fragment OH is

$$G[OH] = 1 \times 10^5 \left(\frac{28}{17} \right)^{1/2} = 1.28 \times 10^5$$

Since $T \sim 1$ for <20 amu, then $T[H_2O] \sim 1$, $T[OH] \sim 1$

Employing Equation (20), the calculated partial pressure is

$$p[H_2O] = \frac{1}{(1.5 \times 10^{-3})0.83}\left[\frac{1 \times 10^{-8}}{1.25 \times 10^{5}(1)} + \frac{2.3 \times 10^{-9}}{1.28 \times 10^{5}(1)}\right]$$

$$p[H_2O] = 7.9 \times 10^{-11} \text{ torr}$$

The actual measurement from the calibrated mass spectrometer was

$$p[H_2O] = 1.5 \times 10^{-10} \text{ torr}$$

Although the approximate answer from the calculation is instructive and useful, there is still significant error. Clearly, the calculation does not contain every correction necessary to make an accurate assessment of the partial pressure! In general, the ion source limitations, the filter and detector errors, and anomalies make quantitative measurements quite difficult. Calibrations in the selected gases of interest are absolutely essential to obtain meaningful data.

3.3.4.2 Gas Identification

Two primary techniques enable the user to unambiguously analyze an unknown gas species or mixture of gas species to clearly quantify the individual fractions of each. They are the isotope analysis [28] and the fragmentation analysis [29]. In most general instances, a combination of both methods will be employed. Although isotopic analysis is a useful technique, it cannot supplant the use of fragmentation analysis. Fragmentation analysis typically works from a much greater number of larger-amplitude mass peaks and is thus likely to give a more complete rendition of the species present. Obviously, to get an accurate species identification, a calibration of peak height for parent and fragments (using the instrument making the analysis) is required. Even without the calibration the spectral standards are useful as guides for determining species that may be present in the system being analyzed. When analyzing for species, one should always attempt isotopic analysis first, followed by fragmentation analysis for confirmation.

ISOTOPE ANALYSIS

The main advantage of using the isotope method is that normally no calibration of the analyzer is required for identification of a species, although it would be required for quantification of results. The isotopic abundances have been accurately measured for the elements of interest. In using isotopic identification, it is important to keep in mind several factors. The analyzer detector system must be capable of linear output over three to four orders of magnitude for best results because the

less abundant isotopes of most of the elements of interest are at the 0.1 to 1.0% level. Modern amplifiers are linear and stable over five–seven orders of magnitude or better, so this factor should not be an issue. It is also important that the analyzer sensitivity not be strongly dependent on the mass number. Most quadrupole instruments have a mass sensitivity that decreases at 1–2% per amu beyond 12 amu. Most of the isotopic species of interest appear above this range. Species appearing below the linear range are uniquely identified on other bases. Finally, due to the low peak height of less abundant isotopic peaks, it is important that they not interfere with other fragments. Ordinarily this is not a problem because isotopes of species of interest do not overlap and with few exceptions, the less abundant isotopes appear at higher mass numbers. One exception is that due to argon. Natural Ar at 40 amu is the most prevalent isotope and is the decay product of an abundant element, thus $^{40}Ar/^{36}Ar = 300$. The abundances of the isotopes of all naturally occurring elements are given in the *American Institute of Physics Handbook*, Table 8.6.1 [30]. Most analyses in vacuum systems and on vacuum processes involve only the lightest 18 elements in the periodic table.

The results of an isotopic analysis at the 1% level should be within a few percent of the level predicted. At the 0.01% level, the results should be within several tens of percent of the level predicted from natural abundance. If one arrives at values outside these ranges, then interference is surely present and an isotopic analysis will need to be supplemented with fragmentation analysis.

When looking at pure elemental gases, such as nitrogen and oxygen, one may find the heavier isotopes to be slightly depleted. This occurs because these pure gases are obtained by fractionation. The lower-molecular-weight species boil at lower temperatures. The rarer higher-boiling-point species may be separated and sold at a greater price of various isotopic tracer experiments. A slight depletion can also occur if the gases have been introduced by diffusion or molecular flow, because the flow is dependent on $(m)^{-1/2}$. Any of these effects typically alter the ratios by a maximum of ~3%.

When looking at an atomic species, the expected ratio of the isotopes can be obtained directly from the aforementioned table. It is noteworthy that the rarer isotopes of the elements of primary interest—i.e., hydrogen, carbon, nitrogen, oxygen, and chlorine with an atomic number (Z) of 18 or less—are 1 or 2 amu greater than the major isotopes for the element. This means that the largest rare isotopic peaks for all molecules will appear at 1 or 2 amu greater mass than the peak of the molecule in question. To calculate the relative abundance of the isotope,

$$\text{Relative abundance (\%)} = \frac{\text{peak height of isotope species}}{\text{peak height of most abundant species}} \times 100 \qquad (22)$$

For illustration, consider the spectra of Ne. The spectra for Ne at masses 20, 21, and 22 amu was observed to have peak heights of 90.92, 0.257, and 8.82 arbitrary

units, respectively. The relative abundance in percent is, then,

For ^{20}Ne: ($m/e = 20$)
$$\frac{[^{20}Ne]}{[^{20}Ne]} \times 100 = \frac{90.92}{90.92} \times 100 = 100$$

For ^{21}Ne: ($m/e = 21$)
$$\frac{[^{21}Ne]}{[^{20}Ne]} \times 100 = \frac{0.257}{90.92} \times 100 = 0.282$$

For ^{22}Ne: ($m/e = 22$)
$$\frac{[^{22}Ne]}{[^{20}Ne]} \times 100 = \frac{8.82}{90.92} \times 100 = 9.70$$

The resulting spectrum of Ne and its isotopes should therefore appear in these proportions. A second example, CO, will show peaks at masses 28, 29, 30, and 31 because CO has isotopes of both carbon and oxygen. The relative abundances should be

For $^{12}C^{16}O$: ($m/e = 28$)
$$\frac{[^{12}C] \times [^{16}O]}{[^{12}C] \times [^{16}O]} \times 100 = \frac{98.89 \times 99.76}{98.89 \times 99.76} \times 100 = 100$$

For $^{13}C^{16}O + {}^{12}C^{17}O$: ($m/e = 29$)
$$\frac{[^{13}C] \times [^{16}O]}{[^{12}C] \times [^{16}O]} \times 100 + \frac{[^{13}C] \times [^{17}O]}{[^{12}C] \times [^{16}O]} \times 100 =$$
$$\frac{1.11 \times 99.76}{98.89 \times 99.76} \times 100 + \frac{98.89 \times 0.0374}{98.89 \times 99.76} \times 100 = 1.16$$

For $^{13}C^{17}O + {}^{12}C^{18}O$: ($m/e = 30$)
$$\frac{[^{13}C] \times [^{17}O]}{[^{12}C] \times [^{16}O]} \times 100 + \frac{[^{12}C] \times [^{18}O]}{[^{12}C] \times [^{16}O]} \times 100 =$$
$$\frac{1.11 \times 0.034}{98.89 \times 99.76} \times 100 + \frac{98.89 \times 0.204}{98.89 \times 99.76} \times 100 = 0.205$$

For $^{13}C^{18}O$: ($m/e = 31$)
$$\frac{[^{13}C] \times [^{18}O]}{[^{12}C] \times [^{16}O]} \times 100 = \frac{1.11 \times 0.204}{98.89 \times 99.76} \times 100 = 0.00228$$

Clearly, in order to observe the full isotope spectrum contributing to the CO signal, the instrument must be sensitive enough to detect at least five decades of signal intensity less than that of the parent peak at $m/e = 28$ amu. If the partial pressure of CO is $\sim 1 \times 10^{-9}$ torr, one requires an instrument with a minimum detectable partial pressure of $\sim 1 \times 10^{-14}$ torr to fully see the isotope signal at $m/e = 31$. In addition to sufficient sensitivity, it is also necessary for the instrument to be linear over the full range to accurately compare the relative peak heights. Armed with this spectrum, one can now unambiguously identify this gas if it is the only species in this mass range. Unfortunately, the situation is seldom this simple unless one is backfilling with CO from a pure source. Molecular nitrogen and ethylene are also contributors to $m/e = 28$, thus requiring the decon-

volution of the overlapping data. As previously discussed, a leak-tight vacuum system will not have nitrogen as part of the residual spectrum, but there are instances where a leak may exist or some operation involving N_2 may occur. Thus, in this case, $m/e = 28$ is obviously not pure CO or pure N_2. A similar analysis to that just conducted for CO shows that pure N_2 has normalized mass peaks of 100 at $m/e = 28$, 0.75 at $m/e = 29$, and 1.4×10^{-5} at $m/e = 30$. Assume that the measured spectrum has 100 units of CO and N_2 at $m/e = 28$, 1.0 unit at $m/e = 29$, 0.14 units at $m/e = 30$ and 1.7×10^{-3} units at $m/e = 31$. The following analysis can be applied:

$$\begin{aligned} &\text{Let } x = \text{CO fraction and } y = N_2 \text{ fraction;} \\ &\text{For } m/e = 28: \qquad x + y = 100 \\ &\text{For } m/e = 29: \qquad 0.75y + 1.16x = 1.0 \end{aligned}$$

$$x = 0.61 \text{ and } y = 0.39 \text{ (61\% CO and 39\% } N_2\text{)}$$

To check if this is reasonable, the $m/e = 30$ peak can be examined to see if the expected relative fraction correctly corresponds to 61% CO. Since for comparable amounts of CO and N_2, the value of the N_2 isotope contribution at $m/e = 30$ is so small (1.4×10^{-5}), the $m/e = 30$ signal must be predominantly CO. If we divide the observed peak height of 0.14 at $m/e = 30$ by 61 at $m/e = 28$, one obtains 0.23×10^{-2}, which is in reasonable agreement with the 0.21% for $m/e = 30$ for CO. If C_2H_4 had also been expected as a contribution to $m/e = 28$, the $m/e = 30$ peak would have to be used to determine the fraction of CO in the three-way mixture. Although the fractional contributions in this example give us an idea as to the relative amounts of each species, this does not mean that the partial pressures are related in the exact same ratio. The partial pressures would have to be evaluated by calibrating the instrument for these gases. The complexity of calculating the expected values for isotopic peaks obviously increases as the number of atoms in a molecule increases. For more detailed analysis of this procedure, please see the notes of L. C. Beavis [28].

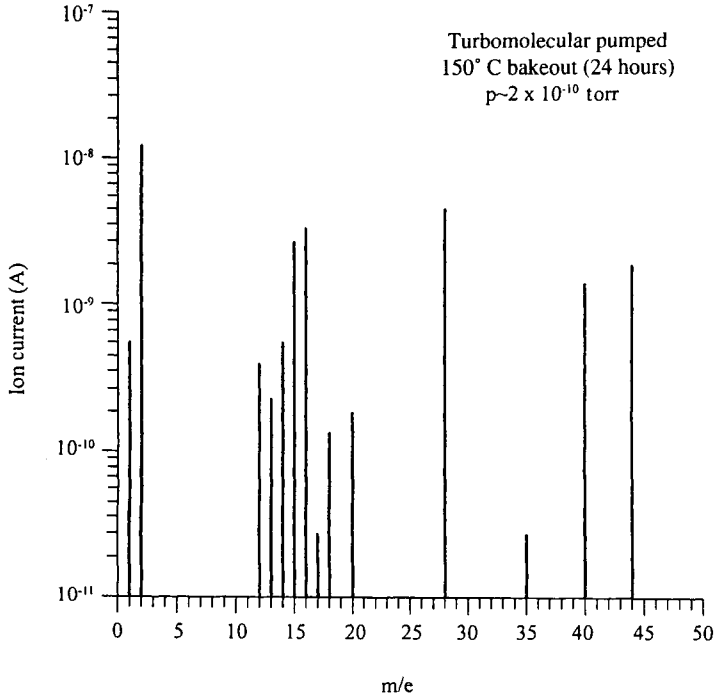
FRAGMENTATION ANALYSIS

In general, a clean, dry-pumped vacuum system will have mostly inorganic residual gases (usually with the exception of CH_4). If the system is unbaked, the primary residual gas will be H_2O and perhaps some solvent residues like acetone and/or ethanol. The inorganic gases are, in most cases, 2- or 3-atom molecules such as H_2, H_2O, CO, O_2 and CO_2 with even mass numbers. Further, ionization is more probable than is molecular dissociation, so the fragmentation patterns are dominated by the parent peak. The inert gases are stable, nonreactive atoms. If the vacuum system is pumped by an untrapped rotary vane and/or a diffusion pump, the residual gases will be dominated by organic vapors and gases. If the pumps are well trapped, organics will be present, but dominated by inorganic gases. Or-

ganics contain many more atoms than an inorganic molecule and are more likely to dissociate before ionization, so the parent peak will, very often, not be predominant. Dissociation usually occurs at the carbon–carbon single bonds, which means that a CH_2 ($\Delta m = 14$) or a CH_3 ($\Delta m = 15$) group is the first to break away. The resulting fragment molecules are usually an odd mass number. For example, ethanol (C_2H_6O) is an often-used solvent cleaner in vacuum systems and usually leaves a residual vapor. Under electron impact, the methyl group in this molecule is the the most probable dissociation fragment leaving the CH_2OH ($m/e = 31$) as the dominant peak. Subsequent to that, the single bonds associated with hydrogen are the next most likely to dissociate, so the subsequent fragmentation proceeds in descending steps of one mass unit.

Consider a stainless steel turbomolecular pumped system that has been moderately baked at 150°C for 24 h. Figure 10 shows the quadrupole spectra for such a system. The water vapor is substantially reduced and the remaining gases are primarily H_2, CH_4, CO, and CO_2. Some Ar is also present from an earlier backfill. Analysis of this system by fragmentation involves the fragmentation patterns of all these gases and they are shown in the previous section (see Chapter 3.2,

Fig. 10.



The mass spectrum of a turbomolecular-pumped system that has been moderately baked ($p \sim 2 \times 10^{-10}$ torr).

Lieszkovszky). The peak height at a given mass number, I_m, is the sum of all the peak height contributions, i_{mn}, over all gases n:

$$I_m = \sum_n [i_{mn}] = \sum_n [S_n p_n \gamma_{mn}] \tag{23}$$

where S_n is the analyzer sensitivity for the gas n, p_n is the partial pressure for gas n, and γ_{mn} is the peak height contribution of gas n at mass number m.

From the figure we can tabulate the signal peaks (in amperes) as follows:

m/e	i_{mn}(A)
44	2.20×10^{-9}
40	1.60×10^{-9}
28	5.10×10^{-9}
20	1.72×10^{-10}
18	1.30×10^{-10}
17	2.99×10^{-11}
16	3.50×10^{-9}
15	2.83×10^{-9}
14	5.15×10^{-10}
13	2.54×10^{-10}
12	3.79×10^{-10}
4	5.80×10^{-11}
2	1.10×10^{-8}
1	5.50×10^{-10}

The total pressure for this spectra is $p_T \sim 2.0 \times 10^{-10}$ torr. The overall sensitivity of the quadrupole mass spectrometer with the gain of the multiplier adjusted to 1×10^5 is 150 A/torr (N_2). The calculated sensitivities for the other relevant gases are

$$S[CO_2] = 160 \text{ A/torr}$$

$$S[Ar] = 132 \text{ A/torr}$$

$$S[CO] = 150 \text{ A/torr}$$

$$S[H_2O] = 150 \text{ A/torr}$$

$$S[CH_4] = 153 \text{ A/torr}$$

$$S[\mathrm{He}] = 83 \text{ A/torr}$$

$$S[\mathrm{H_2}] = 223 \text{ A/torr}$$

The analysis should start at the highest mass and continue in descending order of molecular mass to ensure that all fragment peaks are considered. If we ignore small contributions from isotopes and doubly ionized species, CO_2 has significant peaks at *m/e* of 44, 28, 16, and 12 amu. The parent peak is at mass 44 so the fragmentation ratio, $\gamma_{44,\,CO_2} = 1$. Since there are no other contributions to mass 44, the partial pressure of CO_2 can be determined from Equation (23),

$$p_n = i_{mn}/S_n\gamma_{mn} = p_{CO_2} = 2.2 \times 10^{-9}/160\,(1) = 1.38 \times 10^{-11} \text{ torr}$$

With this partial pressure and the fragmentation ratios (γ_{mn}) for the remaining peaks, the fragment peak heights may now be determined from Equation (23),

$$\gamma_{44,\,CO_2} = 1 \qquad i_{44} = 2.20 \times 10^{-9}\,\mathrm{A}$$

$$\gamma_{28,\,CO} = 0.114 \qquad i_{28} = 2.50 \times 10^{-10}\,\mathrm{A}$$

$$\gamma_{16,\,O} = 0.085 \qquad i_{16} = 1.87 \times 10^{-10}\,\mathrm{A}$$

$$\gamma_{12,\,C} = 0.06 \qquad i_{12} = 1.32 \times 10^{-10}\,\mathrm{A}$$

The CO_2 fragment peak heights may now be subtracted from the original signal levels to give,

m/e	i_{mn} (A)
44	.....
40	1.60×10^{-9}
28	4.85×10^{-9}
20	1.71×10^{-11}
18	1.30×10^{-10}
16	3.31×10^{-9}
15	2.83×10^{-9}
14	5.15×10^{-10}
13	2.54×10^{-10}
12	2.97×10^{-10}
4	5.80×10^{-11}
2	1.10×10^{-8}
1	5.50×10^{-10}

The next peak to be evaluated is that of Ar at mass 40 and at mass 20 (Ar^{++}). Since it has no contributions to the other species it can be subtracted by inspection. The partial pressure is calculated to be $p_{Ar} = 1.21 \times 10^{-11}$ torr.

Carbon monoxide has a parent peak at mass 28 and has only one *significant* fragment peak at mass 12. The partial pressure is $p_{CO} = 4.85 \times 10^{-9}/150\ (1) = 3.23 \times 10^{-11}$ torr from which the peak height of the C fragment peak is $i_{12} = 3.23 \times 10^{-11}\ (150)\ 0.045 = 2.18 \times 10^{-10}$ A. Subtracting from the peak heights gives

m/e	i_{mn} (A)
44	.....
40	.....
28	.....
20	.....
18	1.30×10^{-10}
17	2.99×10^{-11}
16	3.31×10^{-9}
15	2.83×10^{-9}
14	5.15×10^{-10}
13	2.54×10^{-10}
12	7.92×10^{-11}
4	5.80×10^{-11}
2	1.10×10^{-8}
1	5.50×10^{-10}

The next peak in descending order is H_2O at mass 18, with fragments at 17 and 16 amu. The partial pressure is $p_{H_2O} = 1.30 \times 10^{-10}/150\ (1) = 8.67 \times 10^{-13}$ torr from which the peak height of the fragments are $i_{17} = 2.99 \times 10^{-11}$ A and $i_{16} = 1.43 \times 10^{-12}$ A. Subtracting from the remaining peak heights gives

m/e	i_{mn} (A)
44	.....
40	.....
28	.....
20	.....

18	.....
17	.....
16	3.30×10^{-9}
15	2.83×10^{-9}
14	5.15×10^{-10}
13	2.54×10^{-10}
12	7.92×10^{-11}
4	5.80×10^{-11}
2	1.10×10^{-8}
1	5.50×10^{-10}

Methane at mass 16 has significant fragments at 15, 14, 13, and 12 amu. The partial pressure $p_{CH_4} = 3.30 \times 10^{-9}/153\ (1) = 2.16 \times 10^{-11}$ torr from which the peak heights of the fragments are $i_{15} = 2.83 \times 10^{-9}$ A, $i_{14} = 5.15 \times 10^{-10}$ A, $i_{13} = 2.54 \times 10^{-10}$ A and $i_{12} = 7.92 \times 10^{-11}$ A. Subtracting from the remaining peaks gives

m/e	i_{mn} (A)
44	.....
40	.....
28	.....
20	.....
18	.....
17	.....
16	.....
15	.....
14	.....
13	.....
12	.....
4	5.80×10^{-11}
2	1.10×10^{-8}
1	5.50×10^{-10}

The remaining peaks are unambiguous (if there is no deuterium in the system); He at mass 4, H_2 at mass 2 and the atomic hydrogen fragment at mass 1. The partial pressures are then $p_{He} = 5.8 \times 10^{-11}/83(1) = 6.99 \times 10^{-13}$ torr and $p_{H_2} = 1.1 \times 10^{-8}/223\ (1) = 4.9 \times 10^{-11}$ torr.

The sum of the resulting partial pressures in the system are

CO_2	$p = 1.38 \times 10^{-11}$
Ar	$p = 1.21 \times 10^{-11}$
CO	$p = 3.23 \times 10^{-11}$
H_2O	$p = 8.67 \times 10^{-13}$
CH_4	$p = 2.16 \times 10^{-11}$
He	$p = 6.99 \times 10^{-11}$
H_2	$p = 4.90 \times 10^{-11}$
	$p_T = 2.0 \times 10^{-10}$ torr

This example has clearly been adjusted for illustration. The pressures seldom come out exactly in actual practice. Further, this represents a very simple case of residual gases in a UHV system. Most often, the spectra are far more complicated, especially if organics are involved. It does, however, illustrate the methodology for analyzing any spectra by the fragmentation method.

3.3.5

CALIBRATION OF VACUUM SYSTEM MASS SPECTROMETERS

In general, most mass spectrometer users are only concerned with the qualitative aspects of the vacuum environment and never calibrate the instrument, but there are many cases when it is absolutely necessary. Accurate partial pressure measurements for gas composition-sensitive instruments are controlled by the quality of their calibration. In other words, the measurement is only as good as the calibration. Considering a high-accuracy calibration, the subsequent demounting of the instrument, exposure to air, transfer from the calibration system to the operation system, remounting of the instrument, and finally, the bakeout and degassing required to reestablish calibration conditions, it is very difficult to maintain an actual sensitivity near the calibrated sensitivity. Ideally, the calibration system and the operational system should be one and the same. Unfortunately, this is seldom the case and the scientific worker must endure the aforementioned scenario primarily because a carefully calibrated instrument is absolutely necessary for good

research, development, or process reliability. Several techniques that can be used to calibrate vacuum instrumentation are presented next. Detailed descriptions of these methods are provided in Berman [31] and Ellefson and Cain [32].

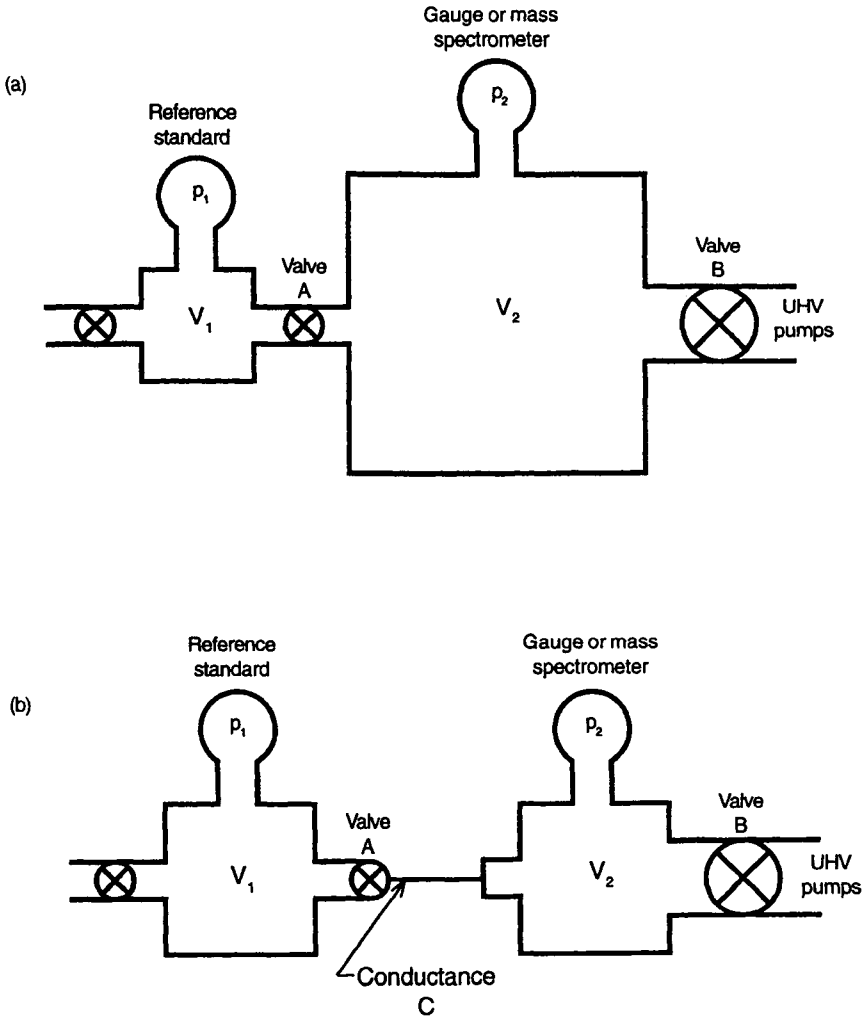
3.3.5.1 Static Volume Expansion

The static volume expansion simply involves establishing a pressure that is measurable by a primary standard within a very small volume and expanding that volume of gas into a much larger volume to obtain a substantial reduction in pressure (Boyle's law). Figure 11(a) is a schematic of the technique. The small volume, V_1, typically a few cm^3, and the large volume V_2, typically $> 10^3$ cm^3, are initially cleaned and evacuated to the UHV range to minimize outgassing contributions to the signal. The small volume is then pressurized to $p_1 = 10^{-2}$ or 10^{-1} torr with the separation valve closed. The small-volume pressure p_1 is measured with a primary standard or a working standard such as a fused silica Bourdon tube , a capacitance manometer, or a spinning rotor gauge and then the separation valve (which has a large-diameter bore) is fully opened. The equilibrium pressure then becomes

$$p_2 = \frac{V_1 p_1}{V_1 + V_2} \tag{24}$$

Vacuum chamber wall pumping effects such as chemisorption and/or incorporation can be serious problems and several runs must be made to saturate the walls and obtain repeatability. Furthermore, if hydrogen is the calibration gas, the walls must be significally oxidized to inhibit the seemingly endless hydrogen incorporation. This surface condition can be achieved with a bakeout of ~400°C in an atmosphere of oxygen for several hours. Thicker oxides, of course, can be obtained using longer times and higher temperatures. Another concern is gauge pumping. The ordinary ion source in a mass spec operating at an emission current of ~1 mA may have a pumping speed for N_2 approaching 0.1 L/s, which can have a substantial effect on the pressure diminution. Extrapolation techniques can be used to estimate the correct expanded pressure, but this is not a totally acceptable procedure. If wall effects and the gauge pumping are not serious problems, calibrations as low as 1×10^{-6} torr , depending on the gas, can be achieved. If the two volumes are initially degassed to a base pressure of 1×10^{-10} torr, wall outgassing in the static system in V_2 may increase the background pressure two–three decades over the time of the individual measurements. Thus, there is a 1% or greater uncertainty in the calibration gas. In practice, the low-pressure calibration limit is probably 1×10^{-5} torr. This technique is not generally used because of the aforementioned effects and because of its one-decade-or-less range overlap with ionization-type instruments.

Fig. 11.



(a) Schematic of volume expansion (Boyle's law) calibration technique. (b) Schematic of rate-of-rise calibration technique.

3.3.5.2 Pressure Rate-of-Rise Technique

A more desirable method of instrument calibration is a linear rate-of-rise of pressure in the measurement volume with which the instrument output is compared. The volume on the left (V_1) in Figure 11b and the volume on the right (V_2) are

typically small and are initially cleaned and evacuated to the UHV range to minimize the outgassing contribution to the signal. The volume, V_1, is then pressurized from 1 to 1000 torr, which is measured by a reference standard such as a quartz Bourdon tube, capacitance manometer, or a deadweight tester. The valve is then opened to the calibrated conductance C, which connects the two volumes. The resulting pressure rise in V_2, which has been closed off from its pumps, is then compared with the instrument output, so that the change in pressure is given by

$$V_2 \frac{dp_2}{dt} = C(p_1 - p_2) \tag{25}$$

Since $p_1 >> p_2$,

$$p_2(t) = \frac{p_1 C}{V_2} t \tag{26}$$

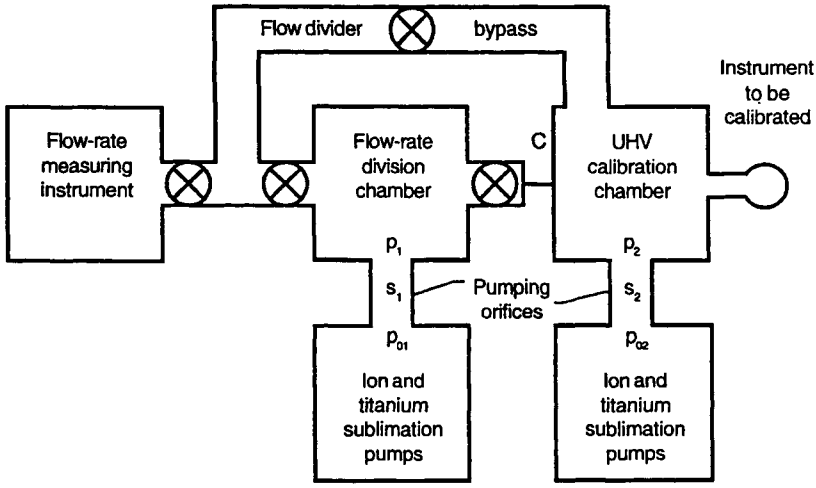
where $p_2(t)$ is the pressure in V_2 at any time t. The conductance, C, is usually selected to give a rate-of-rise that is fast enough to overwhelm gauge pumping, but slow enough so that instrument response is not a problem. Since V_2 is small and exposure to the calibration gas for several cycles tends to minimize wall chemisorption and incorporation, wall effects are not a serious problem. Since C is also small enough that p_1 remains constant, the flow in V_2 is assumed to be linear (after a few seconds). Furthermore, the low-pressure limit for this technique is probably better than an order of magnitude less than the static expansion technique, since only a short time is spent in the low-pressure range (e.g., $\sim 1 \times 10^{-7}$ torr). The rate-of-rise technique has been studied and has been found to be an acceptable calibration method.

3.3.5.3 Orifice Flow Technique

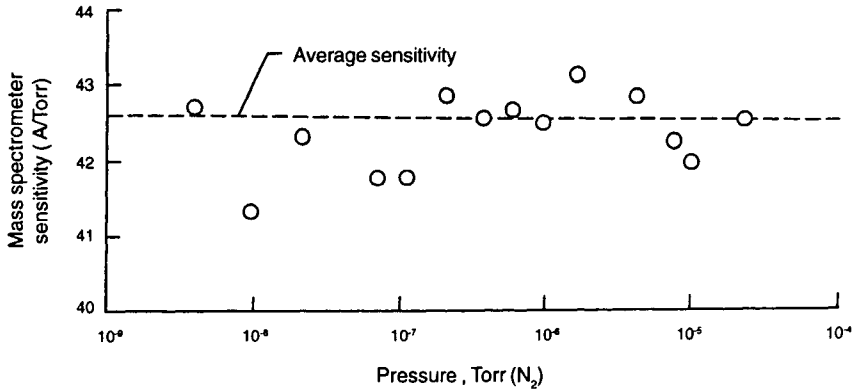
The orifice dynamic flow method is the primary technique used by the National Institute for Standards and Technology (NIST) for the calibration of ionization gauges and mass spectrometers, basically because it can extend the range of calibration several decades below the two methods previously discussed (i.e., from 5×10^{-5} to 5×10^{-10} torr). The orifice method is a dynamic technique that uses a flowmeter to provide a known flow Q_1 into the first of two volumes, which are pumped at known constant speeds (the two volumes have conductance-limiting orifices). The second volume, which is the second stage of pressure reduction, is where the instruments to be calibrated are located for the full range of calibration. Figure 12(a) is a diagram of the two-staged, ion-sublimation, turbomolecular pumped orifice calibration system. The entire system goes through a bakeout and is sufficiently degassed to achieve an ultimate pressure of $\sim 5 \times 10^{-12}$ torr or less

Fig. 12.

(a)



(b)



(a) Schematic of orifice calibration system for calibrations down to 1×10^{-10} torr (N_2). **(b)** Typical variations in mass spectrometer instrument sensitivity over the full calibration range using the orifice calibration method.

in both volumes, although this low background pressure is not absolutely necessary in the first volume. A flow Q_1 is admitted in V_1, and the steady-state pressure is

$$p_1 = \frac{Q_1 + S_1 p_{01}}{S_1} \sim \frac{Q_1}{S_1} \tag{27}$$

where S_1 is the conductance-limited orifice in V_2 or $S_1 << S_{p1}$ and p_{01} is the pressure in pump 1. When the valve is opened to V_2, the flow Q_2 into that volume is controlled by the calibrated conductance C and given by $Q_2 = C(p_1 - p_2)$. The steady-state pressure in V_2 is

$$p_2 = \frac{CQ_1 + CS_1 p_{01} + S_1 S_2 p_{02}}{S_1 S_2 + CS_1} \tag{28}$$

where S_2 is the conductance-limited orifice in V_2, or $S_2 << S_{p_2}$, and p_{02} is the pressure in pump 2 (S_{p_1} and S_{p_2} are the intrinsic pumping speeds of pumps 1 and 2). The magnitude of Q_1 could be a value that permitted a steady-state pressure of the calibration gas as low as 5×10^{-10} torr. Thus, the capability of a calibration from 5×10^{-5} to 5×10^{-10} torr or a range of 5 decades was provided. Once the calibration is complete, this method may also be used to measure outgassing from a particular volume where Q unknown is Q_1. The parameter of interest that characterizes the instrument over the calibration range is the sensitivity. The sensitivity of a mass spectrometer for N_2 as determined by an orifice system is shown in Figure 12(b). The relative variation of this instrument approaches 5%, but with some care it is possible to achieve 3%. In general, the conditions under which the calibrations are conducted should be nearly identical to those that the instrument would experience in the system where it is to be used. For example, it is recommended that mass spectrometers in the system have the same geometric metal (or insulator) envelope around the ion source of the instrument in the calibration system, because the electron density around the grid can be substantially different due to the alteration of potentials in the ion source, resulting in errors exceeding 50 percent.

The total uncertainty of the calibration is governed by several factors. First, there is the uncertainty in the standard input flow, which may be controlled to about 1 percent. Then there is an uncertainty in the conductances to the pumps in V_1 and V_2 as well as the conductances between the two volumes. There is the difference in gas composition in the pumps compared with the main volumes, but this should only be a minor concern. Perhaps, the two most significant issues are that no change occurs in the physical positioning of the instrument elements (which will substantially change the potentials) and that no chemical contamination occurs in the instrument transfer back to the operational system (which could increase ESD effects as well as alter the gain of the multiplier).

3.3.5.4 Transfer Standards

Considering the total number of mass spectrometer users, users seldom directly calibrate most instruments by sending the instrument transducer and electronics to an organization such as NIST. Users that do require some quantitative precision in their work most often employ a transfer standard, e.g., a spinning rotor gauge or ionization gauge that has been previously calibrated by one of the aforementioned methods. This instrument is then added to and remains as an integral part of the system. Periodically a gas (previously employed as calibration gas) is backfilled to the desired pressure range, and the mass spectrometer output is then compared to that pressure. The resulting sensitivity is thus determined as a mass spectrometer signal current out divided by the pressure as indicated by the calibrated SRG or ionization gauge, e.g., in (A/torr). The SRG calibration is good down to $p \sim 1 \times 10^{-6}$ torr and provides a two-decade overlap with the mass spectrometer range. The resulting sensitivity is then assumed to be a constant and independent of pressure for use in the lower pressure ranges. Direct comparison to a calibrated ionization gauge covers a much broader pressure range (1×10^{-4} to 1×10^{-11} torr), but it is less palatable as a transfer standard, because the gauge has the same problems as does the ion source in the mass spectrometer. Another benefit of transfer standards is that frequent checks on the mass spectrometer sensitivity can be conducted as desired without the problems of exposing the instrument to atmospheric pressure, physically moving the instrument, and the vagaries of subsequent bakeouts.

3.3.6 SOME APPLICATIONS

3.3.6.1 Industrial Applications

The following section represents examples of the types of information obtained from vacuum system mass spectrometers in industrial environments.

LEAK DETECTION

The basic information desired from a vacuum system mass spectrometer is, of course, very much dependent on the application. In general, all users first examine the spectrum to see if the vacuum system is leak tight and whether there is a normal spectrum. It, therefore, is initially a leak detector. If the spectra show a high atomic nitrogen fragmentation peak at $m/e = 14$ amu, a high H_2O peak at $m/e = 18$ amu, or a high molecular oxygen peak at $m/e = 32$ amu, then there is

most probably a leak. If the leak is too severe to continue, then the mass spec becomes a leak detector and normal leak detection procedures are implemented. If not, then the user proceeds to normal operation procedures. In the case of a leak, most often, the instrument is tuned to He (m/e = 4 amu) and the search to locate the leak initiated. The output of the He peak can be visually studied on an oscilloscope, strip chart recorder or on a computer by repetitive scans over m/e = 4 to indicate when the leak is found. If it is a very small leak, it may be sealed by low-vapor-pressure epoxies. If the leak partially or fully seals, the spectrum immediately shows a reduction in the aforementioned peaks.

VAPOR DEPOSITION

The acceptable pressure in which one can conduct thin film depositions without it being compromised by an unwanted residual gas contaminant, is a function of the application. The partial pressure of the contaminant species is directly related to the incident collision rate, ν, with the surfaces or volume of interest.

$$\nu = \frac{p}{(2pmk_BT)^{1/2}} \tag{29}$$

where p is the partial pressure, m is the mass, k_B is Boltzmann's constant, and T is the gas temperature. The well-known rule of thumb that a monolayer of surface contamination can be formed by an incident contaminant species (sticking coefficient of 1) in 1–2 seconds at approximately 1×10^{-6} torr illustrates the level of the problem.

EVAPORATION

In deposition by evaporation or physical vapor deposition, the source material is heated (usually by an electron bombardment) to a point where the resulting vapor pressure provides a sufficient deposition rate to coat the substrate. This is done in a vacuum environment that ordinarily has an ultimate pressure of $\sim 1 \times 10^{-9}$ torr and degrades to 1×10^{-6} torr during the deposition. If the partial pressure of the contaminant gas is too high (CO, for example), the purity, growth, and properties of the thin film may be severely altered by the dissociating carbon and oxygen. Basically, the C and O nucleate new phases that alter the crystallinity, electronic, mechanical, and chemical properties [33].

SPUTTER DEPOSITION

Sputter deposition is inherently a faster deposition rate than evaporation, but more of a contamination problem because the system is backfilled to pressures several orders of magnitude higher, $p \sim 1 \times 10^{-4}$ torr or greater, which gives a

proportionally higher contamination rate (often the vacuum system is not as good as those employed in an evaporation system). In addition, the purity of the backfill gas can create other contamination problems.

Mass spectra in these systems give an indication of what the species are as well as what their partial pressure is, therefore providing the corresponding contamination rate. This information, will ultimately provide some insight into the resulting film quality. Not all high background pressures are detrimental to the process, however. For example, one might have a high background of He (perhaps from permeation through elastomer O-rings or a glass envelope), which will play no part in contamination at all. It is therefore essential to know the gases that comprise the total background pressure.

In these last 2 deposition methods it is possible that inert gases such as He and Ar are physically trapped in the deposit. Although chemically the film is unaltered, it may have different physical characteristics due to the entrapment.

CHEMICAL VAPOR DEPOSITION

This application is primarily one of process control. The application can vary from atmospheric pressure down to the low-pressure regime and include plasma-enhanced methods. In general, the method requires that the mass spectrometer be differentially pumped through a capillary or an orifice [34] to allow normal operation of the instrument ($<1 \times 10^{-4}$ torr). The various gases then can be observed on the input and output sides of the reaction cell to control the ultimate deposition. Contaminants such as oxygen- and carbon-bearing gases are controlled to the ppm level and can be monitored in real time to allow maximum refinement of the process and therefore minimize contaminant-related effects.

3.3.6.2 Research Applications

As more sophisticated system requirements evolve and more demand is made on device performance, lower and lower pressures are desired. Today, research laboratories as well as many other scientific and industrial disciplines require ultrahigh vacuum ($p < 1 \times 10^{-9}$ torr). In the research discipline of surface science, even a background partial pressure of 1×10^{-10} torr may be too high. At this pressure, according to our rule of thumb, a monolayer of contamination will occur in 1.5 h. Some experiments can be seriously affected with less than 1% of a monolayer and may require experiment times longer than a 1.5 h to conduct. This would mean that even a partial pressure of 1×10^{-12} torr is not sufficient! Mass spectra at the operational pressure of the system are invaluable in assessing the potential contamination to a given experiment, thus providing an indication of whether the experiment can proceed and what effects it may have had on the re-

sults of that work. In research systems (where there are ion guns, electron guns, X-ray sources, cylindrical and hemispherical analyzers, and a variety of other instruments and hardware), the resulting partial pressures are most certainly determined by the different construction materials contained therein, their total area, temperature, and history as well as numerous other factors that control the outgassing. Mass spectra provide revealing information as to the relative state of the system cleanliness. A figure of merit often quoted in research papers is the ultimate pressure and residual gas species in the experimental system employed, because, in part, it basically reflects the quality of the work.

Mass spectrometers are used in a variety of different ways for research studies. Some examples are as follows:

TEMPERATURE DESORPTION SPECTROSCOPY (TDS)

The mass spectrometer looks at the desorption of gases adsorbed on surfaces that have been rapidly heated (2–10°C/s) with a specific temperature ramp (usually linear). The gas desorbs directly into the ion source and is subsequently analyzed. The analysis of the rate of gas evolution allows activation energies for the process to be determined and give some insight into the adsorption state proceeding the desorption [35].

SECONDARY ION MASS SPECTROSCOPY (SIMS)

The mass spectrometer analyzes ions (positive and negative) and neutrals that are sputtered by an incident ion beam of some inert gas (e.g., He, Ar) or some reactive ion (e.g., O_2, Cs). The sputtered ions are energy filtered to narrow the energy spread and the ions analyzed with a mass filter (e.g., quadrupole) [36].

CHEMICAL KINETICS ANALYZER

The mass spectrometer ion sources that are pulsed ~10k Hz and the ions pulsed into a long time-of-flight (TOF) drift tube allow kinetic changes of various reactions (conducted in a reaction chamber) to be analyzed for ms chemical changes. The arrangement of the reaction cell adjacent to the ion source where the ions are pulsed into the long flight tube allows high-resolution TOF analysis and permits chemical changes to be monitored at the rate of better than 1 full spectrum per 1 ms time intervals.

TIME-OF-FLIGHT (TOF) ENERGY ANALYZER

The neutral atomic and molecular kinetic energies can be analyzed by TOF methods where the neutrals are ionized by electron impact perpendicular to the axis of

the drift tube and then chopped or gated into the drift tube. The measure of the transient time then gives the kinetic energy [37].

GAS CHROMATOGRAPH–MASS SPECTROMETRY (GCMS)

The gas chromatograph columns initially separate different specific gas species by using selective adsorbents in a column the egress of which ultimately is measured by a mass spectrometer with much higher accuracy and sensitivity [38].

REFERENCES

1. F. A. White, *Mass Spectrometry in Science and Technology* (Wiley, NY, 1968); F. W. Aston, *Mass Spectra and Isotopes* (Longmans, Green, New York, 1942); R. W. Kiser, *Introduction to Mass Spectrometry and Its Applications* (Prentice-Hall, Englewood Cliffs, NJ, 1965).
2. P. H. Dawson, *Quadrupole Mass Spectrometry and Its Applications* (Elsevier, Amsterdam, 1976/American Institute of Physics Press, New York, 1995).
3. M. J. Drinkwine and D. Lichtman, *Partial Pressure Analysis*, American Vacuum Society Series (American Institute of Physics Press, New York, 1990).
4. H. Semat, *Introduction to Atomic and Nuclear Physics* (Rinehart, New York, 1958).
5. A. O. C. Nier, *Advances in Mass Spec.*, **6** (1981) 212.
6. D. G. Torr, *Photochemistry of Atmospheres*, edited by J. S. Levine (Academic Press, Orlando, FL, 1985).
7. G. F. Weston, *Vacuum*, **25** (1975) 469.
8. W. G. Perkins, *J. Vac. Sci. Technol.*, **10** (1973) 543.
9. R. A. Outlaw and M. R. Davidson, *J. Vac. Sci. Technol.*, **A12** (1994) 854.
10. *Handbook of Chemistry and Physics*, 56th ed. (CRC Press, Cleveland, 1975).
11. L. Lieszkovsky, *Handbook of Vacuum Technology* (Academic Press, Boston, 1997), Chapter 3.2.
12. *Mass Spectral Data*, American Petroleum Institute Research Project #44 (Carnegie Mellon University, Pittsburgh, PA, 1962). See also National Institute of Standards and Technology, Standard Reference Data Base, Gaithersburg, MD, for current mass spectral library of compounds.
13. S. Rezaie-Serej and R. A. Outlaw, *J. Vac. Sci. Technol.*, **A12** (1994) 2814.
14. P. A. Redhead, J. P. Hobson and E. V. Kornelsen, *The Physical Basis of Ultrahigh Vacuum* (Chapman and Hall, London, 1968/American Institute of Physics Press, New York, 1993).
15. W. H. Kohl, *Handbook of Materials and Techniques for Vacuum Devices* (Reinhold, New York, 1967/American Institute of Physics Press, 1995).
16. P. A. Redhead, *Can. J. Phys.*, **42** (1964) 886.
17. D. Menzel and R. Gomer, *J. Chem. Phys.*, **41** (1964) 3311.
18. M. L. Knotek and P. J. Feibleman, *Surf. Sci.*, **90** (1979) 78.
19. P. R. Antoniewicz, *Phys. Rev.*, **B21** (1980) 3811.
20. T. E. Madey and J. T. Yates, Jr., *J. Vac. Sci. Technol.*, **8** (1971) 525.
21. M. J. Drinkwine and D. Litchman, *Prog. Surf. Sci.*, **8** (1970) 78.
22. H. A. Engelhardt and D. Menzel, *Surf. Sci.*, **57** (1976) 591.
23. A. von Engel, *Ionized Gases* (Oxford University Press, London, 1965/American Institute of Physics Press, New York, 1994).
24. J. B. Hasted, *Physics of Atomic Collisions*, 2nd ed. (Butterworths, London, 1972).

25. T. A. Flaim and P. D. Ownby, *J. Vac. Sci. Technol.*, **8** (1971) 661.
26. J. F. Hennequin and R. L. Inglebert, *Intl. J. Mass Spec. Ion Phys.*, **26** (1978) 131.
27. R. Young, *Rev. Sci. Instrum.*, **37** (1966) 1414.
28. L. C. Beavis, *Partial Pressure Analyzers*, American Vacuum Society Education Committee, New York, 1996.
29. F. W. McLafferty, *Interpretation of Mass Spectra* (Benjamin, New York, 1967).
30. *American Institute of Physics Handbook* (McGraw-Hill, New York, 1972).
31. A. Berman, *Total Pressure Measurements in Vacuum Technology* (Academic Press, New York, 1985).
32. R. E. Ellefson and D. Cain, *J. Vac. Sci. Technol.*, **A5** (1987) 134.
33. P. L. Swart and H. Aharoni, *J. Vac. Sci. Technol.*, **A3** (1985) 1935.
34. N. Muller, G. Rettinghaus, and G. Strasser; *J. Vac. Sci. Technol.*, **A8** (1990) 2822.
35. J. B. Hudson, *Surface Science: An Introduction* (Butterworth-Heinemann, Boston, 1992), p.292; R. J. Madix, *Surf. Sci.*, **299/300** (1994) 785.
36. A. Benninghoven, *Surf. Sci.*, **299/300** (1994) 246.
37. G. Samorjai, *Surf. Sci.*, **299/300** (1994) 849.
38. I. Morisako, I. Kato, and Y. Ino, *Intl. J. Mass Spec. Ion Phys.*, **48** (1983) 19.

CHAPTER 3.4

Mass Flow Measurement and Control

L. D. Hinkle
MKS Instruments, Inc.

3.4.1 GENERAL PRINCIPLES OF MASS FLOW MEASUREMENT

3.4.1.1 Units of Measure

The measurement of the mass of gaseous material as it enters a vacuum system is critical to much of the work related to vacuum technology. Analogous with the scale on which substances are weighed in a chemistry lab, the mass flow meter provides the ability to control the mixture of gaseous substances as they enter the process chamber. The necessity for a *mass* flow meter is not universal, however, and other techniques for measuring the velocity or volumetric flow of a gas are also quite common and useful in certain applications.

The true *units* of mass flow are not often used in vacuum practice, probably due to the difficulty in picturing, for example, a kilogram of gas going by in a second. On the other hand, volumetric flow units are intuitively simple — liters per minute or cubic centimeters per minute. The problem with volumetric flow is that it usually does not provide enough information. Since gases are compressible, the amount of material in a certain volume depends on the pressure and temperature. So without specifying temperature and pressure, the units of volumetric flow cannot be used effectively for measuring gas flows.

ISBN 0-12-325065-7
$25.00

What has evolved over time is a system of practical units based on volumetric flow relative to a standard condition temperature and pressure, 0°C and 1 atm respectively. The most commonly used forms are sccm (standard cubic centimeters per minute) and slm (standard liters per minute). This convention assumes the ideal gas law and uses the fact that one mole of an ideal gas occupies 22.4 liters of volume. Therefore, standard volumetric flow relates directly to molar flow and then to mass flow by the molar mass. For example,

500 sccm of nitrogen
= 0.5 slm of nitrogen
= 0.5 slm/22.4 standard liters per mole = 0.0223 moles per minute of nitrogen
= 0.0223 moles per min × 28 grams per mole of nitrogen = 0.624 grams per minute.

Other standard volumetric units (with an implied reference temperature of 0°C) can be applied with the proper conversions as shown here:

1 mbar liter/second = 59.2 sccm
1 Pa m^3/second = 592 sccm

3.4.1.2 Methods of Measure

For continuous mass flow measurement of gases into vacuum systems, two general approaches can be commonly found in modern instrumentation: pressure based and thermal based. Of these, the majority of instruments are thermal based. Flow measurements may be made by measuring pressures upstream and downstream of a laminar flow element, an orifice plate, a Venturi element, or a molecular flow element. Also, simple mechanical devices known as *rotameters* (variable area, constant differential pressure) can serve as convenient volumetric flow indicators. As mentioned in the section of this handbook that discusses gas flow equations, each of these embodiments must also involve temperature, gas properties, and geometrical considerations to properly determine mass flow. The use of pressure-based flow instrumentation has typically been in areas where some aspect of a thermal-based device, such as temperature or required pressure drop, conflicts with the application.

The principle used in thermal mass flow sensors is that the thermal transfer from upstream to downstream relates to the flow of thermal mass, which in turn relates to the "true" mass flow by the *specific heat* c_p of the gas. For monatomic and diatomic gases, the specific heat is relatively constant as a function of pressure and temperature; the relationship between thermal transfer and mass flow is nearly independent of the gas pressure and temperature. For more complex molecular gases, the dependence of c_p on temperature and pressure may have an ef-

fect on the linearity and pressure sensitivity of the measurement. Note that there is an inherent gas type dependence in that each gas has its own c_p value.

3.4.2

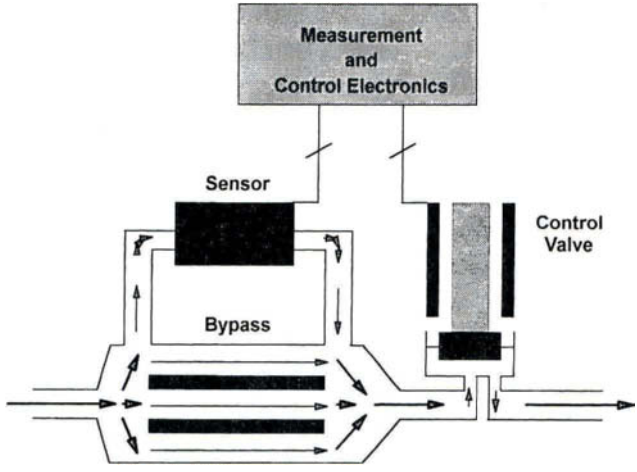
OVERVIEW OF THERMAL MASS FLOW CONTROLLER TECHNOLOGY

A thermal mass flow meter (MFM) is often combined with a control valve and electronics to form an instrument known as a mass flow controller (MFC). A representation of these components and their typical arrangement in an MFC is shown in Figure 1.

3.4.2.1 Sensing Techniques

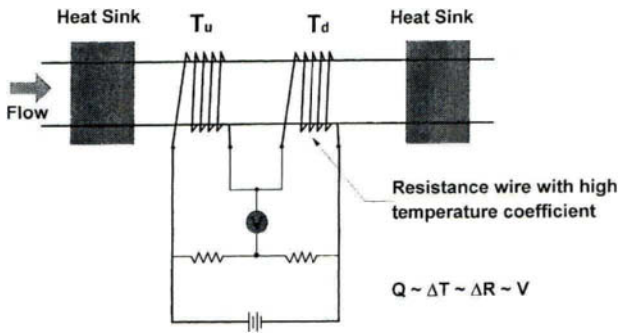
The construction of the thermal flow sensor consists of a stainless steel capillary tube with heaters and temperature sensors arranged symmetrically on the outer surface. The ends of the tube are clamped, both mechanically and thermally, to the body of the instrument such that the thermal environment is symmetric around the center of the tube. Therefore, any imbalance detected by the temperature sensors is interpreted as a mass flow within the tube, the gas absorbing heat from the upstream region and transferring heat to the downstream as it flows. A typical sensor is designed with two matched windings of high-temperature coefficient resistance wire connected in a Wheatstone bridge circuit as shown in Fig-

Fig. 1.



Thermal MFC overview.

Fig. 2.



Thermal flow sensor technique based on differential temperature.

ure 2. In this arrangement, each winding performs the heating and temperature-sensing functions. The resulting bridge voltage represents the windings' resistance difference and thus the temperature difference. Other methods of excitation and sensing have been developed based on a "constant temperature, differential voltage" approach or variations of these techniques.

The electronics that detect and condition the output of the sensor bridge must amplify, linearize, and "speed up" the signal. It should be noted that this conditioning results in a flow signal output in the case of an MFM; in an MFC, it also serves as an input to a controller circuit (discussed in Section 3.4.2.4). A typical sensor output exhibits 3–10% of full-scale nonlinearity and a time constant of between 0.5 and 5 seconds. Various standard methods of linearization are used to adjust the instrument output to within the typical 0.5% of full-scale linearity specification, while an anticipation circuit, based on an active differentiator, is effective in making the output indicate the actual flow in real time.

3.4.2.2 Bypass Ranging

Most thermal flow sensors have a maximum flow limit of between 5 and 20 sccm, above which they become too nonlinear to correct for in the electronics. Since many applications for flow measurement require higher rates, a laminar bypass is placed in parallel with the sensor such that a fixed ratio of gas is diverted away from the sensor. This is analogous to a shunt resistor used to select the range of an ammeter. The bypass technique allows measurement range selection from 10 sccm to greater than 100 slm—all with one sensor. The standard ranges divide each decade into three fairly equal regions (on a logarithmic scale) 10, 20, 50, 100, 200, etc. Calibration of the meter must occur after the bypass has been installed to account for the slight variations in the conductance from bypass to bypass and sensor tube to sensor tube. Recall from Chapter 1.1 that the conductance for a

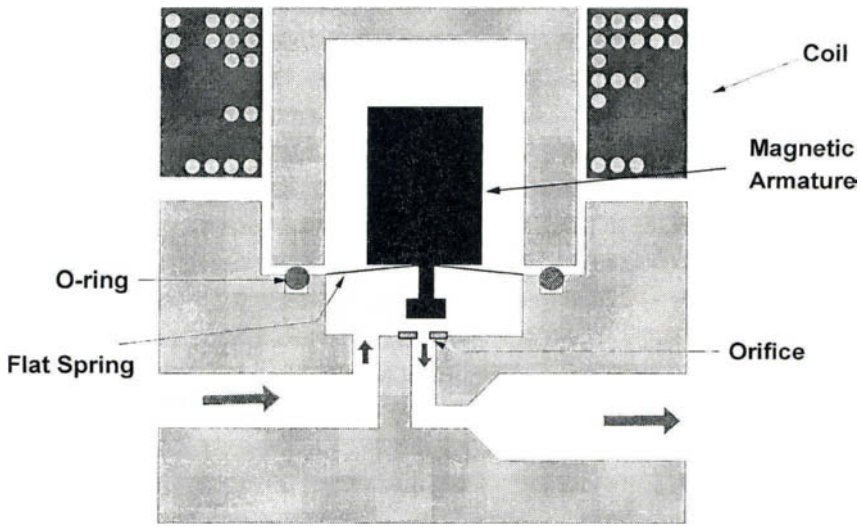
tube in laminar flow is related to the *fourth power* of diameter. For the level of flow accuracy desired for these instruments, better than 1%, the mechanical tolerances are effectively unachievable unless calibration is performed after assembly of the individual sensor and bypass combination.

3.4.2.3 Control Valve Design

The sensor, associated electronics, and bypass (where one is needed) form a mass flow meter. To make the instrument a mass flow controller, a control valve is placed downstream from the meter and a control circuit is added to the electronics. Control valves used in typical MFCs allow continuous adjustment of the conductance as opposed to the on/off action of a shutoff valve. The resolution of conductance is much better than 0.1% of the open conductance to obtain the desired control range of the MFC. Generally, the control valve is not capable of providing positive shutoff, and a separate valve intended for that purpose should be used. The valve's open conductance can be set by selecting the orifice diameter and, as with the bypass, is determined by the full-scale range of the MFC. Two varieties of control valves are in common use: solenoid actuated and piezoelectric actuated.

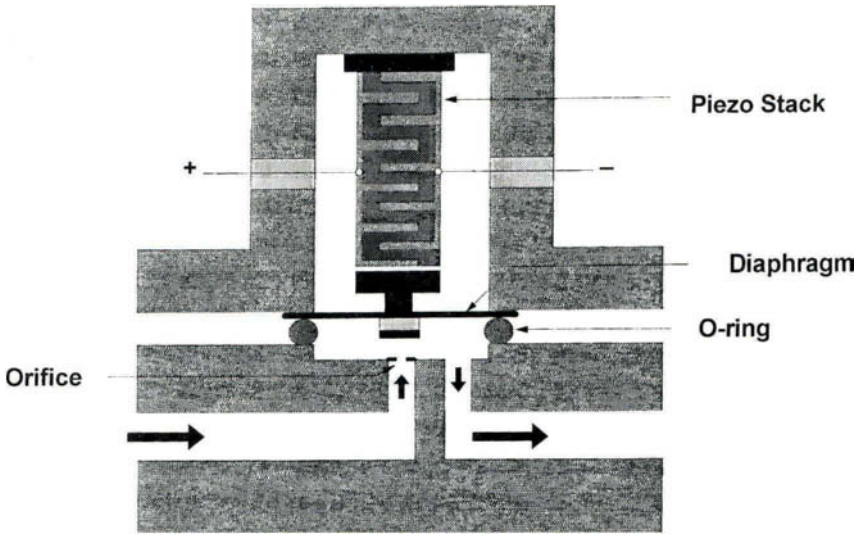
Solenoid valve designs are more traditional and more commonly used in MFCs. An overview of the typical construction of a solenoid valve is shown in Figure 3. The operating principle relies on current in the solenoid to generate a magnetic field that acts on the armature to open the valve. A flat spring is normally under

Fig. 3.



Solenoid valve overview.

Fig. 4.



Piezo valve overview.

tension (preload) to oppose the force provided by the solenoid actuator. This configuration results in a *normally closed* valve whereby if power is removed the device will assume a closed condition. *Normally open* valves can also be designed with solenoid actuation, but they tend to be more complicated and are not nearly as common.

Flow controllers with piezo valves were developed in the mid-1980s and have been used primarily in applications where high purity and corrosion resistance are primary concerns. The construction features of the piezo valve are represented in Figure 4. The piezo actuator is a stack of 50–100 individual disks made from a special ceramic material that exhibits the piezoelectric effect where an electric field generated across it causes the material to expand along that axis. Interdigitated electrodes generate the electric field when voltage is applied to the actuator terminals. This causes each disk to expand, and the combined effect acts to close the valve. Thus, piezo valves tend to be designed to operate as normally open, although normally closed versions are also possible. Also, their capability for generating high force allows piezo actuators to be used in diaphragm isolation valve designs as shown in Figure 4.

3.4.2.4 Control System Design

The control circuit is based on PID (proportional integral derivative) feedback on the *controllor error* signal generated by the flow sensor versus the set point signal

supplied to the MFC. Note that the controllor error is simply the instantaneous difference between the flow signal and the set point; it should not be confused with the overall instrument error, which is represented by the difference between the flow signal and the actual flow. The control action is tuned to drive the valve so that the controller error is reduced to zero as rapidly as possible with little or no overshoot. Tuning involves changing parameters governing the action of the PID gains to achieve the desired response to set point changes. The tuning can be optimized for a given set of conditions such as gas type, upstream pressure, downstream pressure, physical orientation, and set point; but if the conditions change significantly, controller performance may be affected. Symptoms are typically slow response to set point changes or control oscillation.

3.4.3 PERFORMANCE CHARACTERISTICS

The proper installation procedure, gas line components, and ancillary electronic instrumentation help ensure optimal performance and reliability from an MFC. Environmental conditions such as temperature, orientation, vibration, and electromagnetic interference may also affect an MFC's ability to measure and control flow.

3.4.3.1 Installation

For all gas line components including the MFC, the proper fittings and mechanical installation are necessary to provide a leak-tight system. The fittings on most MFCs are either face seal (VCR®, registered TM of Cajon Co., Macedonia, OH) or ferrule seal (Swagelock®, registered TM of Crawford Fitting Co., Solon, OH) for $\frac{1}{4}$-inch pipe. However, if the full-scale flow is greater than 20 slm, the fittings may be for $\frac{1}{2}$-inch pipe. For vacuum system or high-purity applications, the MFC fittings should be helium leak tested if possible. This is extremely important in the case of toxic or corrosive gas service. In addition to ensuring the leak integrity of the system, proper purging to reduce atmospheric contamination (moisture and oxygen) to acceptable trace levels is crucial for reliable performance. The specific method to achieve this condition depends on the system and level of contamination desired. At a particular facility, standard purging procedures should be established for each case. Likewise, procedures for removing the toxic or corrosive gas before removing an MFC (or any other gas line component) must be established and followed.

After the recommended warmup period, a check of the indicated zero is advisable. The user's manual for the particular instrument should be consulted regarding the proper procedure, but in general, the following guidelines may be used:

- Ensure there is no flow by closing a separate shutoff valve in line with the MFC.
- Command a set point flow to the MFC. (This equalizes the pressure upstream to downstream.)
- After the indicated flow has stabilized, command a set point flow of zero.
- Adjust the potentiometer, usually accessible from the outside of the instrument, labeled "zero."
- Alternatively, the adjustment may be performed at the readout/set point device.

This process avoids the problem of "zeroing in" a small leak, which may normally occur through the MFC control valve while a differential pressure remains across the MFC during the zero adjustment.

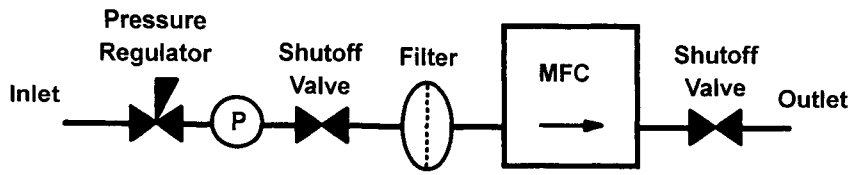
3.4.3.2 Gas Line Components

In line with the MFC is often a combination of other instruments and components including a pressure regulator, a particle filter, shutoff valves, and a pressure transducer. In some industries, such as semiconductor manufacturing, this is called a *gas stick*. A typical gas stick layout is shown in Figure 5. The pressure regulator and filter can affect MFC performance if a fluctuation in pressure occurs or if the differential pressure becomes too low or too high.

3.4.3.3 Electrical Requirements

The electrical connector on an MFC is usually either a "card edge," 15-pin D, or 9-pin D. The D connectors provide a degree of electromagnetic shielding when

Fig. 5.



Typical gas stick.

used with a shielded cable. Unfortunately, there is no standard for the electrical connections (pinouts) between manufacturers, so caution must be used to ensure that proper electrical compatibility is maintained among power supply, cable, and MFC. The industry standard power supply for MFCs is +/−15 V DC with a typical current draw of 100 mA (200 mA maximum at startup). The set point voltage and indicated flow are each 0 to 5 V, but it should be noted that most instruments are capable of indicating some amount of reverse flow (negative voltage) and controlling to well above full-scale flow (greater than 5 V).

Other functions that may be available as signals at the connector include commands for overriding the internal valve control to drive it completely open or closed; a voltage for monitoring the internal valve drive signal (an important diagnostic); an optional control signal input; and a 5 V reference.

There are usually three separate common lines with separate functions. The *signal common* is used as a 0 V reference for the set point and indicated flow. The *power common* is used to return current imbalances in the power supply draw. A *chassis ground* is intended to provide a connection to the body and case of the instrument. Although some designs are more tolerant of combining the signal common and power common, it is recommended to keep these three connections separate in any precision instrumentation such as an MFC.

3.4.3.4 Temperature Effects

The ways in which the environmental temperature may affect the behavior or performance of an MFC include shifts in the indicated zero, shifts in the indicated span or sensitivity, reduced performance of the control capability, or inability to reach set point due to valve conductance loss.

The most easily recognized problem due to temperature effects is a zero shift expressed quantitatively as a zero temperature coefficient (zero TC). The specification for zero TC is usually less than 0.05 % FS/°C. So for example, if the temperature around an MFC changed by 10°C (18°F) the zero could change by as much as 0.5 % FS. This can be compensated for by re-zeroing the device at the operating temperature following the guidelines in Section 3.4.3.1.

The shift in span or sensitivity as a function of temperature is more difficult to recognize and account for in a typical application. For gases that do not exhibit a large temperature dependence of c_p (see Section 3.4.1.2) the span TC is typically less than 0.1% of Rdg/°C. However, gases having a temperature sensitive c_p will exhibit additional span TC behavior.

The operating environmental temperature range for MFCs is usually rated at 0 to 50°C.

3.4.3.5 Orientation (attitude) Effects

Since flow measurement and control instrumentation must be mounted in a variety of locations and orientations relative to gravity, the effect known as "thermal siphoning" may be exhibited in MFCs. This effect occurs when the heated sensor is oriented other than horizontally and the buoyancy difference between the gas in the sensor and bypass causes a circulating flow within the MFC. Although the effect is present anytime such orientation exists, it is most noticeable when a particularly dense gas is used (e.g., SF_6, Kr, or the various HFCs and CFCs and related compounds). The effect is also line pressure sensitive, with higher pressures resulting in higher circulating flows.

For small offsets, the thermal siphoning influence exists independent of the flow through the device and therefore acts as a fixed offset to all flow measurements. In this case (less than 5% FS shift), the device may be re-zeroed at line pressure by following the guidelines in Section 3.4.3.1 and used normally.

3.4.3.6 Gas Type Effects

The type of gas flowing through an MFC has a direct effect on the calibration of the measurement and the control performance.

A gas correction factor (GCF) may be applied to the output signal to convert the sensitivity from the calibration gas to another gas. Recommended GCF values are listed in most user's manuals and may vary from one manufacturer to another. Caution must be exercised when using GCFs as the accuracy of the standard gas correction method is not better than 1% for many gases and for certain situations may introduce as much as 10% error. The accuracy of a particular gas correction depends on several factors, including the two gases involved, the instrument model, and the range (bypass or no bypass). For a particular case, specific recommendations may be available from the instrument manufacturer.

Since the issue of gas correction is, practically speaking, very complex, an ideal approach is to perform a calibration or verification with the actual gas in the actual environment if possible. For some applications, *in situ* or point-of-use calibration methods have been developed based on pressure rate-of-rise (ROR) procedures to address the issue. With these methods, though, care must be exercised to properly account for such influences as temperature, surface conditions, and gas compressibility and condensation behavior.

The technique of surrogate gas calibration, where a gas with similar molecular properties is used for the calibration of an instrument for hazardous gas use, has received some acceptance. However, there is very little published experimental evidence that correlates the performance of surrogate gases to actual gases.

3.4.4 TROUBLESHOOTING

Since an MFC has a set point signal, an output signal, and usually an internal valve drive signal available at the electrical connector, some troubleshooting is possible without removing the instrument or otherwise opening the gas supply line. *In situ* trouble shooting can potentially avoid time-consuming and costly replacement procedures.

3.4.4.1 Initial Checks

With an appropriate breakout connector and a voltmeter, prepare to measure the various signals available on the device under test. Measure with respect to the signal common through a high-input impedance meter.

1. Verify that the power supply and grounding lines are within the specified tolerances.
2. Verify that the output signal is being displayed properly by the readout instrument. Check for:
   a. Proper settings for range and gas correction factor
   b. Unintentional zero offsets at the display
   c. Display voltage calibration
   d. Cable losses

3.4.4.2 Operational Diagnostics and Troubleshooting

Ideally, the device under test should have been characterized while working properly as part of a routine "fingerprinting" exercise for the entire system. A record of the typical level and variability of each signal becomes extremely useful when troubleshooting or predicting a potential failure. In addition to measuring the signals available at the MFC, it is important to have an independent indication that flow is occurring. This need not be an accurate measurement; only an indication that there is or is not flow. Often a pressure gauge on the vacuum system into which the flow is directed can serve as the "actual flow" indication.

The troubleshooting chart in Table 1 and symptom explanation serves as a guide to possible problems encountered in a gas delivery system. It is worth noting that some of the potential problems can be remedied without removing components or otherwise exposing the system to atmospheric conditions.

Table I

MFC Troubleshooting Chart

Set Point (V)	Actual Flow	Sensor Output (V)	Valve Indication	Problem
0	Large	Large	Closed	Valve won't close
0	Large	Large	Open	Control loop problem
5	Large	0	Open	Meter problem (false low)
5	0	Large	Closed	Meter problem (false high)
5	0	0	Open	Valve won't open or no Δp
5	Some	<5 V	Open	Valve slightly clogged or low Δp
5	Some	OK but oscillating	Oscillating	Control loop not stable

Valve Won't Close:
Mechanical
  Actuator failure
  Out of mechanical adjustment
  p too large
Electrical
  Valve override applied
  Coil failure (N/O)
  Connector failure (N/O)

Valve Won't Open:
Mechanical
  Actuator failure
  p too large
  Clogged (valve or sensor)
  Swelled elastomer plug
Electrical
  Valve override applied
  Coil failure (N/C)
  Connector failure (N/C)

Meter Problem:
False low
  Sensor failure (electrical)
  Clogged sensor (in bypassed ranges)
False high
  Sensor failure (electrical)
  Clogged bypass (large ranges)
  Thermal syphoning

Control Loop Problem:
► Latch up
► Circuit failure

Control Loop Not Stable:
► Valve adjustment
  Orifice size
  Plug to seat distance
► Electronic adjustment
  Control gain
  Control phase
► Gas type
► Line pressure
► Pressure fluctuations upstream
► Pressure fluctuations downstream

REFERENCES

General and Historical
Benson, J. M., and W. C. Baker, *Instrum. Control Syst.*, **43** (1970) 85.
Brown, A. F., and H. Kronberger, *J. Sci. Instrum.*, **24** (1947) 151.
Burggraaf, Pieter, *Semiconductor International*, **14**(3) (1991) 60.
Hawk, C. E., and W. C. Baker, *J. Vac. Sci. Technol.*, **6** (1969) 225.
Miller, R. W., *Flow Measurement Engineering Handbook* (McGraw-Hill, New York, 1983).

Olin, J. G., *Intech*, **40** (1993) 8.
Sullivan, J. J., S. Schaffer, and R. P. Jacobs, Jr., *J. Vac. Sci. Technol.*, **A7** (1989) 2387.
Tison, S. A., *J. Vac. Sci. Technol.*, **A14** (1996) 2582.

Calibration
Clark, R., *MKS Applications Engineering Note*, **5** (1989).
Le May, Dan, and David Sheriff, *Solid State Technology*, **39**(11) (1996) 83.
Mattingly, G. E., Gas flow metrology, *NCSL Newsletter*, **29**(1) (January 1989).
Scott, R. W. W., *Developments in Flow Measurement-1* (Applied Science Publishers, Englewood, NJ, 1983), chap. 2.
Widmer, A. E., R. Fehlmann, and K. Ohtani, *J. Phys E.*, **15** (1982) 213.

Sensor Technology
Benson, J. M., W. C. Baker, and E. Easter, *Instrum. Control Syst.*, **43** (1970) 85.
Hinkle, L. D., and C. F. Mariano, *J. Vac. Sci. Technol.*, **A9** (1991) 2043.
Komiya, K., F. Higuchi, and K. Ohtani, *Rev. Sci. Instrum.*, **59** (1988) 477.
Segal, R., E. M. Sparrow, and T. M. Hallman, *Appl. Sci. Res.*, **7** (1958) 386.
Singh, S. N., *Appl. Sci. Res.*, **A7** (1958) 325.

Gas Correction, Time Response, Diagnostics
Dille, J. C., *Semiconductor International* May 1988) 234.
Gray, D. E., N. M. P. Benjamin, and B. N. Chapman, *Microelectronics Manufacturing Technology* (August 1991), p. 35.
Hinkle, L. D., *Semiconductor International*, **14**(3) (1991) 68.

PART 4

Systems Design and Components

4.1 Selection Considerations for Vacuum Valves 391
- 4.1.1 Introduction 391
- 4.1.2 Valves for Shutoff 391
- 4.1.3 Valves for Control 397
- 4.1.4 Valve Construction 398
- 4.1.5 Specialty Valves 404
- 4.1.6 Installation Considerations for Vacuum Valves 407

References 408

4.2 Flange and Component Systems 409
- 4.2.1 Introduction 409
- 4.2.2 Selecting a Flange System 410
- 4.2.3 Common Flange Systems 410
- 4.2.4 Components with Flanges Attached 425

Trademarks 430

References 432

4.3 Magnetic-Fluid-Sealed Rotary Motion Feedthroughs 433
- 4.3.1 Basic Sealing Principle 433
- 4.3.2 Application Factors 434
- 4.3.3 Impact of Feedthrough on Process 436
- 4.3.4 Impact of Process on Feedthrough 437
- 4.3.5 Materials Considerations 438
- 4.3.6 Application Examples 440
- 4.3.7 Comparison to Other Types of Feedthroughs 442

4.4 Viewports 444
- 4.4.1 Materials 444
- 4.4.2 Mounting Systems and Precautions 445

4.5 Construction Materials 446
- 4.5.1 Properties Defining Material Performance 446
- 4.5.2 Vacuum Chamber Materials 451
- 4.5.3 Special-Purpose Materials 455
- 4.5.4 References 462

4.6 Demountable Seals for Flanges and Valves 463
4.6.1 Sealing Overview: Polymer and Metal Seals 463
4.6.2 The Elastomeric and Nonelastomeric Polymers Used in Vacuum Sealing 464
4.6.3 Metal Seals 474
References 482
4.7 Outgassing of Materials 484
4.7.1 Relationships Among System Pressure, Pumping Speed, and Outgassing 484
4.7.2 Initial Pumpdown from Atmospheric Pressure 494
4.7.3 Pressure vs. Time During Outgassing 495
4.7.4 The Outgassing Rate of Elastomers and Plastics 497
4.7.5 The Outgassing Rate of Metals and Ceramics 501
4.7.6 The Outgassing Rate of Preconditioned Vacuum Systems After Short Exposure to the Atmosphere 504
4.7.7 Methods of Decreasing the Outgassing Rate 506
4.7.8 Measurement of the Outgassing Rate of Materials 507
References 508
4.8 Aluminum-Based Vacuum Systems 509
4.8.1 Outgassing 509
4.8.2 Demountable Seals 512
4.8.3 Cleaning and Surface Finishing 518
4.8.4 Mechanical Considerations 520
4.8.5 Thermal Conductivity and Emissivity 536
4.8.6 Corrosion 538
4.8.7 Welding Aluminum for Vacuum Applications 541
References 548
4.9 Preparation and Cleaning of Vacuum Surfaces 553
4.9.1 Surface Modification 554
4.9.2 External Cleaning 567
4.9.3 Assembly, Handling, and Storage 587
4.9.4 *In Situ* Cleaning 591
4.9.5 Documentation 599
4.9.6 Conclusion 601
Trade Names 601
References 601

CHAPTER 4.1

Selection Considerations for Vacuum Valves

Neil T. Peacock
HPS Division of MKS Instruments, Inc.

4.1.1 INTRODUCTION

Most vacuum equipment uses valving for pump isolation, gas flow control, or gas admission. When selecting a vacuum valve, there are many considerations. Some of these are whether it is for shutoff or for control, the type of valve, the pressure range over which the valve is to be used, how it is to be actuated, and the seal materials. Compatibility of construction materials exposed to the process or the ability of the operating mechanism to tolerate opening against a pressure differential may also be important.

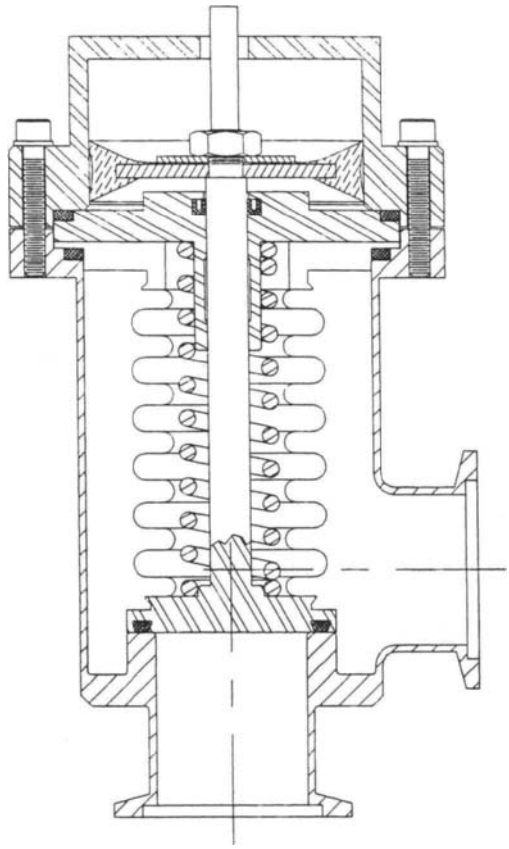
4.1.2 VALVES FOR SHUTOFF

Valves intended to isolate sections of a vacuum system completely seal off a port of the valve, and generally are intended to be used only open or closed. In most

ISBN 0-12-325065-7
$25.00

Fig. 1.

Pneumatically actuated right-angle poppet valve.



This valve has polymer nosepiece and bonnet seals, and uses a bellows for the transfer of actuation motion into the body. The spring within the bellows, concentric with the stem, closes the valve when air pressure is removed from the pneumatic cylinder. (Courtesy of the HPS Division of MKS Instruments, Inc., Boulder, Colorado.)

valves a movable plate containing a seal element is forced against a surface in the valve body to shut off the gas flow. The seal member can be either a polymer or a metal. Two common valve types are the poppet valve and the gate valve. A poppet valve is shown schematically in Figure 1. In this valve, the seal plate or nosepiece is attached to a stem that moves perpendicularly to the sealing surface. Closure and sealing are made by pushing the nosepiece onto the valve seat. To open the valve, the stem is lifted, pulling the nosepiece away from the seat. Poppet

valves are available as right-angle valves where the ports are at 90 degrees to each other, or as "in-line" valves where the ports are spaced 180 degrees apart. In some in-line valve designs the gas path is straight through, while in others it makes several bends within the valve. Figure 2 illustrates these different gas paths.

A gate valve is shown schematically in Figure 3. Here the ports are in-line, and the seal plate is part of a gate or carriage assembly that moves in and out of the gas flow path. To make the seal when the gate is positioned between the ports, a means of expansion is usually provided to force the seal plate against a sealing area concentric to one of the ports.

In most poppet valves, the stem provides the location and guidance of the nosepiece. This is also true in some smaller gate valves, but as the port size of a gate valve increases, this becomes impractical. Instead, the gate assembly is usually guided by the valve body, and some folding linkage is used to decrease the overall length of the valve.

Standard poppet valves are made in port sizes from 0.25 to 10 inches. Gate valves with port sizes from 0.625 to 20 inches are readily available, but both types of valves have been made in larger sizes. Gate valves with rectangular ports, sometimes called *slit valves*, are useful for applications such as handling wafers in semiconductor fabrication or other flat stock processed in vacuum.

Gate valves which use a means of gate expansion, have more moving parts in vacuum and are more complicated than poppet valves, because the seal plate must move in two directions. Other disadvantages include higher cost than a poppet valve, and the gate expansion may produce mechanical jarring unacceptable in some applications. Gate valves without expanding gates are commercially available and have less mechanical jarring on closure. Advantages of gate valves include relatively thin bodies, a large conductance when open, and an optically straight-through flow path. Some gate expansion mechanisms mechanically latch in the closed position, allowing the valve to remain closed and leak-tight in the event of a power or compressed air failure. Advantages and disadvantages of poppet and gate valves are summarized in Table 1.

Diaphragm valves (Figure 4) and butterfly valves (Figure 5) are sometimes used in vacuum systems. In the diaphragm valve, the diaphragm is clamped at its edge and the stem is attached to its center. The diaphragm is often made of a polymer, and to make the seal, it is pushed against the valve body. These valves are made for vacuum service with port sizes up to 2 inches. Butterfly valves rotate a circular vane to open or close the gas flow path. To form a seal in the closed condition, an O-ring is mounted in the periphery of the vane. "Scuffing" or twisting of the O-ring can prevent a leak-tight seal. To help overcome this problem, some butterfly valves cam the seal disk into position at the end of its rotation.

Fig. 2.

"In-line" poppet valve body styles and flow paths.

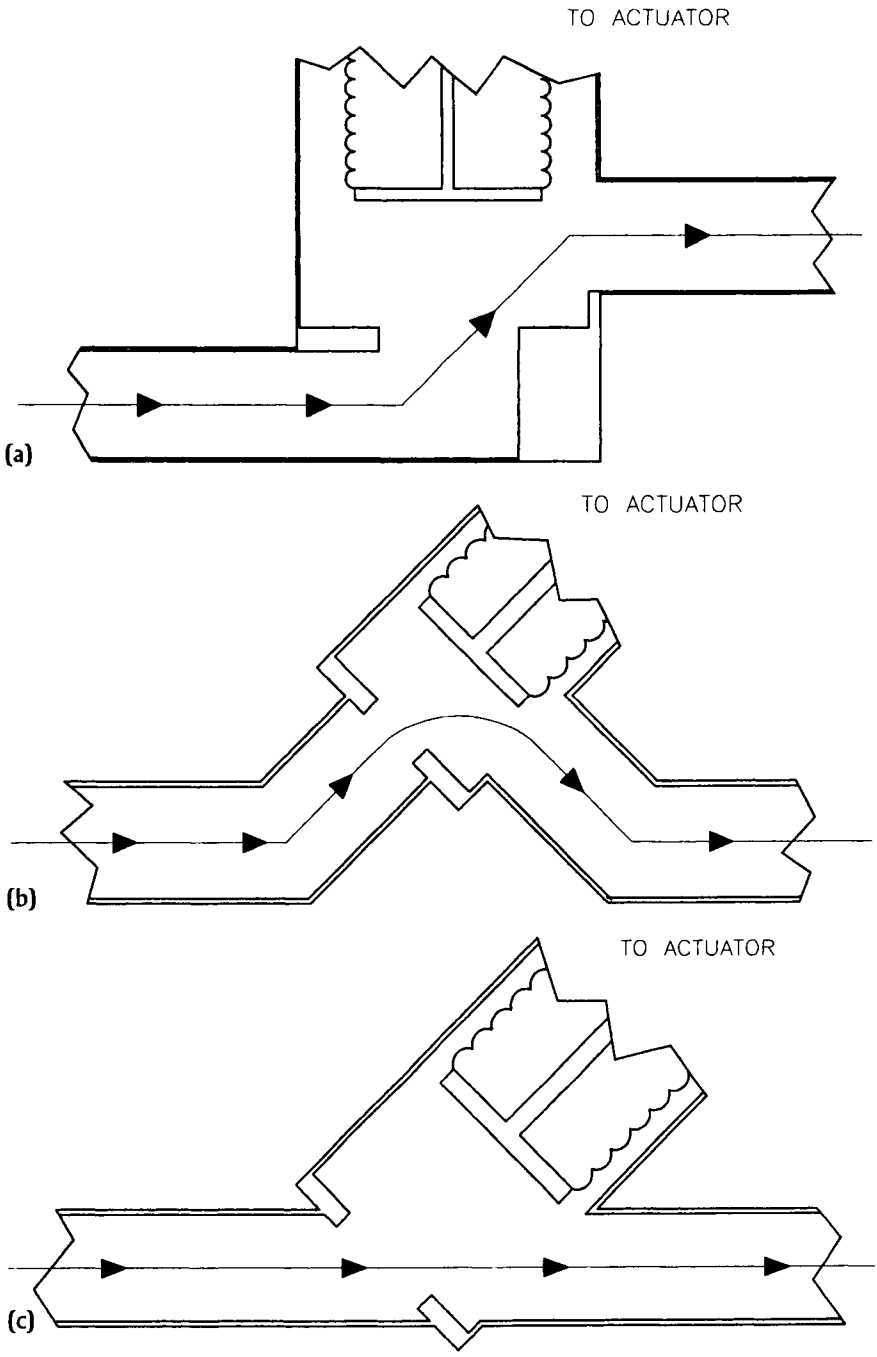
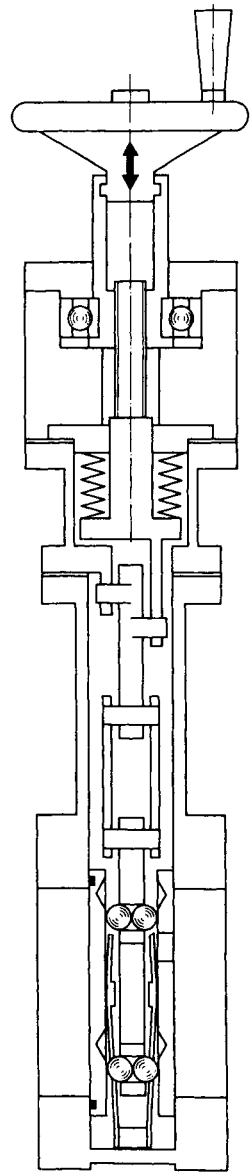


Figure 2(a) illustrates a body where the ports are "in-line," but the attached tubing is offset. In Figure 2(b) the ports are truly in-line, without an offset. The flow path is not "optically straight through." Figure 2(c) illustrates an optically straight through in-line valve body. This style is useful for transfer into vacuum systems or for optical applications.

Fig. 3.

Manually actuated gate valve.



This valve has a polymer nosepiece seal and a metal bonnet seal. Actuation motion into the body is sealed with a bellows. The gate expands and latches the valve closed when the end of the carriage assembly contacts the body, pushing the ball bearings out of their detents (VATLOCK mechanism, U.S. Patent No. 4,052,036). (Reprinted with the permission of VAT Vakuumventile AG, Haag/SG Switzerland.)

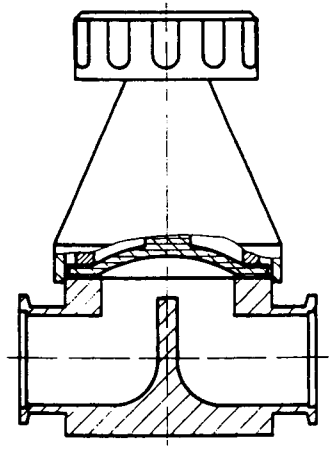
Table I

Comparison of Poppet and Gate Valve Advantages and Disadvantages

Valve Type	Advantages	Disadvantages
Poppet	Simple operating mechanism.	Valve body can be bulky, especially for large port sizes.
	Few moving parts within the vacuum enclosure.	Lower conductance than a gate valve.
	Ports can be in-line.	Gas path with in-line ports may not be optically straight through.
	Lower purchase price.	
	Most designs can tolerate opening with a pressure differential.	
	Spring-closed (or open) types automatically close (open) with power failure, placing the valve in a known state.	
	Some valves incorporate a soft-start bypass feature.	
Gate	Thin body.	Some designs may not tolerate opening against a pressure differential.
	Good conductance.	Operating mechanism is typically more complicated, with more moving parts within the vacuum enclosure.
	Many designs lock in closed position.	Designs with expanding and locking gate mechanisms may produce unacceptable jarring or particulate contamination.
	Short, optically straight-through path for use in load locks or other transfer applications.	Higher purchase price.

Fig. 4.

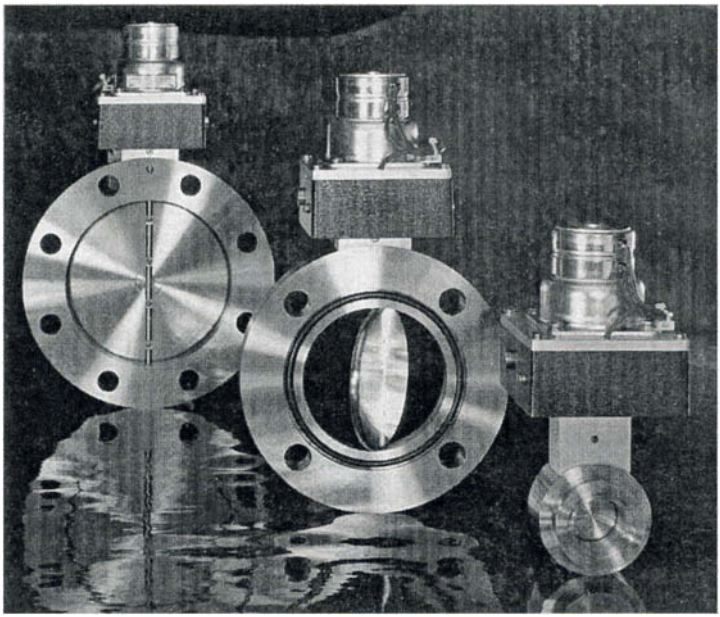
Manually operated diaphragm valve.



The diaphragm is clamped and sealed between the body and bonnet at its periphery. The stem is attached to the center of the diaphragm. For shutoff, the diaphragm is forced against the valve body, isolating the inlet and outlet ports. (Reprinted with the permission of VAT Vakuumventile AG, Haag/SG Switzerland.)

Fig. 5.

Butterfly valve in a partially open condition.



The valves illustrated are actuated with a stepper motor for precise control of the vane angle. This configuration is useful for closed-loop pressure control. When the vane is sealed by an O-ring, butterfly valves can provide shutoff. (Reprinted with the permission of MKS Instruments, Inc., Andover, Massachusetts.)

4.1.3

VALVES FOR CONTROL

Valves can also be used to control pressure in a vacuum system. There are two types of control valves — those that regulate the admission of gas, and those that control the pumping or removal of gas from a system. These are also referred to as *upstream* and *downstream* control valves, respectively. Some control valves do provide shutoff, although it is a misapplication of a control valve to use it for shutoff only.

The upstream/gas admission valves are frequently referred to as *leak* or *metering* valves. One common type is the needle valve, where a tapered needle is moved into an orifice to vary the gas flow. Another type resembles a poppet or diaphragm valve, but the nosepiece seal uses hard materials such as stainless steel or saphire against a soft metal. Polymers are rarely used because creep of the polymer would cause large changes in conductance. Mechanical reduction is frequently used in the drivers of these valves to obtain the minute motion of the nosepiece necessary to accurately control the gas flow. Metering or gas admission valves may not open quickly, and often have only a small conductance in the "full open" case. Piezoelectric leak valves are also available. They do not have mechanical linkages and require only a voltage input as the control signal. Response times are typically several milliseconds.

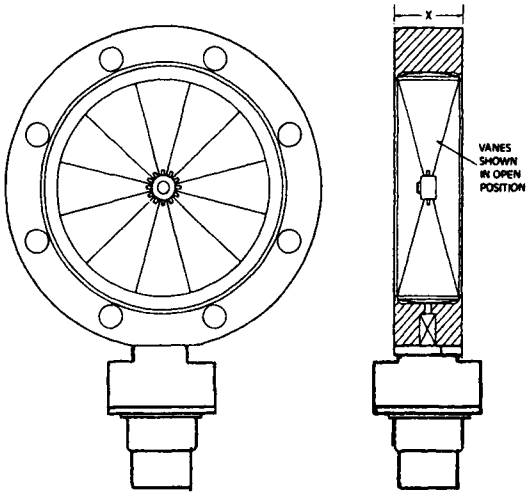
Leak valves are very sensitive to contamination. The gases being controlled should be dry, or a "liquid seal" may form. A liquid seal occurs when adsorbed material completes a seal between the valve nosepiece and the seat. Then, as the nosepiece is retracted, there is no conductance through the valve until the liquid pulls away, breaking the seal. This gives a sudden opening, without any control. Because of the small motions and openings used, leak valves are susceptible to damage from particulates. A vacuum-grade filter made of sintered metal installed on the supply side of the valve can prevent debris from entering the valve.

Downstream control valves or throttling valves are usually mounted near the high-vacuum pump, and regulate the gas flow into the pump. Butterfly valves are often used, because the vane angle can be varied to regulate the flow. When used for control, the vane may not have a seal for complete shutoff. If an O-ring seal is used in the vane, control can be difficult as the vane begins to move. Other downstream control valves move blades across the port to vary the size of the valve opening and conductance. A valve of this type is illustrated in Figure 6. The advantage of this valve is that the blades do not protrude past the body, minimizing the need for clearance in the adjacent components. Specially designed poppet and gate valves are also available for control purposes.

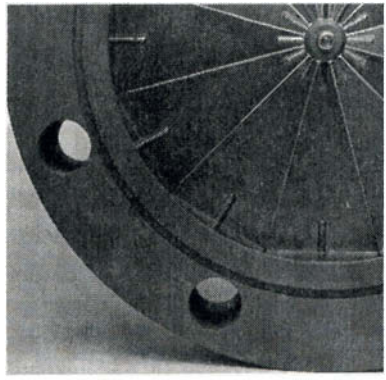
4.1.4 VALVE CONSTRUCTION

Valves differ greatly in method of actuation, means for transferring the actuating motion into the valve body, and the seal materials used. After a valve type has been selected, these points need to be considered.

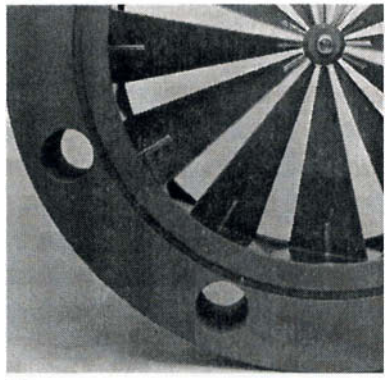
Fig. 6. Multiblade valve for pressure control.



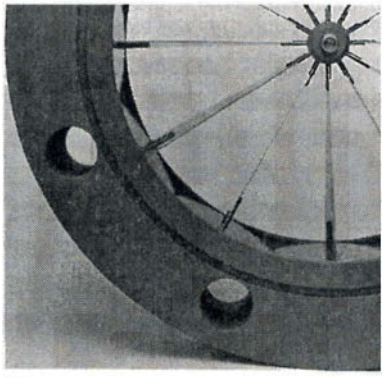
(a)



(b)



(c)



(d)

Flat vanes or blades are mounted on radii of the valve port. A mechanism within the body rotates the vanes simultaneously, changing the conductance of the valve. Three vane positions are illustrated in Figures 6(b) to 6(d). Remote actuation is usually with a stepper motor. While more complicated than a butterfly valve, the vanes do not protrude beyond the valve body. (Reprinted with the permission of MKS Instruments, Inc., Andover, Massachusetts.)

4.1.4.1 Sealing of the Actuating Mechanism

The seal for the transfer of actuating motion into the valve body is usually either an O-ring or a metal bellows. In diaphragm valves, the diaphragm serves as the seal between the valve interior and the actuating motion. When an O-ring is used to seal linear motion, the stem slides through the O-ring as the valve is opened or closed. Although simple in conception, reliable sliding seals are difficult to make and are not suitable for pressures below 10^{-6} torr. O-ring seals used for rotary motion perform better than those for linear motion. In either case, the O-ring will be damaged very quickly unless the stem or rotating shaft has a polished surface and the stem is well lubricated. The grease in a lubricated sliding seal may be a contaminant, and at the least it will "carry in" gas every time the valve closes. Also, initial movement of the stem or shaft may cause a momentary leak past the O-ring, producing a gas burst. Better elastomer seals that minimize these problems use two or three O-rings in series, perhaps with pump out of the region between the O-rings, and provide a reservoir for maintaining lubrication.

A motion feedthrough that overcomes the problems of the sliding O-ring seal uses a metal bellows as illustrated in Figure 1. One end of a bellows is fixed to the valve body, while the other end is attached to the stem. Valves with bellows seals may have a higher initial purchase price, but their vacuum performance is better. Many are suitable for use into the UHV regime. After considerable use, the bellows may fail, but valves are available with lifetimes in excess of 500,000 cycles.

Bellows used in valves are of either the formed or the so-called welded and nested types. Welded ones are made by alternately welding thin rings together on their internal diameter and outside diameter. The convolutions in formed bellows are made by rolling and other forming techniques on lengths of thin-wall metal tubing. Welded bellows are more expensive than formed ones. Also, in a compressed welded bellows there is very little space between adjacent convolutions. With formed bellows, space remains between adjacent convolutions in the compressed condition. This is an advantage in systems where particulates are generated, because the formed bellows is less likely to be damaged. Valves with welded bellows are frequently more compact because the welded bellows for a given stroke is shorter than the formed bellows for the same stroke.

4.1.4.2 Actuation of Valves

Valves can be opened and closed manually, with a pneumatic cylinder, or with a motor. The type of actuation is independent of the stem seal used.

Manual or hand actuation is usually the simplest and the least expensive. Fre-

quently a screw arrangement is used for producing linear motion of the valve stem. Another form of manual actuation sometimes found on smaller valves uses a toggle lever and linkage to the stem.

Manual actuation is not practical for valves cycled frequently or for valves controlled remotely, and pneumatic or motor actuation is then used. For pneumatic operation, a pneumatic cylinder is built onto the valve, often with the piston attached directly to the stem. The forces provided by the actuating cylinder and mechanism determine how much overpressure the valve will seal against, and what internal pressurization the valve can tolerate. The cylinder can be a double-acting type, where the valve is both opened and closed with air pressure, depending on which side of the piston is exposed to the air supply. Single-acting cylinders, where the compressed air either opens or closes the valve, are also used. A spring is then used to move the valve in the opposite direction. Valves with the pneumatic/spring combination have the advantage that in the event of a power or compressed air failure, the valve goes to a known state. Whether this is open or closed depends on the design.

Solenoid control valves are used to regulate the air to the pneumatic actuator cylinder. In some installations, each valve might have its own solenoid, while in others, multiple valves may be operated from the same solenoid. For critical clean room applications, the solenoid or the pneumatic cylinder is vented to a "dirty" area to minimize contamination. When the user provides the pneumatic control, it is necessary to match the solenoid control valve specifications to the vacuum valve's air supply requirements or operation may be adversely affected.

In the case of metal-sealed valves, manual actuation has advantages. Pneumatic actuators tend to be large and expensive because of the higher sealing forces needed. With many metal-sealed valves, it is necessary to reform the seal occasionally. This is done by forcing the soft and hard elements of the seal together with a force greater than the previously used sealing force. This procedure is easier with a manual actuator, particularly when a torque wrench is used on the valve actuator.

When motor actuation is used, a screw or a rack converts the motor's rotary motion to linear motion. In the case of butterfly valves, a gear motor may be coupled directly to the shaft carrying the vane. An advantage of motor operation is that only electrical power need be supplied to the valve. A disadvantage can be that in the event of a power failure, the valve will not go to a known state.

Remotely actuated control valves are often driven by a stepper motor. Together with the valve controller, geared stepper motors can move the valve mechanism by very small and precisely regulated amounts. Some manually actuated valves can be used for control, but the ease of control depends greatly on the valve design.

Limit switches are available on most remotely actuated valves. These switches provide an indication that the valve stem is at one end of its travel or the other and can be used to tell if a valve has obeyed a command to open or close.

4.1.4.3 Seals Within Valves

Valves that provide shutoff have either a metal or a polymer seal between the valve plate and the body. Polymers are inexpensive, but may be damaged by the process, have higher permeation and outgassing rates, can be a source of particulates, and usually are not suitable for UHV pressures. Bakeout temperatures of valves are usually limited by any polymers used—either within the vacuum, or for pneumatic seals of the actuator. Metal-sealed valves often have higher bakeout temperatures, but some of these valves cannot be baked in the closed position. For a more complete discussion and comparison of sealing materials, the reader is refered to Chapter 4.6 on sealing.

O-rings used as gate or nosepiece seals must be properly retained, or they may come out of their groove during operation. Suitable O-ring grooves have a trapezoidal cross section and are vented. Polymer seal materials can also be bonded directly to the valve plate. Valves with metal gate or nosepiece seals are more expensive, because the operating mechanism is more complicated and better surface finishes are needed on sealing surfaces. The number of closures that can be obtained with metal plate seals depend on the application and the care during the use of the valve. Metal seals are more susceptible to damage by particulates than are polymer seals.

Many valves have a demountable body seal, often called the *bonnet* seal, which allows the components within the vacuum to be removed for cleaning and service. The bonnet flange seal can be an O-ring or other shaped elastomer, or a metal seal. On some valves, metal and polymer seals can be used interchangeably. Valves for UHV applications may eliminate the body seal by using an all-welded construction for the body.

Many valves use a metal seal for the bonnet, and a polymer for the valve plate seal. This has the advantages of a metal seal for the static body seal without the increased cost and complexity of a sealing mechanism for a metal plate seal. With this arrangement, the polymer is immersed in vacuum. As explained in Chapter 4.6, an O-ring in vacuum is much less of a gas source than an O-ring exposed to atmosphere.

4.1.4.4 Flanges, Tubulations, and Valve Conductance

If a valve is attached with flanges, it can be removed for service or for modification of the system. Flange systems are described in Chapter 4.2. In some cases, valves are permanently installed by welding or brazing directly to the system piping.

The conductance of the valve and any associated piping should be considered

Table 2

Measured Conductances of Angle Valves as a Function of Pressure

Angle Valve Dimensions (inches(mm))		Conductance (L/sec)			
Inlet Diameter	Axial Length	At 10^{-3} torr	At 10^{-2} torr	At 10^{-1} torr	At 1 torr
1.26 (32)	5.39 (137)	14	27	100	500
1.97 (50)	6.10 (155)	26	47	170	1000
2.56 (65)	7.28 (185)	50	160	850	4500
3.94 (100)	9.25 (235)	90	340	—	20,000

Source: Based on information given by H. A. Steinherz, in *Handbook of High Vacuum Engineering* (Reinhold, New York, 1963), p. 21.

in system design. Formulas and a comprehensive discussion of conductance calculations in general can be found in Roth [1] or Berman [2]. The literature does not contain many papers on the conductances of valves. Conductances of butterfly valves were calculated and measured by Hummer, Halter, and Grössl [3].

Angle valve conductances are frequently taken to be "similar to an elbow of the same nominal size." Steinherz gave some measured values for the conductance of several sizes of angle valves [4]. These are tabulated in Table 2. These data illustrate how the conductance of a component depends on the pressure–flow regime. Rozanov, Shchenev, and Akimov [5] calculated angle valve conductances using Monte Carlo methods. They also modeled these valves analytically, as a series of tubes, apertures, and baffles. Actual valves were then experimentally tested, and good agreement to both models was found. This work showed that clearance around the bellows was important for high conductance. More recently Hablanian, Nuzzi, and Pflanz measured the conductance of angle valves where the body dimensions had been optimized for high conductance [6]. They also found that clearance between the body and the bellows had more effect on conductance than the valve opening. In a valve where the nosepiece cleared the sideport, conductance did not increase during the last 20% of the valve's opening stroke. Rozanov and associates also observed that after a certain valve opening, conductance did not greatly increase. The advantage of designs with a shorter stroke is that bellows life is significantly increased. Hablanian and associates examined an electromechanically operated valve without a bellows, and found that its conductance was approximately 90% of an elbow of the same nominal size.

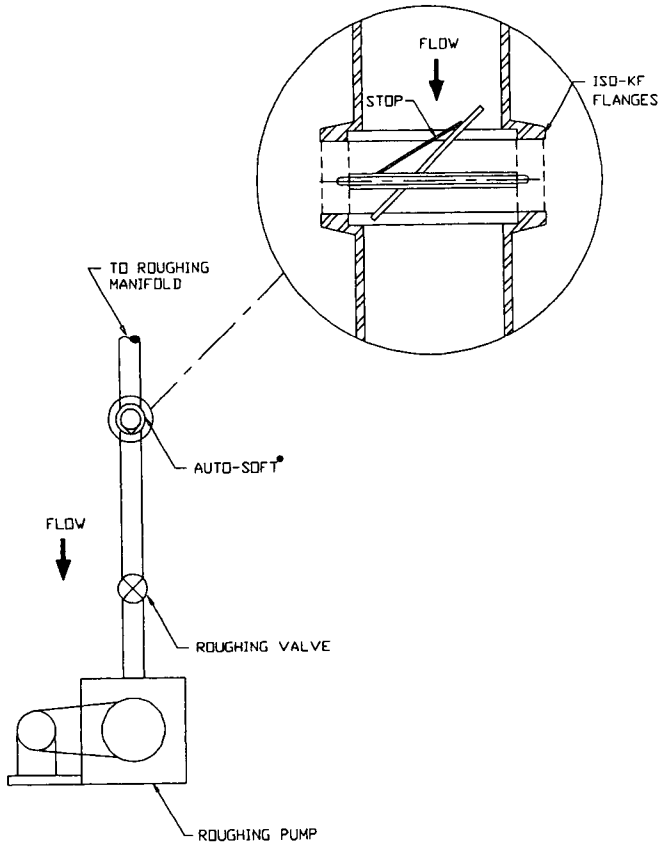
4.1.5

SPECIALTY VALVES

Some vacuum valves are designed for specific applications. For instance, it may be desirable to rough a process chamber slowly to reduce damage to fragile items like wafers or to decrease particle counts. Simple vane-type valves are available that automatically perform this function and are illustrated in Figure 7. These are

Fig. 7.

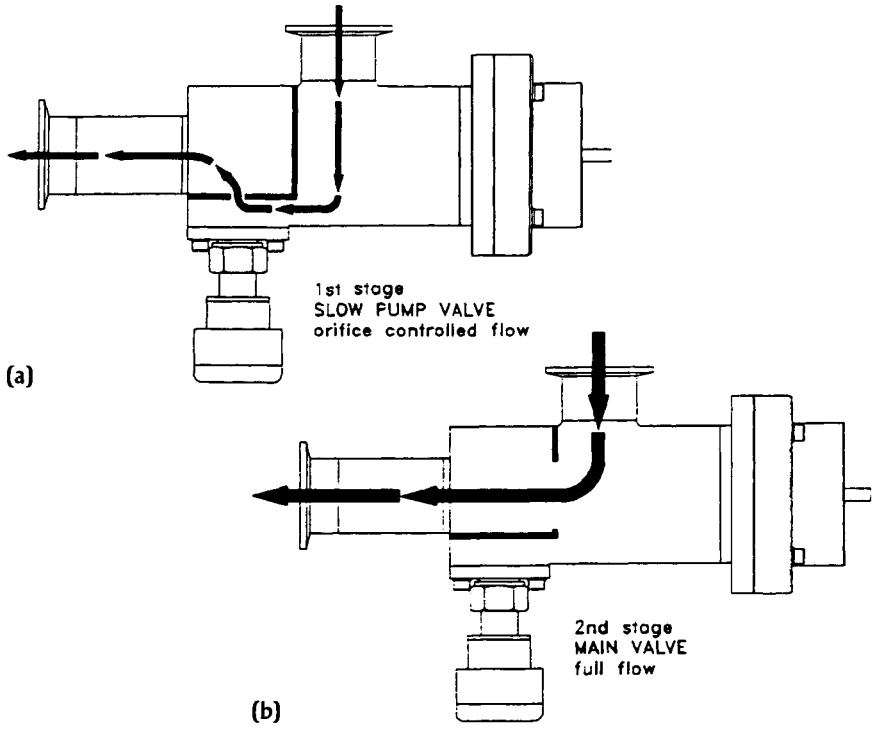
Small automatic valve for "soft starts."



The inset shows a small spring-operated valve for automatic soft roughing to reduce turbulence during system roughing. Installed in the roughing line, initial gas flow closes the valve and pumping continues around the periphery of the vane. When the flow has decreased, the valve springs open for greater conductance. (Reprinted with the permission of the HPS Division of MKS Insruments, Inc., Boulder, Colorado. Auto-Soft is a registered trademark of the HPS Division of MKS Instruments, Inc., Boulder, Colorado.)

Fig. 8.

Poppet valve incorporating a bypass for soft starts.



Initial roughing is through a small orifice as shown in Figure 8(a). After preliminary roughing, the main poppet is opened for increased conductance, as in Figure 8(b). In the valve illustrated, the bypass orifice can be closed with a poppet mechanism for shutoff. (Reprinted with the permission of the HPS Division of MKS Instruments, Inc., Boulder, Colorado.)

installed in a flange joint of the roughing line, and the gas movement at the start of pumpdown closes the vane. Once the gas flow has decreased, the valve automatically springs open for greater conductance.

Poppet valves that include this soft start function are also available. Figure 8 illustrates the flow paths in these valves. Initial roughing is through a bypass with a small orifice, and the nosepiece is opened later in the pumpdown. Such valves are sometimes called *two-stage valves*. The bypass contains a small poppet mechanism to provide shutoff and can be integral with the body. In some cases, pumpdown characteristics can be varied to suit a particular application and system by adjusting the small poppet mechanism.

Another specialized vacuum valve is designed to provide isolation and venting of mechanical pumps in the event of a power failure. This helps to maintain vacuum in the system and prevent system contamination by oil backing up from the

Fig. 9.

Valve for automatic system isolation and pump venting on power loss.

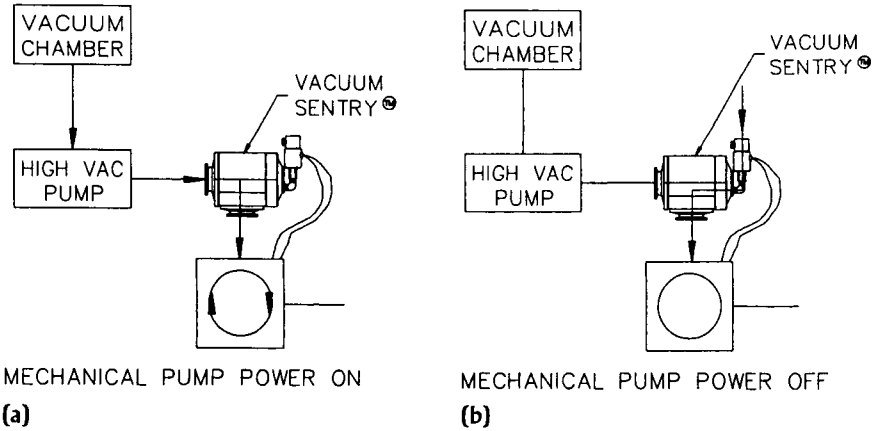


Figure 9(a) illustrates normal, power-on operation. In Figure 9(b), power to the pump and the solenoid valve has been interrupted. Atmospheric pressure has closed the valve, isolating the system, and vented the mechanical pump. (Reprinted with the permission of the HPS Division of MKS Instruments, Inc., Boulder, Colorado. Vacuum Sentry is a registered trademark of the HPS Division of MKS Instruments, Inc., Boulder, Colorado.)

pump. Use of this valve is illustrated in Figure 9. The valve is installed near the pump inlet, and its control solenoid is wired in parallel with the pump motor. When power fails, the valve closes very quickly, isolating the system from the pump. The pump side of the valve is automatically vented to prevent oil backup and for easier restarting of the pump. After power is restored, the valve does not open until the pump side has been evacuated, minimizing any pressure burst to the system on opening. It is important that the valve be installed with the proper flow orientation, so the pump side, not the system side, is vented on closure.

Some applications may require a modified or custom valve. Common modifications to a standard valve include the addition of ports for pumpout or gauge installation. Special materials may also be used to achieve a low magnetic permeability or for increased chemical resistance. Gate valves can be modified to include optical windows in the gate or water-cooled gates for metallurgical processes. Valves for nuclear accelerators can be very specialized, providing rf shielding of the body cavity, or closing times on the order of milliseconds to protect equipment.

Table 3 summarizes the common valve types that have been discussed in this section.

Table 3

Primary Use and Features of Commonly Available Vacuum Valves

Valve Type	Primary Use or Function	Features/Comments
Poppet	Shutoff/isolation	Usually inexpensive and reliable; available with metal seals for UHV applications; bodies with "in-line" ports available; may have control as well as shutoff capabilities.
Gate	Shutoff/isolation	Bodies are frequently thin and ports are in-line; available in a wide variety of sizes and with either elastomer or metal seals.
Diaphragm	Shutoff/isolation	Very low cost, usually only available in port sizes less than 2.0 in; diaphragms are frequently elastomer, although metal ones are used in small valves suitable for gas handling.
Butterfly	Shutoff or control	Manually actuated valves are inexpensive; motor-driven types are suited to closed-loop pressure control.
Iris/Blade	Control	For downstream pressure control; often motor driven.
Leak or Needle	Gas admission	For controlled admission of minute quantities of gas into vacuum systems; many designs can provide shut-off of the gas supply.
Pump Protection		Isolates the vacuum system and automatically vents the pump in the event of a power failure.
Soft Start	Soft pumping	Slows initial pumping of a chamber to reduce particulates. This may be the only function of the valve or it can be incorporated in another valve.

4.1.6

INSTALLATION CONSIDERATIONS FOR VACUUM VALVES

The location of a valve in a system requires thought. Pressure differentials across the valve or loading due to atmospheric pressure can prevent a valve from sealing or opening and may shorten its lifetime. The valve location and orientation should be chosen to minimize the formation or trapping of contamination. Mild contamination may prevent a valve from sealing, while gross contamination can prevent a valve mechanism from operating. Some valves are available with heating mantles to help prevent condensation of process products. Care should also be taken to see that the valve is not subjected to temperatures higher than those recommended by the manufacturer. Overtemperature can decompose polymer seals or other parts and cause the mechanism to seize up. If there could be loose objects in the system, the valve should be oriented so these will not find their way into the valve. The location should also allow enough space for service access, including removal of the operating mechanism.

Care must be taken to see that valves are not damaged during installation. Seal surfaces on flanges should be protected. Valve bodies have been destroyed by using bolts that are too long in flanges with tapped bolt holes. Any valve can be damaged by undue stress, although gate valves are more susceptible. Long runs of piping or other heavy objects attached to the valve body are a common source of stress. To prevent damage in such cases, stress relief must be provided.

Finally, the manufacturer's recommendation for maintenance should be followed. This is especially true with metal-sealed valves and other valves subjected to bakeout, high temperature, or continuous use where the driver or operating mechanism may need periodic lubrication.

REFERENCES

1. A. Roth, *Vacuum Technology*, 2nd ed (Elsevier Science, Amsterdam, 1982).
2. A. Berman, *Vacuum Engineering Calculations, Formulas and Solved Excercises* (Academic Press, San Diego, 1992).
3. G. Hummer, G. Halter, and M. Grössl, *Vacuum*, **41** (1990) 2126.
4. H. A. Steinherz, *Handbook of High Vacuum Engineering* (Reinhold, New York, 1963), p. 21.
5. L. N. Rozanov, V. V. Shchenev, and Yu. D. Akimov, *Sov. Phys. Tech. Phys.*, **22**(10) (1977)1249. Originally published in *Zh. Tekh. Fiz.*, **47** (October 1977) 2151–2156.
6. M. H. Hablanian, F. J. Nuzzi, and T. L. Pflanz, *J. Vac. Sci. Technol.*, **A9**(3) (1991) 2062.

CHAPTER 4.2

Flange and Component Systems

Neil T. Peacock
HPS Division of MKS Instruments, Inc.

4.2.1 INTRODUCTION

Demountable flanges have made today's vacuum systems possible by allowing nearly anyone to quickly and easily design and assemble a complex vacuum system. Until the early 1960s, building a vacuum system required skilled personnel. Smaller laboratory systems were frequently made of glass, and construction or modification required a scientific glassblower. Larger systems were often assembled by sweating or brazing copper plumbing fittings and adapting other components to vacuum service. Producing clean, reliable, easy to modify, and leak-free vacuum systems was difficult. Standardized flanges and components with flanges designed for vacuum service have eliminated many of these problems.

Many flanges are part of a "system," with components like Tees and crosses forming modular "building blocks." This allows for easy modification or substitution of components. With the system concept, components can be quickly assembled for a particular experiment, then disassembled and reused in another arrangement. When these flanges and components are built to international or industry standards, parts from different sources or manufacturers can be intermixed.

In a few applications and installations, demountable joints are undesirable. Ex-

ISBN 0-12-325065-7
$25.00

amples include the handling of hazardous gases or materials that would damage the joint or its seal. In these cases, components are usually welded or brazed together.

4.2.2 SELECTING A FLANGE SYSTEM

When selecting a flange system, there are many points to keep in mind. The first is whether the joint must be metal sealed, or if an elastomer is acceptable. Metal and polymer seal materials are compared in Chapter 4.6. Next is the material for the flange. Typically the chamber, the flanges, and components of a system are made of the same material. Flange pairs of dissimilar metals should be avoided on baked systems because differences in thermal expansion rates could cause leaks. Also, the material needs to be suitable for the method of attachment. When brazing, the material must be compatible with the braze alloy. Welding is generally straightforward when the flange and the component are the same weldable alloy. For stainless steel flanges, parameters such as the ferrite numbers of the metals being welded influence the weldability [1–3].

Other flange selection considerations include the pressure range, with any overpressure requirements, and operating temperature extremes. How flange pairs are held together is an additional consideration. Flanges are often bolted together, but for frequently disassembled joints, or components moved from system to system, a flange held together with an easily removed clamp may be advantageous. Band clamps also facilitate assembly with remote handling equipment as required in high-radiation environments.

Flanges for critical UHV applications are subject to additional considerations. In these cases, vacuum remelt or forged material should be specified to help avoid porosities and leakage paths. The hardness of the flange material is important for flanges with knife edges, so that the flange is not damaged with repeated use. Some specialized applications may require additional material specifications. For example, 316 stainless steel is used when greater chemical resistance is required, and low-magnetic-permeability stainless steel is often used in particle accelerators.

4.2.3 COMMON FLANGE SYSTEMS

Numerous flange systems are in current use. Features of the most widely used ones are compared in Table 1. In this section, descriptions and basic dimensions

Table 1

Comparison of Flange Systems

Flange Type	Tubing Sizes Accommodated[a]	Primary Seal Material[b]	Minimum Usable Pressure (torr)[c]	Applications
Beaded-Edge Face-Seal[d]	$\frac{1}{8}$″ to 1.0″	Metal	10^{-9}	Gas-handling lines, gauge attachment
ISO 2861/I[e] (smaller ISO series, often called ISO-KF)	$\frac{3}{8}$″ to 2.0″	Elastomer	10^{-8}	Forelines, general high-vacuum use, frequently modified systems
ISO 1609/1986[e] (larger ISO series, multifastener type)	2.5″ to 12″	Elastomer	10^{-8}	Pumps, high-conductance forelines, general vacuum use, frequently modified systems
ASA (ANSI)	1.5″ to 12″	Elastomer	10^{-8}	Found on high-vacuum pumps, older vacuum systems
Knife-Edge[d] (CF)	0.75″ to 10″	Metal	10^{-11}	Baked systems, UHV systems
Wire-Sealed	12″ to 18″	Metal	10^{-11}	Bell jars or ports on large systems
MESA or SEMI	Rectangular	Elastomer	10^{-8}	Designed for semiconductor fabrication equipment; ports are sized for passage of 200-mm or 300-mm wafers

[a] Additional sizes may be available.

[b] Primary seal material refers to the designer's intended seal material. Both elastomer and metal seals are available for some flange systems.

[c] With elastomer seals, the minimum usable pressure is often determined by outgassing of the seal material. The values given are conservative. The maximum usable pressure for a flange depends on how well the gasket is contained. In knife-edge and beaded-edge face seal flanges, the gasket is well contained and large internal or overpressures are acceptable. Unless the O-ring outside diameter is confined by an overpressure ring, ISO flanges will not tolerate internal pressure differentials much greater than 5 psig.

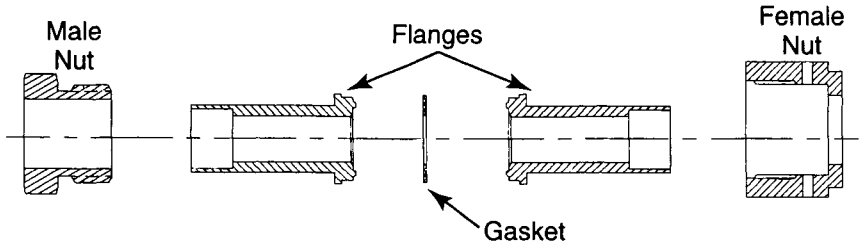
[d] Flanges of this type are sold under a variety of trade names.

[e] Refers to flanges made according to this ISO (International Standards Organization) standard. These flanges are sold under a variety of names.

of common flanges are provided. The designation of flange sizes can be a source of confusion. The outside diameter (O.D.) is sometimes used as the size, while other times the largest tubing size a flange can accommodate is used. With ASA (ANSI) flanges, the nominal pipe size used with the flanges in steam service is still used as the flange size. When specifying and ordering flanges, enough dimensions should be given to identify the flange with certainty.

Fig. 1.

A beaded-edge face-sealed coupling. The gasket is usually copper or nickel.



4.2.3.1 Beaded-Edge Face-Sealed Flanges

Figure 1 illustrates a coupling frequently used on gas supply lines and for gauge attachment. It is made up of two identical flanges, a gasket, and a male and a female nut. The nuts are installed before the flanges are welded into place. On the face of each flange is a raised bead of semicircular cross section. The seal is made when the beads penetrate a gasket as the nuts are tightened. Flanges of this style are available from many manufacturers, and are often interchangeable. Some manufacturers refer to the flanges as *glands*. Elbows, Tees, crosses, and various adaptors in this system are commercially available.

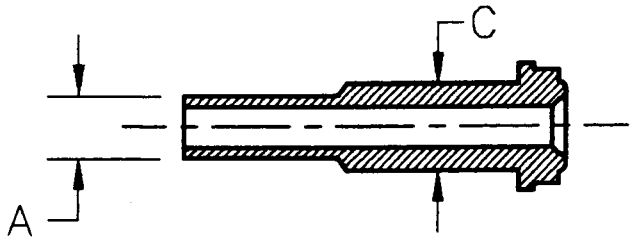
Viton®, copper, nickel, and stainless steel are common gasket materials. Some gaskets have a groove in each side to help align the flanges. In some situations, gaskets with retaining clips can aid installation. Gaskets incorporating sintered filter elements that are capable of trapping particles as small as a few microns are useful for applications where particulates could cause damage.

Table 2 gives dimensions of these flanges. The couplings for $\frac{1}{4}$- and $\frac{1}{2}$-inch O.D. tubing are those most commonly encountered in vacuum work, although flanges are made for up to 1.0-inch tubing. Fittings with electropolished internal surfaces are available for use in ultra-clean applications.

Advantages of these flanges are that they do not require axial clearance to install components, that they are suited to automated welding methods, and that they will tolerate high overpressures. Disadvantages are that the sealing area on the raised bead is not tolerant of scratches or other damage, and that the large sizes require considerable tightening torque to achieve mass spectrometer leak-tight seals.

Table 2

Typical Dimensions of Beaded-Edge Face-Sealed Flanges; Flanges of this type are not standardized within the vacuum industry, and some dimensions may vary



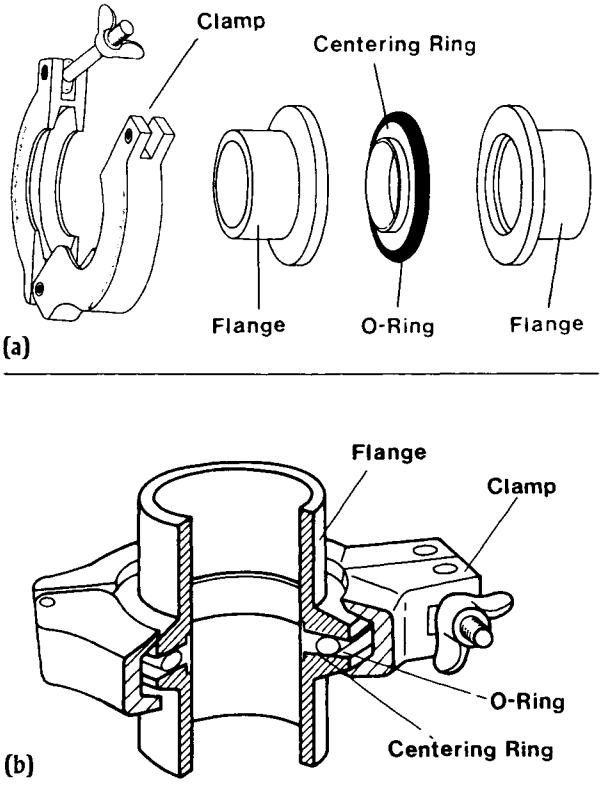
A Nominal Tube Size (in)	C Maximum Body Diameter $\left(\frac{\text{in}}{(\text{mm})}\right)$	Hex Size, Male Nut (in)	Hex Size, Female Nut (in)	Outside Diameter of Gasket $\left(\frac{\text{in}}{(\text{mm})}\right)$
0.125	0.20 (5.1)	$\frac{3}{8}$	$\frac{7}{16}$	0.25 (6.3)
0.25	0.35 (8.9)	$\frac{5}{8}$	$\frac{3}{4}$	0.48 (12.2)
0.375	0.47 (11.9)	$\frac{3}{4}$	$\frac{7}{8}$	0.61 (15.5)
0.50	0.60 (15.2)	$\frac{15}{16}$	$1\frac{1}{16}$	0.78 (19.8)
0.75	0.80 (20.3)	$1\frac{5}{16}$	$1\frac{1}{2}$	1.14 (28.9)
1.0	1.19 (30.2)	$1\frac{5}{8}$	$1\frac{3}{4}$	1.40 (35.6)

4.2.3.2 ISO O-Ring Sealed Flanges

The International Standards Organization (ISO) in Geneva, Switzerland, has published dimensional standards for two different series of O-ring sealed vacuum flanges. The smaller series is often referred to as ISO-KF and the larger one as ISO-MF or ISO-K.

ISO-KF flanges are covered by ISO Standard 2861/1 — "Vacuum technology — Quick release couplings — Dimensions" and are available for tubing sizes up to 2.0 inches. The system originated in Europe and the KF comes from the German

Fig. 2.



(a) A joint using KF flanges is made with two identical flanges, an O-ring carried on a centering ring, and a clamp. The centering ring limits compression of the O-ring and provides alignment of the flanges. (b) The assembled KF joint. (Reproduced with the permission of the HPS Division of MKS Instruments, Inc., Boulder, Colorado.)

word *Kleinflansch*, meaning "small flange." These flanges are also called "QF," "quick flanges" or "NW" flanges by some manufacturers. "NW" was mistakenly picked up from European flange tables, where it was used as the abbreviation for *Nennweite*, meaning "nominal diameter."

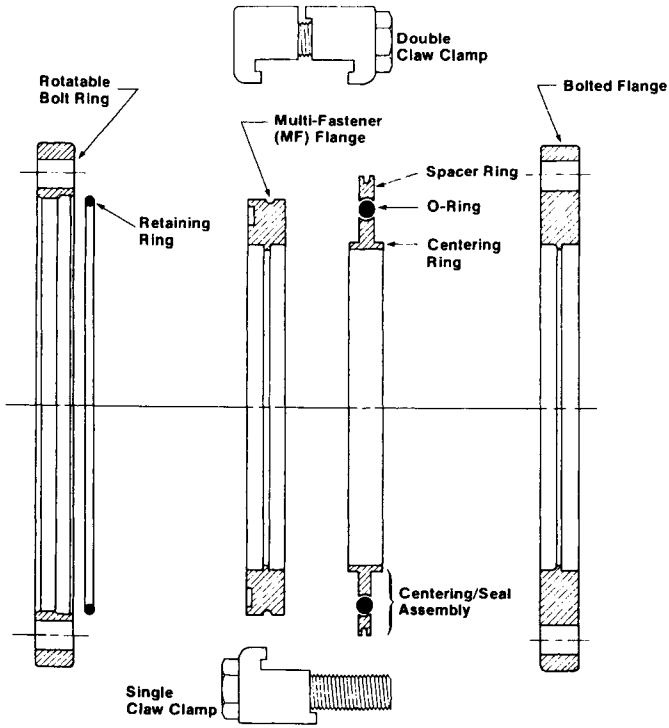
A KF flange joint is illustrated in Figure 2. A joint consists of two identical and infinitely rotatable flanges, a centering ring with seal, and a clamp. The centering ring centers and spaces the flanges, retains the O-ring seal, and limits the O-ring compression. The clamp holds the joint together and provides the sealing force. Some clamps are closed with a wingnut, others with a toggle lever. Lever-closed clamps are convenient when the joints are frequently disassembled. Clamps are correctly tightened when the flange faces contact the spacer portion of the center-

ing ring. This contact is necessary to make reliable joints. Wingnut-closed clamps should be tightened until a sudden, sharp increase in closing torque is felt. Lever-closed clamps require some spring means to provide the correct sealing force and ensure flange to centering ring contact in any situation.

A second series of O-ring sealed flanges is described by ISO Standard 1609, "Vacuum Technology—Flange Dimensions." Of the flanges described by this standard, only those for nominal tubing sizes of 2.5 inches and larger are in common use. This system is variously called ISO-MF, ISO-Universal, ISO-K, or *PNEUROP*. They too are sealed with an O-ring mounted on a centering ring, and the flanges are sexless. Figure 3 shows the hardware used in this system.

Fig. 3.

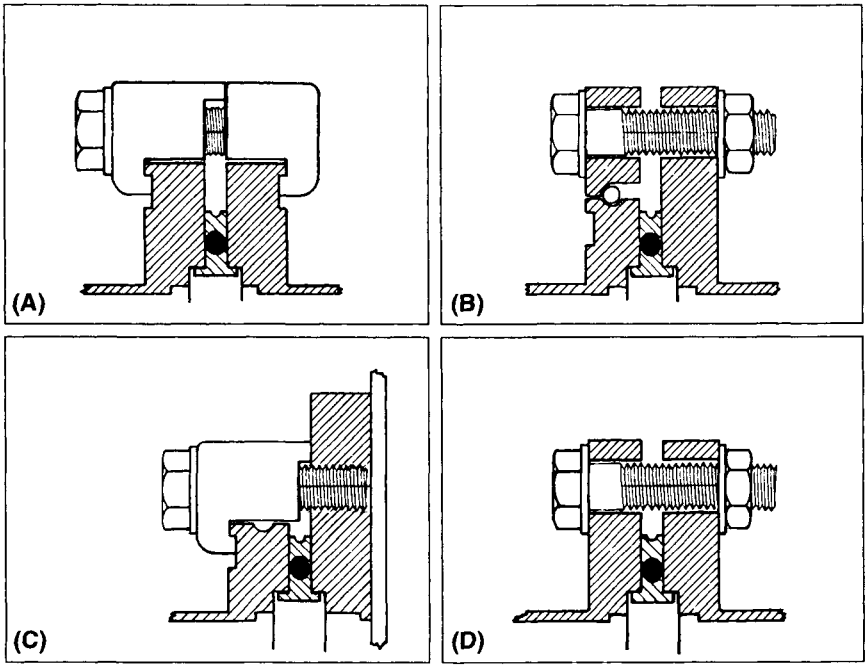
Flanges and hardware fasteners belonging to the larger ISO (International Standards Organization) O-ring-sealed flange system.



The names are for descriptive purposes and do not reflect any universally accepted name. (Reproduced with the permission of the HPS Division of MKS Instruments, Inc., Boulder, Colorado.)

Fig. 4.

The various ways the larger ISO (International Standards Organization) O-ring-sealed flanges can be held together.



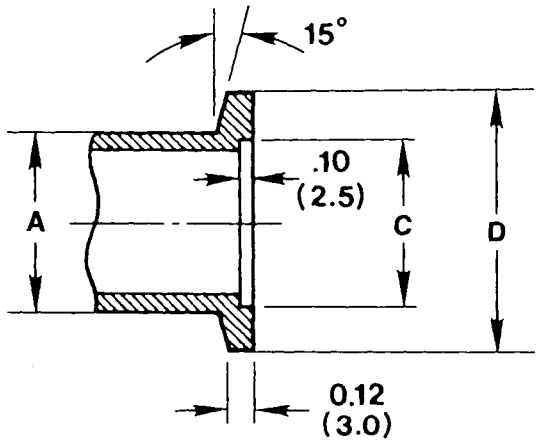
(Reproduced with the permission of the HPS Division of MKS Instruments, Inc., Boulder, Colorado.)

The various larger flange styles and the ways of holding them together are illustrated in Figure 4. The ISO-MF name came about because of the several methods of fastening these flanges together, "MF" standing for "Multi Fastener flange." Both flanges can be infinitely rotatable, held together with double-claw clamps as in Figure 4(a), or they can be bolted as in Figure 4(d). Bolted flanges are primarily found on components such as gate valves, where there is not space for a clamp. Nonbolted flanges can be joined to a tapped flange by using single-claw clamps, as in Figure 4(c), or with the addition of a rotatable bolt ring that fits over the outside diameter of the flange, as illustrated in Figure 4(b). Flanges held together with clamps or rotatable bolt rings have the advantage of infinite rotatability. The centering ring and the seal region for a given size flange are always the same. As with the ISO-KF flanges, MF flange joints are correctly tightened when the seal surfaces contact the spacer portion of the centering ring.

Table 3 gives dimensions of common KF flanges, and Table 4 dimensions of

Table 3

KF Flange Dimensions: Stub dimensions suitable for welding to inch-sized tubing are given. All other dimensions in the table are according to the ISO (International Standards Organization) or PNEUROP Standards. The ISO standard specifies a minimum length of 10 mm (0.39 inches). (Reproduced with the permission of HPS Division of MKS Instruments, Inc., Boulder, Colorado.)

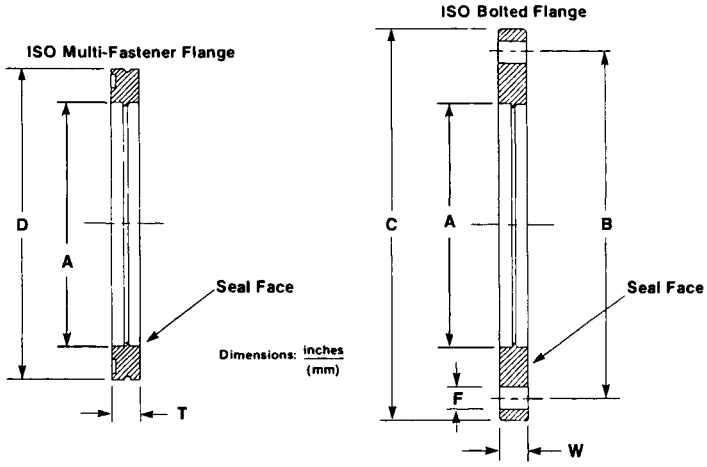


Nominal Size (mm)	A Nominal U.S. Tube Diameter	D Outside Diameter	C Counterbore Diameter
10	0.50 (12.7)	1.18 (30)	0.48 (12.2)
16	0.75 (19.0)	1.18 (30)	0.68 (17.2)
25	1.00 (25.4)	1.57 (40)	1.03 (26.2)
40	1.50 (38.1)	2.16 (55)	1.62 (41.2)
50	2.00 (50.8)	2.95 (75)	2.05 (52.2)

Dimensions: $\frac{\text{inch}}{(\text{mm})}$

Table 4

Dimensions of the Larger ISO (International Standards Organization) O-Ring-Sealed Flanges. The weld socket dimensions are typical for flanges made to fit inch-sized tubing. All other dimensions in the figure are according to the ISO standard. (Reproduced with the permission of HPS Division of MKS Instruments, Inc., Boulder, Colorado.)



Nominal[a] Flange Size Designation (mm)	A Suggested Tube/Pipe Dims.[b] o.d. × Wall (inches)	ISO Multifastener Flange (ISO-MF) D (o.d.)	T (thickness)	No. Clamps Required	ISO Bolted Flange C (o.d.)	W (thickness)	B (bolt circle)	F (bolt holes)	N (no. holes)	Wrenc (mm
NW 63	2.5 × .065	3.74	.47	3–4	5.12	.47	4.33	.35	4	13
		(95)	(12)		(130)	(12)	(110)	(9)		
NW 80	3.0 × .065	4.33	.47	4–8	5.71	.47	4.92	.35	8	13
		(110)	(12)		(145)	(12)	(125)	(9)		
NW 100	4.0 × .065	5.12	.47	4–8	6.50	.47	5.71	.35	8	13
		(130)	(12)		(165)	(12)	(145)	(9)		
NW 160	6.0 × .083	7.09	.47	4–8	8.86	.63	7.87	.43	8	17
		(180)	(12)		(225)	(16)	(200)	(11)		
NW 200	8.0 × .120	9.45	.47	6–12	11.22	.63	10.24	.43	12	17
		(240)	(12)		(285)	(16)	(260)	(11)		
NW 250	10.0 × .125	11.42	.47	6–12	13.19	.63	12.20	.43	12	17
		(290)	(12)		(335)	(16)	(310)	(11)		
NW 320	12.75 × .165	14.57	.67	8–12	16.73	.79	15.55	.55	12	19
		(370)	(17)		(425)	(20)	(395)	(14)		
NW 400	16.0 × .188	17.72	.67	8–16	20.08	.79	18.90	.55	16	19
		(450)	(17)		(510)	(20)	(480)	(14)		
NW 500	20.0 × .188	21.65	.67	12–16	24.02	.79	22.83	.55	16	19
		(550)	(17)		(610)	(20)	(580)	(14)		
NW 630	25.0 × .25	27.16	.87	16–20	29.53	.94	28.35	.55	20	19
		(690)	(22)		(750)	(24)	(720)	(14)		

[a]NW is retained here by convention. It does not specify the flange type but is from the German *Nennwei* meaning "nominal diameter."

[b]Schedule 5 pipe for sizes 320 and above.

Dimensions: $\frac{\text{inch}}{(\text{mm})}$

MF flanges. The flanges described in the ISO standards are dimensioned for welding to metric tubing. Flanges made in the United States generally have inch-sized stubs or tube sockets, while the flange interface dimensions are made to ISO standards. This allows flanges from different manufacturers or countries to be used together. As of this writing, the ISO standards do not include the KF 50 size. This is because the 50 size is not from the "first ISO preference series" or progression. For more explanation, the reader is referred to the ISO Standard, ISO 3, "Preferred Numbers-Series of Preferred Numbers." Two other KF sizes from the second preference series have been made — 20 and 32, but these are no longer in common use. PNEUROP standard 6606/1981, prepared by the European Committee of Manufacturers of Compressors, Vacuum Pumps and Pnuematic Tools, includes the KF 50 flange. The dimensions in this standard are used by most flange manufacturers and are under consideration for addition to the ISO standard.

Either ISO flange is available with weld stubs, weld or braze sockets, unbored or blank for user modification, or in adaptors to other flange or fastening systems such as pipe threads. Both types of ISO flanges are primarily designed to be O-ring sealed, but metal seals are available. To obtain the higher sealing forces necessary for metal seals, special stronger clamps must be used with KF flanges, while with MF flanges additional clamps must be used.

KF, and smaller MF flanges, are particularly suited for forelines. They are also convenient for building vacuum systems for short-term use, or where frequent disassembly for maintenance or modification is required. The lower pressure limit is usually determined by outgassing of the O-ring seal material. KF flanges are not very tolerant of large pressure differentials. If these flanges are exposed to more than a 5 psig differential, over pressure rings that surround and confine the O-ring are necessary.

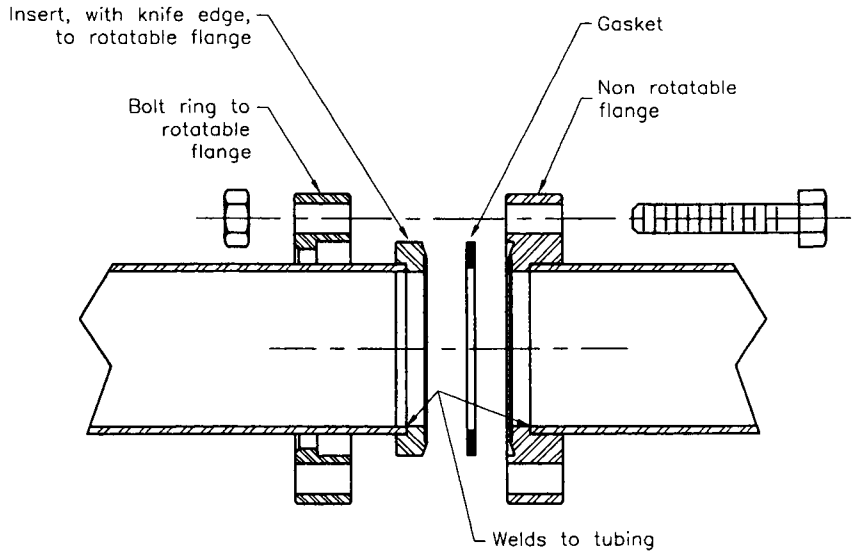
4.2.3.3 Knife-Edge-Sealed Flanges

In the early 1960s Varian Associates introduced the ConFlat® series of metal sealed flanges for UHV applications. Interchangeable flanges, often referred to as CF flanges, are available from numerous manufacturers.

A joint with these flanges is illustrated in Figure 5. The sealing mechanism of the CF flange is described and discussed in Section 4.6.3.4 on seal design and materials. Briefly, the seal is formed by knife edges in each flange penetrating a soft metal gasket between them. The knife edges are slightly recessed below the flange face for protection. Bolts hold the flange pair together and provide the seal force. The flanges are sexless, but because of the bolts they are not infinitely rotatable. To provide rotational alignment, a two-piece rotatable version is used. This is il-

Fig. 5.

Flanges and hardware for a knife-edge flange joint.



Nonrotatable and a rotatable flange are used together to provide rotational alignment. Bolt holes are aligned by rotating the bolt ring.

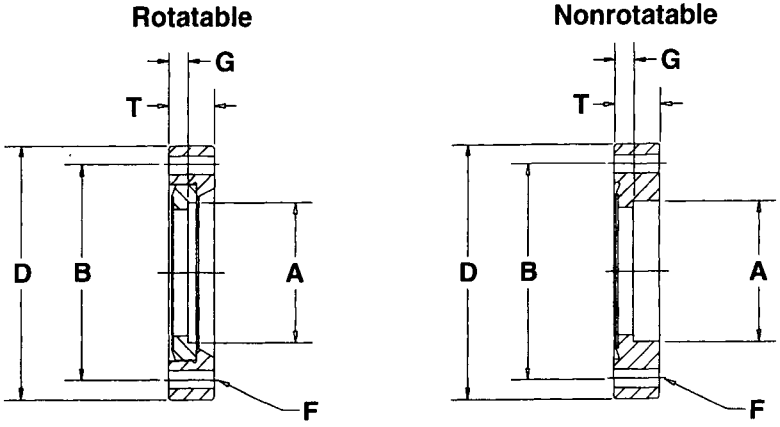
lustrated in Figure 5 and consists of an insert with the knife edge and a concentric bolt ring. The insert is welded to the tube or component, and the bolt ring is rotated to the desired alignment. Flanges with tapped holes are available for use where nuts would be inconvenient. The threads in tapped flanges may be either metric or U.S. system, depending upon origin.

The flanges are commercially available in a variety of stainless steel alloys and aluminum. Aluminum flanges have a hard coating over the flange face and knife edge so that the knife edge is hard enough to penetrate the gasket without damage. The bolts in a CF joint should be the same alloy as the flanges to ensure uniform expansion during bakeout. This helps to prevent leaks. The tensile strength of the bolts must be adequate to withstand the necessary sealing forces. Copper is the most common gasket material and provides the most reliable sealing, but aluminum and nickel alloys are also used. Nickel gaskets require higher sealing forces, so high-tensile strength bolts must be used. Nickel or other hard materials can damage the knife edge, limiting the number of seals that can be made with the flange.

CF flanges are available for use with tubing from $\frac{1}{4}$ inch to 16 inches O.D. Dimensions of CF flanges for up to 8-inch-diameter tubing are given in Table 5. Se-

Table 5

Typical Dimensions of Knife-Edge Flanges Manufactured in the United States. Flanges made outside the United States may have slightly different dimensions and metric threads. Bolt rings for flanges larger than 3.375-inch O.D. may be thicker than the values given. (Reproduced with the permission of HPS Division of MKS Instruments, Inc., Boulder, Colorado.)



Flange Size (in)	A Nominal Tube Diameter	D Flange Outside Diameter	B Bolt Circle Diameter	T Thickness	N Number of Holes	F Tapped Holes Thread	F Through Holes Diameter	G Face to Counter-bore
$1\frac{1}{3}$	.75 (19.0)	1.33 (33.8)	1.06 (26.9)	.30 (7.6)	6	8–32	.172 (4.3)	.18 (4.6)
$2\frac{1}{8}$	1.00 (25.4)	2.12 (53.8)	1.62 (41.1)	.50 (12.7)	4	$\frac{1}{4}$–28	.265 (6.7)	.17 (4.3)
$2\frac{3}{4}$	1.50 (38.1)	2.75 (69.8)	2.31 (58.7)	.50 (12.7)	6	$\frac{1}{4}$–28	.265 (6.7)	.21 (5.3)
$3\frac{3}{8}$	2.00 (50.8)	3.38 (85.8)	2.85 (72.4)	.69 (17.5)	8	$\frac{5}{16}$–24	.332 (8.4)	.22 (5.6)
$4\frac{1}{2}$	2.50 (63.5)	4.50 (114.3)	3.63 (92.2)	.69 (17.5)	8	$\frac{5}{16}$–24	.332 (8.4)	.38 (9.6)
$4\frac{5}{8}$	3.01 (76.2)	4.61 (117.6)	4.03 (102.4)	.81 (20.6)	10	$\frac{5}{16}$–24	.332 (8.4)	.28 (7.1)
6	4.00 (101.6)	6.00 (152.4)	5.13 (130.3)	.78 (19.8)	16	$\frac{5}{16}$–24	.332 (8.4)	.43 (10.9)
8	6.00 (152.4)	8.00 (203.2)	7.13 (181.1)	.88 (22.4)	20	$\frac{5}{16}$–24	.332 (8.4)	.50 (12.7)
10	8.00 (203.2)	10.00 (254.0)	9.13 (231.9)	.97 (24.6)	24	$\frac{5}{16}$–24	.332 (8.4)	.50 (12.7)

Dimensions: $\frac{\text{inch}}{(\text{mm})}$

lection of a CF flange for tubing larger than 10 inches is confusing. For instance, 10-inch-diameter tubing is the maximum size for both 12-inch and 13.25-inch O.D. flanges. Likewise, 14.0- and 14.5-inch O.D. flanges are made for 12-inch-diameter tubing. When specifying a large CF flange, important dimensions and the number of bolt holes should be stated. Most interface dimensions on the larger flanges are made to vacuum industry customary standards, so that flanges from different manufacturers are interchangeable.

4.2.3.4 ASA (ANSI) Flanges

Although in declining use today, ASA (American Standards Association) or ANSI (American National Standards Institute) flanges are one of the oldest flange systems used in building metal vacuum systems in the United States. ASA (ANSI) flanges remain an option on many types of high-vacuum pumps. These flanges were adapted to vacuum service from the 150-lb ASA (ANSI) flanges used with steam or compressed air. They are O-ring sealed, held together with bolts and are available for tubing sizes from 1.5 inches to 12 inches. Typical applications are on industrial vacuum systems.

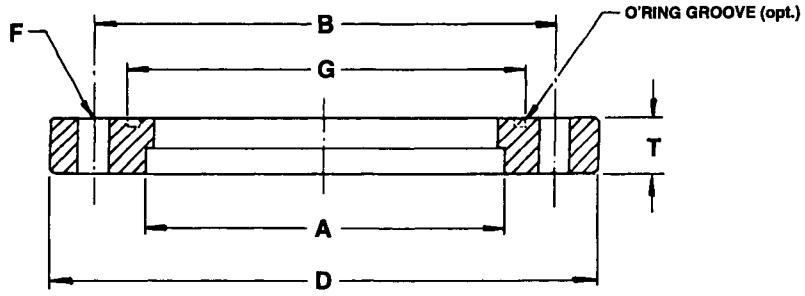
A nonrotatable ASA (ANSI) flange is illustrated in Table 6. Unlike the ISO or CF flanges, these flanges are not sexless because one flange of the pair usually contains an O-ring groove. Occasionally they are sealed with a flat elastomer gasket or with an O-ring in a carrier. To allow alignment of components, rotatable ASA (ANSI) flanges with a bolt ring and a separate insert containing the seal area are made specifically for vacuum service. The rotatable flanges are also available with or without O-ring grooves.

The sizing of an ASA (ANSI) flange can be confusing. Sometimes they are referred to by the maximum bore suggested by the ANSI 150-lb standard, and sometimes by the outside diameter of the flange. An additional complication is that the bore in a flange for vacuum service is frequently larger than the ANSI nominal. For instance, a "6-inch ASA" flange, which has an 11.0-inch O.D., can have a bore as large as 8.0 inches.

Table 6 gives dimensions of ASA (ANSI) flanges used in vacuum service. Flange outside diameter, bolt circle diameter, and number of bolts are standardized. However, the bore, O-ring groove dimensions, the size of the O-ring used, flange thickness, and bolt sizes are not uniform within the industry. When selecting ASA flanges, the user must specify the O.D., the bore, whether it is to be rotatable, whether it is to have an O-ring groove, and if so, the O-ring size or number.

Table 6

Typical Dimensions of ASA (ANSI) Flanges for Vacuum Service.
The ANSI nominal size has been listed for reference. O-ring groove dimensions and flange thickness are not standardized within the vacuum industry. Two-piece rotatable flanges are available, but may be slightly thicker than indicated. (Reproduced with the permission of HPS Division of MKS Instruments, Inc., Boulder, Colorado.)



ANSI Size (in)	A Nominal Tube Diameter[a]	D Outside Diameter	T Thickness	B Bolt Circle Diameter	F Through Hole Diameter[b]	N Number of Holes
1	1.5 (38)	4.25 (108.0)	0.50 (12.7)	3.12 (79.4)	0.625 (15.9)	4
$1\frac{1}{2}$	2.5 (63.5)	5.00 (127.0)	0.50 (12.7)	3.88 (98.4)	0.625 (15.9)	4
2	3.5 (88.9)	6.00 (152.4)	0.50 (12.7)	4.75 (120.7)	0.75 (19.0)	4
3	4.0 (102)	7.50 (190.5)	0.50 (12.7)	6.00 (152.4)	0.75 (19.0)	4
4	6.0 (152)	9.00 (228.6)	0.50 (12.7)	7.50 (190.5)	0.75 (19.1)	8
6	8.0 (203)	11.00 (279.4)	0.75 (19.0)	9.50 (241.3)	0.812 (20.6)	8
10	12 (305)	16.00 (406.4)	0.75 (19.0)	14.25 (362.0)	0.812 (20.6)	12

[a] Given as approximate maximum.
[b] Smaller bolts often used.

Dimensions: $\frac{\text{inch}}{(\text{mm})}$

4.2.3.5 JIS Flanges

An O-ring sealed and bolted flange that might be encountered, particularly in equipment made in Japan, is the JIS (Japanese Industry Standard) flange. These flanges are made in nominal sizes from 10 to 1000 mm, for tubing from 17.3 to 1016 mm in diameter. In a typical joint, one flange contains an O-ring groove and the other does not. However, there are adaptors available that allow two identical flanges to be joined. If both flanges are without O-ring grooves, the adaptor has O-ring grooves on both sides. If both flanges contain O-ring grooves, then the adaptor is smooth on both sides.

4.2.3.6 Large Wire-Sealed Flanges

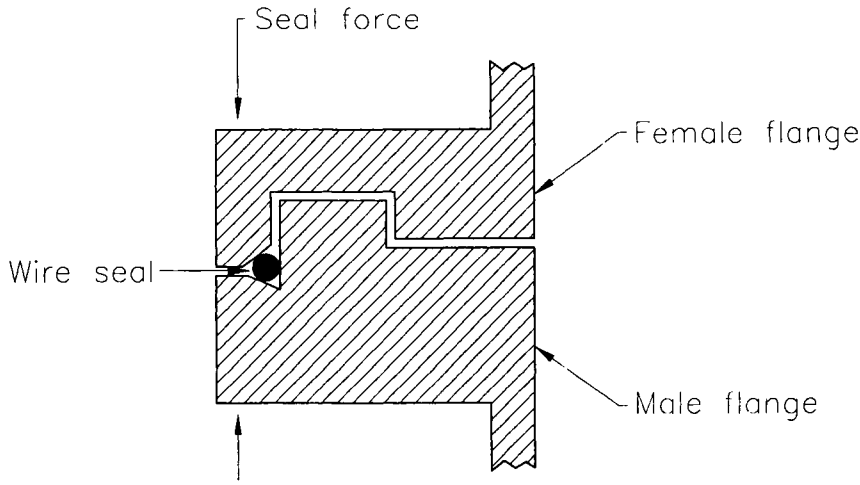
Metal-wire-sealed flanges are commercially available in 12- to 27-inch outside diameters, for tubing from 10 to 24 inches in diameter. Typical applications are metal bell jars or large custom chambers. Figure 6 illustrates the sealing principle of these flanges. The seal is a metal wire, usually copper, contained in grooves on the mating flanges. The flanges are sexed, with a tongue and groove to provide centering and to limit compression of the seal. Some flanges using this sealing principle are held together with bolts, while others use special clamps that provide sealing force directly over the seal.

4.2.3.7 Rectangular Flanges

In some applications, it is advantageous to use a non-circular flange. One rectangular flange for which standards exist is the MESA (Modular Equipment Standards Architecture) or SEMI (Semiconductor Equipment and Materials International) flange [4]. This flange was designed for the transfer of wafers through load-locks in semiconductor fabrication equipment. Valves with these flanges are commercially available.

Fig. 6.

Sealing principle used on large wire-sealed flanges.



The tongue and groove arrangement centers the flanges. The seal force can be provided by bolts or special claw clamps.

Table 7 illustrates this flange and lists some key dimensions. SEMI flanges are O-ring sealed, and a flange pair is held together with the claw clamps used for the larger ISO flanges described earlier. The SEMI standard specifies the seal region on the flange face and also that the smooth flange without the O-ring groove is to face the transport module side. The flange containing the O-ring is specified to face toward a cassette or process module. A dowel pin either side of the port assist in aligning the flanges. Two sizes of the flange are specified in the SEMI standard, one for 200-mm wafers and one for 300-mm wafers. The flange design has been adapted to other sizes by component manufacturers.

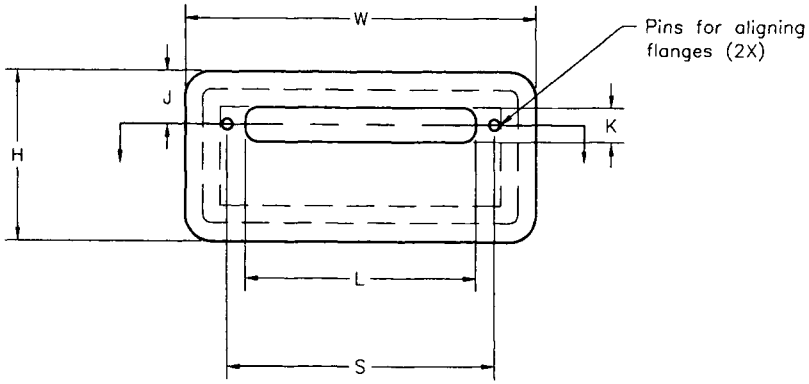
4.2.4

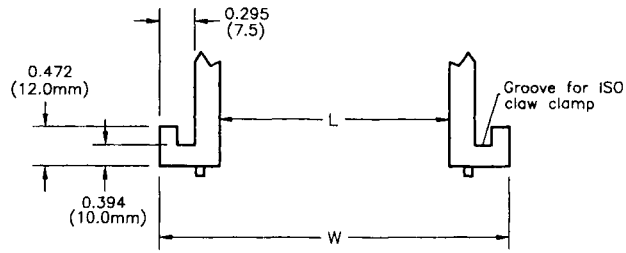
COMPONENTS WITH FLANGES ATTACHED

Components such as valves, elbows, Tees, and crosses are commercially available in many of the flanging systems described above. This allows quick and easy con-

Table 7

Dimensions of SEMI (Semiconductor Equipment and Materials International) or MESA (modular equipment standards architecture) Rectangular Flanges for Passage of 200-mm and 300-mm Wafers. The flanges are O-ring sealed, and a flange without O-ring groove is illustrated. Claw clamps as used on ISO O-ring sealed flanges provide the sealing force.





Intended Wafer Size	Outside Width (W)	Outside Height (H)	Minimum Port Width (L)	Port Height (K)	Locating Pin Spacing (S)	Locating Pin Offset (J)
8 in	13.39	6.38	8.74	1.26	11.81	1.97
(200 mm)	(340)	(162)	(222)	(32)	(300)	(50)
12 in	17.32	6.38	12.68	1.26	15.75	1.97
(300 mm)	(440)	(162)	(322)	(32)	(400)	(50)

Dimensions: $\frac{\text{inch}}{(\text{mm})}$

struction or modification of a vacuum system. In this section, dimensions are given for identification of components and to indicate what is available.

4.2.4.1 Components with ISO-Type Flanges

The ISO standard ISO 9803:1993(E), "Vacuum Technology — Pipeline Fittings — Mounting Dimensions," provides standards for the flange face to center line dimension on elbows, Tees, and crosses with ISO flanges. The dimension for elbows is frequently used on right-angle valves. When the standards are followed, the flanges and components form a very modular system. One component can be easily exchanged with another without disturbing the rest of the system. In addition, interchangeability between manufacturers is assured.

Table 8 gives the flange face to centerline dimensions, sometimes referred to as the "modulus," for KF-flanged components. Dimensions of MF-flanged components are given in Table 9. In all cases the component is referenced by the nominal size of the flange used. The spacing between opposed flanges, dimension "L" in either table, on a cross or Tee is twice the modulus. The components illustrated have the same size flange on all ports. The ISO standard applies to components made with either bolted or clamped-type flanges. However, the larger components are usually made with clamped flanges (ISO-MF), since these flanges are infinitely rotatable.

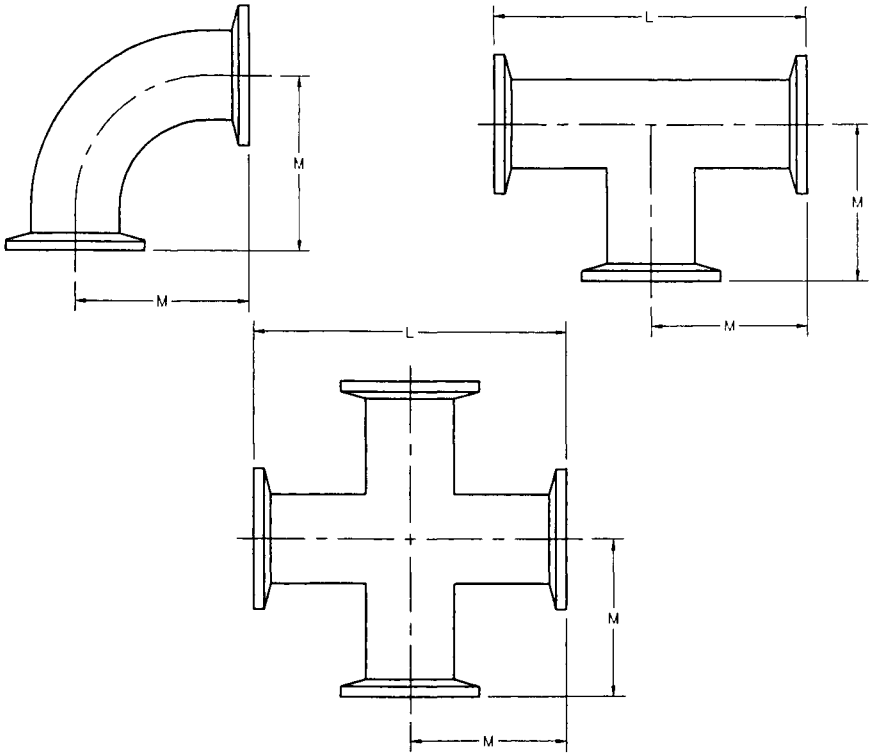
The customary tubing size used in the United States for the component body is listed for reference in Tables 8 and 9. This dimension is not specified in the ISO standard, and may depend on the country of origin. However, it is usually unimportant, and components are fully interchangeable if the modulus is maintained.

Some manufacturers offer components with nominal 320 and larger flanges. However, at the present time, the ISO standard does not address the dimensions of components in sizes larger than 250. Also, the ISO standard does not specify any value for elbow, Tee, and cross dimensions in the 50, 80, and 200 nominal sizes. As explained earlier in the ISO flange section, flanges of these nominal bores are not from the "first preference series." The values given here in these three sizes are used by most U.S. and European manufacturers but may not represent products from all manufacturers.

Many other components such as crosses with five or six arms, Tees with a mixture of flange sizes, as well as adaptors between flange sizes and other flange systems are available. None of these components are described or dimensioned by ISO standards, but they increase the usefulness of the system.

Table 8

Dimensions of Elbows, Tees, and Four-Way Crosses with KF Flanges. Flanges on all ports are the same size. Dimensions in the NW 16, 25, and 40 sizes are set by ISO (International Standards Organization). The values listed for the NW 50 size are used by both U.S. and European manufacturers. At the present time, ISO does not specify these dimensions for the NW 50 size.



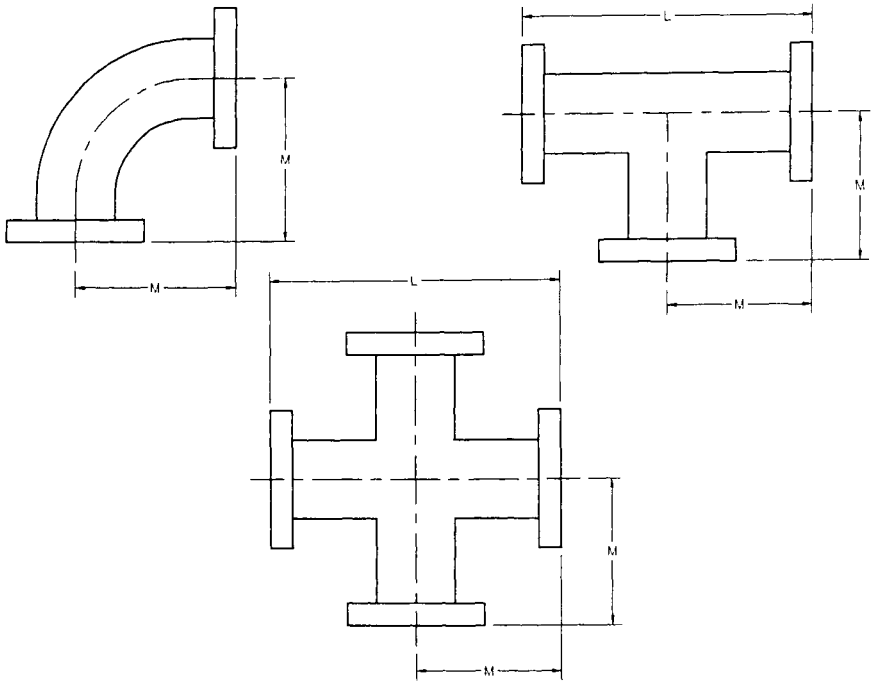
ISO Flange Size[a] (mm)	Customary U.S. Tubing Diameter (reference) (in)	*M*	*L*
NW 16	0.75	1.57	3.15
		(40)	(80)
NW 25	1.00	1.97	3.94
		(50)	(100)
NW 40	1.50	2.56	5.12
		(65)	(130)
NW 50	2.00	2.76	5.51
		(70)	(140)

[a]NW is retained by convention. It does not specify the flange type, but is from the German *Nennweite*, meaning "nominal diameter."

Dimensions: $\frac{\text{inch}}{(\text{mm})}$

Table 9

Dimensions of Elbows, Tees, and Four-Way Crosses with Larger ISO (International Standards Organization) O-Ring-Sealed Type Flanges. Flanges on all ports are the same size. The dimensions are set by ISO except in the NW 80 and NW 200 sizes. The values given for the NW 80 and NW 200 sizes are typical, and may not represent products from all manufacturers.



ISO Flange Size[a] (mm)	Customary U.S. Tubing Diameter (reference) (in)	*M*	*L*
NW 63	2.50	3.46	6.93
		(88)	(176)
NW 80	3.00	3.86	7.72
		(98)	(196)
NW 100	4.00	4.25	8.50
		(108)	(216)
NW 160	6.00	5.43	10.87
		(138)	(276)
NW 200	8.00	6.61	13.23
		(168)	(336)
NW 250	10.00	8.19	16.38
		(208)	(416)

[a]NW is retained by convention. It does not specify the flange type, but is from the German *Nennweite*, meaning "nominal diameter."

Dimensions: $\frac{\text{inch}}{(\text{mm})}$

4.2.4.2 Components with Knife-Edge Seal Flanges

Unlike ISO-flanged components, there is no standard for the dimensions of components with knife-edge flanges. However, most U.S. manufacturers of CF-flanged components make parts to the same customary dimensions so that there is interchangeability between manufacturers.

These dimensions are given in Table 10 for the most common flange sizes. The table is organized by the outside diameter of the flange and the flanges on all ports are the same O.D. To ensure rotational alignment, the indicated combination of rotatable and nonrotable flanges is necessary. The long dimension of a Tee or cross is again twice the modulus. Catalogues from CF flange manufacturers list additional components, such as crosses with five and six arms, and Tees with different-size flanges.

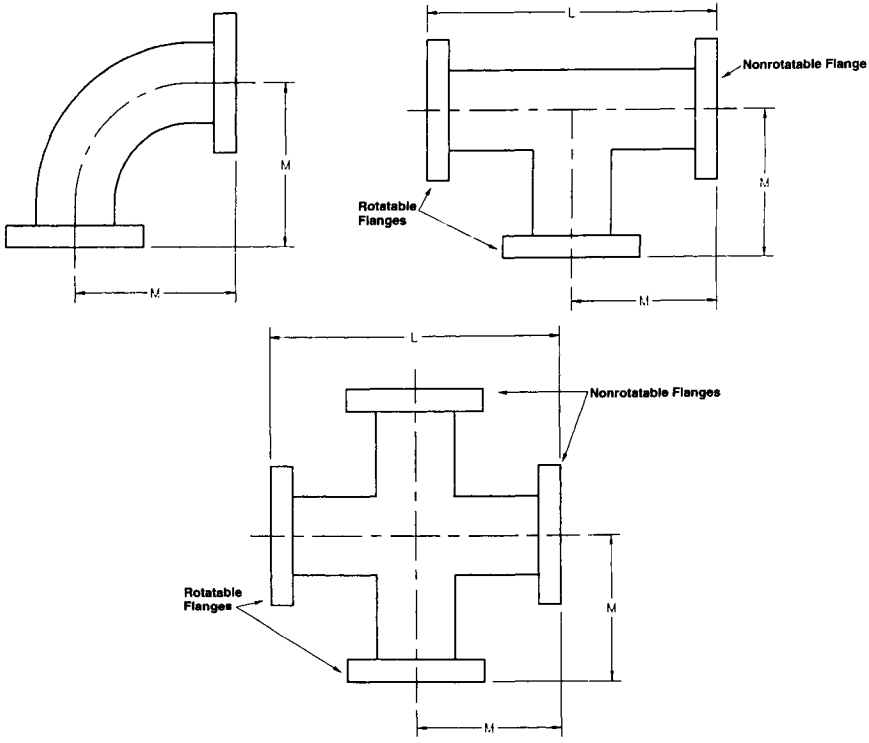
4.2.4.3 Components with ASA-Type Flanges

Valves, elbows, Tees, and crosses with ASA (ANSI) flanges are commercially available. However, there is little standardization of flange face to centerline dimension or tubing size for the component, especially in the larger sizes. For this reason, dimensions of components are not tabulated here. When purchasing components, including valves, for use with existing ones or components from multiple manufacturers, users should verify that critical dimensions are identical. Also, the position and size of O-ring grooves must be considered and specified. Adaptors are also available to other flange systems.

TRADEMARKS

Table 10

Customary Dimensions Used by U.S. Manufacturers of Elbows, Tees, and Four-Way Crosses with Metal-Sealed Knife-Edge Flanges. Flanges on all ports have the same outside diameter. The combination of rotatable and nonrotatable flanges is used to provide rotational alignment.



Nominal Flange Size (in)	Nominal Tubing Diameter	*M*	*L*
1.33	0.75	1.50	3.00
	(19.0)	(38.1)	(76.2)
2.12	1.00	2.06	4.12
	(25.4)	(52.3)	(104.6)
2.75	1.50	2.46	4.92
	(38.1)	(62.5)	(125.0)
3.37	2.00	3.21	6.42
	(50.8)	(81.5)	(166.0)
4.5	2.50	4.13	8.26
	(63.5)	(105.0)	(210.0)
4.62	3.00	4.94	9.88
	(76.2)	(125.5)	(251.0)
6.0	4.00	5.31	10.62
	(102.0)	(135.0)	(270.0)
8.0	6.00	6.56	13.12
	(152.0)	(167.0)	(333.0)
10.0	8.00	8.00	16.00
	(203.0)	(203.0)	(406.0)

Dimensions: $\frac{\text{inch}}{(\text{mm})}$

REFERENCES

1. C. H. Rosendahl, *Sheet Metal Industries* (February 1970), pp. 93–109.
2. R. D. Thomas, Jr., *Welding Research Supplement to The Welding Journal* (December 1984), pp. 335s–368s.
3. N. Suutala, T. Takalo, and T. Moisio, *Metallurgical Transactions A*, **11A** (1980) 717–725.
4. SEMI standards E21-94 and E21.1-92, available from SEMI North America, 805 East Middlefield Road, Mountain View, CA 94043, telephone (415) 964-5111, semihq@semi.org.

CHAPTER 4.3

Magnetic-Fluid-Sealed Rotary Motion Feedthroughs

Walter Helgeland
Rigaku/USA, Inc.

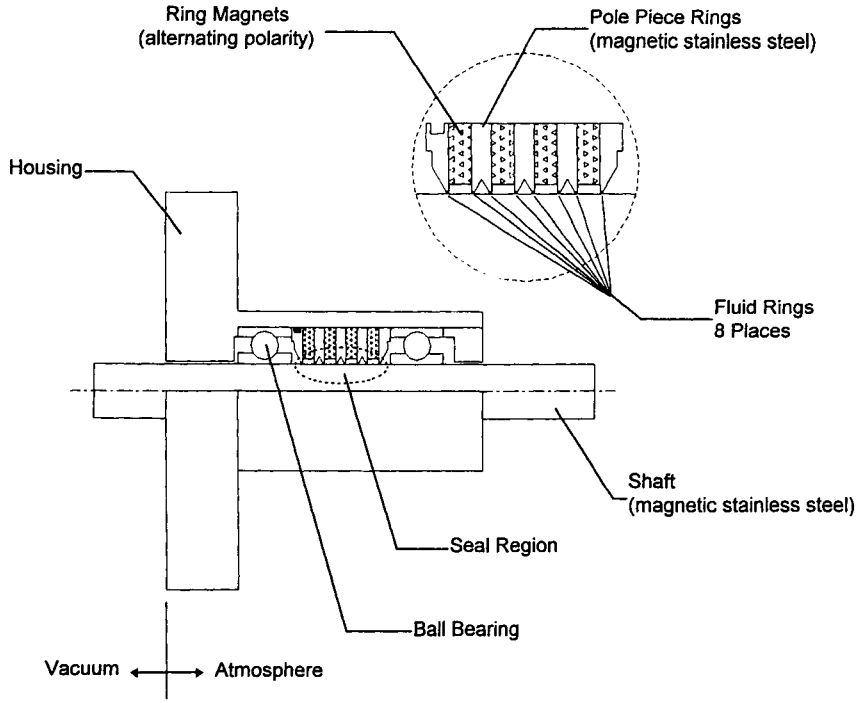
4.3.1 BASIC SEALING PRINCIPLE

In this class of rotary feedthroughs, dynamic sealing is accomplished by a hermetic liquid seal that permits free rotation of the shaft. Figure 1 illustrates the arrangement of the major elements in a typical device.

Magnetic fluids are stable colloids comprising a base liquid, ferromagnetic particles, and a dispersing agent that suspends the particles in the base liquid. The ferromagnetic particles interact directly with an external magnetic field. Because of the strong coupling to the magnetic particles via the dispersing agent, the base liquid interacts indirectly with the external magnetic field. Hence magnetic fluids can be pushed, pulled, and shaped by magnetic fields. In feedthroughs, narrow rings of fluid form liquid barriers filling the annular spaces between a rotating shaft and the tips of a stationary pole piece. Radially, the fluid rings are bounded by the shaft and the pole piece tips. Axially, the rings are free surfaces at gas–liquid interfaces, restrained only by magnetic forces. Magnetic flux density at the pole tips is very large. Hence, any axial displacement of a liquid ring from the neighborhood of the pole tip results in a force that resists the displacement. The

ISBN 0-12-325065-7
$25.00

Fig. 1.



Typical Magnetic Fluid Sealed Feedthrough

isolated volumes between adjacent rings are important in the functioning of the device.

Maximum sustainable pressure difference across a single liquid ring is less than 1 atmosphere. Consequently, practical devices require a series of separate rings with small, isolated gas volumes between each pair of rings. The total pressure difference (typically 1 atmosphere) is divided over several stages.

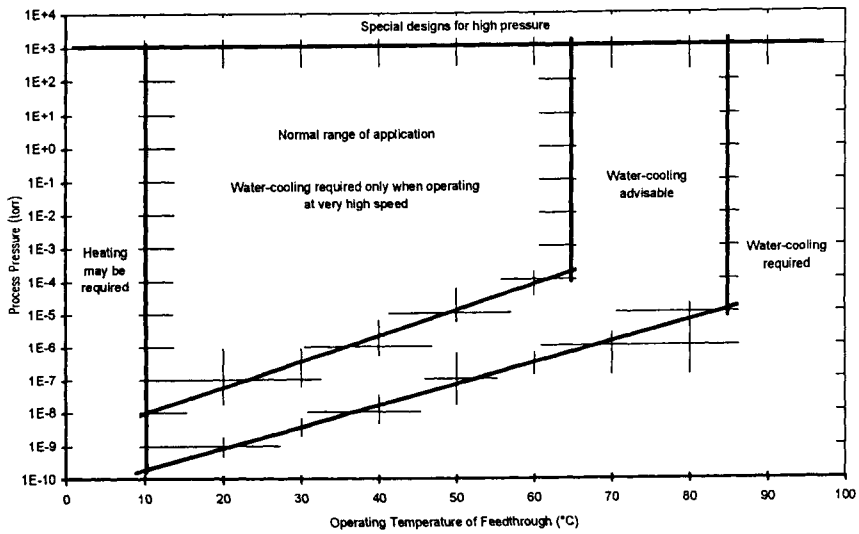
Dimensions and locations of shaft and pole tips are critically important for reliable operation. A complete system of shaft, pole piece, bearings, housing, and other parts provides the necessary precision in a single package that is easily integrated into a vacuum system.

4.3.2

APPLICATION FACTORS

Magnetic fluid feedthroughs are appropriate for many applications, but it is important to consider two groups of application factors when selecting a specific model

Fig. 2.



Operating Region Guidelines

for a particular application. *Environmental factors* always include process operating pressure and ambient temperature, and may include other process-related factors as well. *Mechanical factors* include speed, loads (radial and thrust), duty cycle, and bearing ratings.

Figure 2 shows the recommended operating guidelines for process pressure and chamber temperature. Note that water cooling (or heating) extends the useful temperature range. Note also that units can be designed for use at high pressure.

Temperature limits are based on the properties of magnetic fluids and other construction materials. Although rarely encountered, low temperatures can increase fluid viscosity sufficiently to require unacceptably large driving torque. High temperatures cause deleterious chemical and physical changes in fluids and other materials.

Pressure limits are determined by magnetic design and fluid properties. The lower limit is set by the vapor pressure of the base liquid (temperature dependent) in combination with the process tolerance for fluid vapors. Higher pressure ratings can be achieved by adding more stages. Standard products offer maximum pressure capacity of as much as 2.4 kg/cm^2.

In cases of high-speed operation, the self-heating of the viscous magnetic fluid becomes a significant cause of elevated temperature. Surface speeds up to 7 m/sec are possible without water cooling (DN = diameter $\times$ rpm = 120,000 mm-rpm). With water cooling, surface speed can be increased to about three times this value.

Except for light-duty applications, mechanical operating factors have great impact on the performance and lifetime of feedthroughs. Wholly apart from vacuum

considerations, bearings and shafts must support each application's requirements for speed, load, duty cycle, torque transmission, end play, and so on. Thorough consideration of these factors is beyond the scope of this discussion, but a few key points can be mentioned here.

Shaft sizes must be large enough to transmit the required torque from motor to load over the full range of anticipated acceleration and speed. Possibilities of jamming should be taken into account.

Bearings must be rated to handle the radial and axial loads to be experienced in service. Significant tradeoffs exist between radial load, axial load, operating speed, duty cycle, operating temperature, lubrication, and bearing lifetime. Thus, any "load rating" of a feedthrough as a single number cannot be an adequate characterization over a realistic range of applications. Users should be aware that bearing lifetime varies strongly with load and speed, according to the following equation:

$$L_{10} = A \frac{(C/P)^3}{N} \tag{1}$$

where L_{10} = bearing life expectancy (90% confidence)
A = lubrication factor (normally less than 1)
N = operating speed, RPM
C = "dynamic load rating" of the bearing (depends on bearing size)
P = "dynamic equivalent load" (combination of radial and axial loads)

The main point to note from this equation is that bearing life varies linearly with rpm and exponentially (*as the third power!*) with load and rating. Loading the bearings beyond their rating leads to *very* rapid failure.

As a practical matter, selection of appropriate models can often be guided by previous experience in similar applications. Manufacturers of these products should be consulted for recommendations and application assistance.

4.3.3 IMPACT OF FEEDTHROUGH ON PROCESS

Feedthroughs can make undesirable secondary contributions to the vacuum process environment because of the presence of magnetic fluids, bearings, and magnetic fields. Users should consider whether any of these contributions might be significant in a particular application.

Magnetic fluids are the most mysterious of these elements for most users; their properties are discussed more fully in Section 4.2.5. With respect to their possible impact on the process environment, the main considerations are vapor pressure

and chemical composition. Each feedthrough contains a small free surface of fluid in contact with the vacuum space. The area of this surface is approximately given by the relationship:

$$A \approx 0.3 * D \tag{2}$$

where A = area of free surface, mm^2
D = shaft diameter, mm

Fluid evaporates from this free surface at a rate proportional to vapor pressure (10^{-10} to 10^{-12} torr at 20°C for typical fluids). Not much can be done to reduce the area of the free surface, but it is possible to select fluids with the lowest vapor pressures. However, fluids with the lowest vapor pressure typically have very high viscosity, leading to an increased torque requirement to operate the feedthrough. At high speeds, this leads to increased self-heating, raising the local temperature (and the vapor pressure) of the fluid. The most common fluids are based on hydrocarbons with vapor pressure in the neighborhood of 10^{-10} torr (20°C). Fluids based on perfluorinated polyether (PFPE) materials can also be obtained (vapor pressure approximately 10^{-12} torr).

Bearings can contribute particulates to the process because most feedthroughs employ a "straddle" design in which one ball bearing is placed on each side of the magnetic seal system. The vacuum-side bearing generates particulates that depend on load, speed, and lubrication. Grease lubricants with low vapor pressure (PFPE materials) are commonly used. When exceptionally low particulate generation is required, dry-lubricated bearings can be specified, but load-carrying capacity is substantially reduced and cost is increased. In some cases (e.g., very high speeds) ceramic bearings may be justified (very expensive). Sometimes a "cantilever" bearing arrangement is employed in which both bearings are located on the atmosphere side of the seal. This isolates the bearings and lubricants from the process, but also increases the size and cost of the feedthrough.

Most of the *magnetic energy* is confined to the interior regions of feedthroughs. External fields are also present, and, in general, are undesirable. Some feedthroughs employ a magnetically balanced design in which multiple magnets are used in an opposed-field configuration. This results in substantial (but not complete) nulling of the undesired external field. In applications where external magnetic fields are a significant issue, users should consult with manufacturers for specific guidance before final selection of a feedthrough.

4.3.4 IMPACT OF PROCESS ON FEEDTHROUGH

Most operating environments (except where high temperatures are present) are benign, with no adverse impact on these devices. It is important to be aware of the

unusual cases in which the environment threatens the feedthrough, especially as a result of foreign matter finding its way into the bearings or the fluid region. Four such cases are discussed here.

Reactive process gases wreak havoc with hydrocarbon-based magnetic fluids, causing the feedthrough to leak. In some cases, reactive gases also lead to gummy deposits on bearings, causing mechanical failure. Cantilever bearing arrangements (both bearings on the atmosphere side of the seal) protect the bearings, but not the sealing fluid.

Particulate matter generated by the process can find its way into the seal region. Small amounts of this can become jammed into the narrow magnetic gap between the shaft and pole piece, causing the shaft to vibrate as it rotates. Labyrinth or lip seals are sometimes used to reduce the ingress of material. Lip seals used in this way necessarily result in small, isolated volumes located between the lip seal and the fluid seals. In some processes, the resulting virtual leaks can be tolerated.

Solvents used during cleaning of the system can destabilize the colloid system of the magnetic fluid, leading to instantaneous major leakage. Cleaning solvents must be used with care, limiting the quantities employed in the neighborhood of the feedthrough. When the feedthrough is located at the bottom of the system, with vertical shaft orientation, the chances of flooding are greatly increased.

Solvents sprayed on the atmosphere end of the feedthrough while hunting for gross leaks have damaged many feedthroughs. This leak-testing technique must not be used on feedthroughs.

4.3.5 MATERIALS CONSIDERATIONS

Most applications are well served by feedthroughs of standard design. However, users should be aware of the materials used in constructing feedthroughs to be sure they are compatible with the system environment in which the feedthrough will operate. Table 1 lists the materials commonly employed in these feedthroughs. Manufacturers can provide more specific information.

Typical characteristics of magnetic fluids are summarized in Table 2. Since these fluids are proprietary materials, this table must be understood as a very broad characterization.

For most applications, hydrocarbon-based fluids are entirely satisfactory. PFPE fluids offer lower vapor pressures, but present significantly higher drag and may require more sealing stages (larger feedthroughs).

Table 1

Materials Used in Feedthroughs

Material	Where Used	Comments
Stainless steel, nonmagnetic	Housing, flange	303 and 304 are commonly used.
Stainless steel, magnetic	Shaft, pole piece	17-4 PH (SUS 630) and 416 are both widely used.
Bearing steels	Balls, rings	1. SAE 52100 (SUJ 2) is common. 2. 440 (SUS 440 C) used in all-stainless bearings.
	Retainers (ball cages)	Usually pressed from carbon strip steel. Stainless steel available on special order.
Magnet alloys	Magnets	Magnets are normally isolated from the vacuum space by some of the fluid rings, so are not usually an issue with respect to process compatibility. Commonly used rare-earth magnet alloys are SmCo and NdFeB. AlNiCo magnets are also used.
Elastomers	Static O-ring seals	Viton® O-rings are used to provide static vacuum sealing between housing and pole piece.
Hydrocarbons, fluorocarbons	Magnetic fluids	Formulations are proprietary. Major constituent is the base oil, which is typically alkylnaphthalene, Krytox® or Fomblin®. See Table 2.
Adhesives	Epoxies and thread-locking compounds	Epoxies, if used, are normally isolated from the vacuum space by some of the fluid rings, so are not usually an issue with respect to process compatibility. If threaded retainers are present on the vacuum face, it is possible that some locking compounds have been used. Consult manufacturer for details in such cases.

Table 2

Typical Magnetic Fluid Material Properties

Base Liquid	Hydrocarbon	Fluorocarbon	
Vapor pressure, torr			
at 20°C	6×10^{-10}	2×10^{-11}	Fluorocarbon fluids (Krytox®,
at 100°C	2×10^{-5}	5×10^{-7}	Fomblin®) have lower vapor pressure.
Viscosity, centipoise			
at 20°C	300	5,000	Hydrocarbon fluids have lower viscosity, lower drag.
Magnetization, gauss	450	200	Hydrocarbon fluids are "magnetically stronger," can provide higher-pressure capacity with fewer stages.

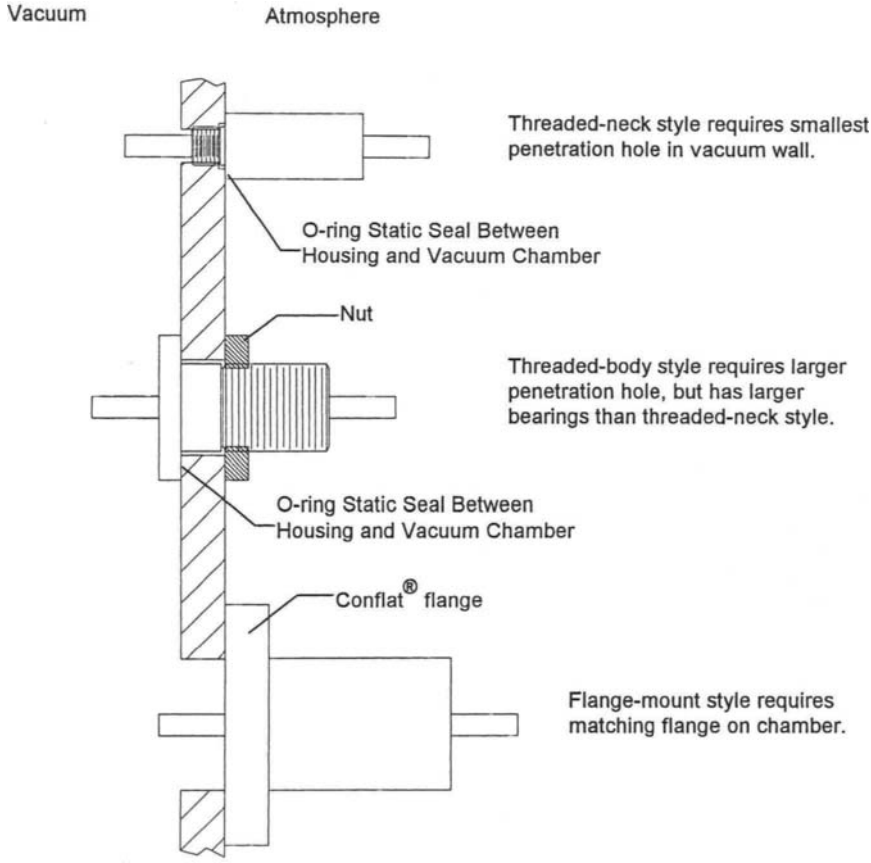
4.3.6

APPLICATION EXAMPLES

4.3.6.1 Light-Duty Feedthroughs

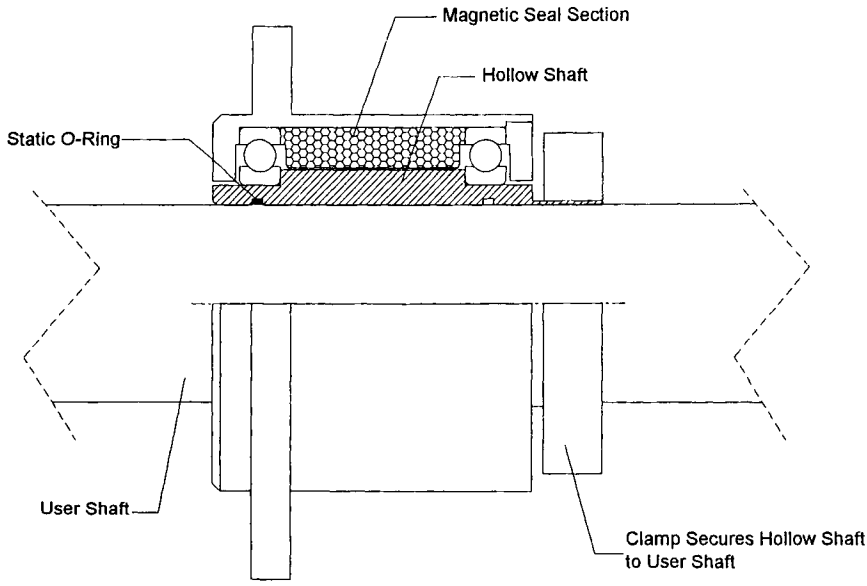
Examples of this category include shutters for evaporation and sputtering deposition sources, positioning systems for multipocket E-beam evaporation sources, sample positioners, and chopper wheels. Shaft diameters are typically 6 mm to 10 mm. Figure 3 illustrates three mounting styles for popular products with 6.35-mm shaft diameter.

Fig. 3.



Light-Duty Feedthroughs, 6.35 mm Shaft Diameter

Fig. 4.



Hollow Shaft Feedthrough

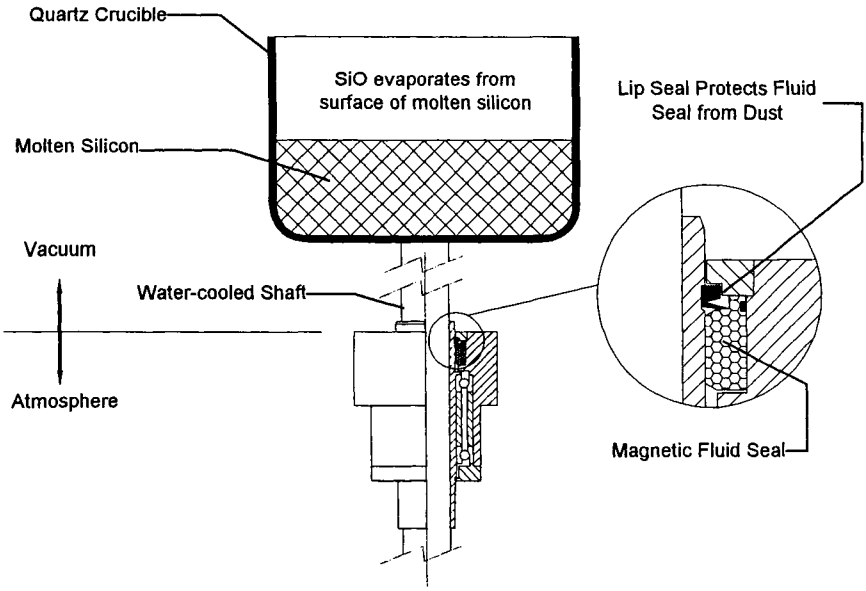
4.3.6.2 Hollow Shaft Seals

In applications where shafts are large, complex, proprietary, or subject to design revision, hollow-shaft feedthroughs can often be used to advantage. Hollow-shaft feedthroughs provide a complete bearing and seal package in which the shaft is a tube or sleeve with closely controlled inside diameter. The user joins this package to the shaft to form a complete assembly suitable for the application. A static O-ring seal between the hollow shaft I.D. and the user shaft O.D. is required. Some positive means of driving the hollow shaft via the user shaft is required to ensure that the user shaft does not move relative to the hollow shaft, an event that would damage the static O-ring seal. Figure 4 illustrates the typical case.

4.3.6.3 Crucible Rotation for Czochralski Silicon Crystal Puller

Figure 5 shows how a magnetic fluid feedthrough is used in a process that generates a great deal of finely divided dust (SiO) that would rapidly degrade bearings and seal. A cantilever bearing design separates bearings and process, and a lip seal

Fig. 5.



Silicon Crystal Growth System

isolates the magnetic fluid seal from the process, increasing the number of process cycles that can be run before the seal must be serviced. This is a custom-designed product.

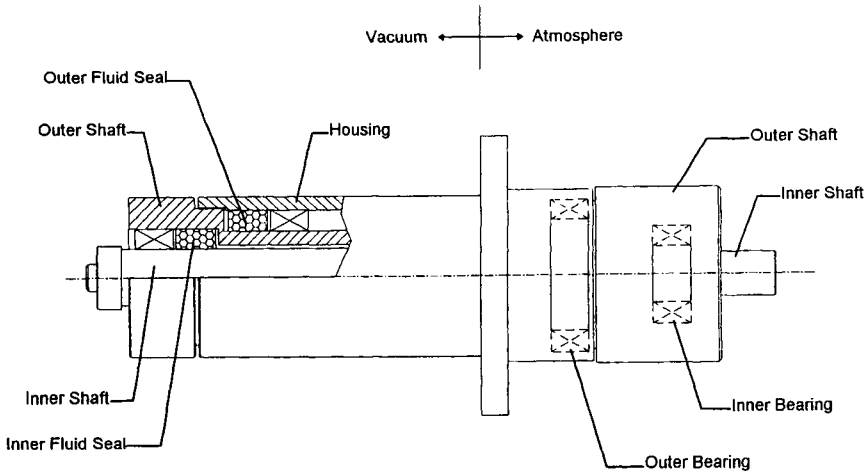
4.3.6.4 Coaxial Feedthrough for Robotic Arm

Robotic arms are widely used to handle silicon wafers and flat-panel displays in cluster tools and other semiconductor processing equipment. Two independent motions are required. One method of achieving this is to nest one feedthrough inside another, as illustrated in Figure 6. This is a custom-designed product.

4.3.7 COMPARISON TO OTHER TYPES OF FEEDTHROUGHS

Table 3 compares magnetic fluid feedthroughs with three other types. Magnetic fluid feedthroughs provide very high reliability and require no maintenance be-

Fig. 6.



Coaxial Feedthrough for Robotic Arm

cause the liquid sealing element is not subject to any inherent wearout mechanism. Shafts are continuous from atmosphere to vacuum, resulting in high torque and zero backlash. Auxiliary devices (differential pumping, guard vacuums) are not normally required. Because magnetic fluids are viscous, the feedthroughs exhibit noticeable drag, a factor that becomes significant at high speeds or with low-powered drive systems.

Table 3

Comparison of Four Types of Feedthroughs

	Magnetic Fluid Seals	Lip Seals and O-Ring Seals	Bellows Seals with Bent Shaft	Magnetically Coupled Seals
Reliability	High	Low	Moderate	High
Maintenance	None	High	Moderate	None
Torque Capacity	High	High	Low	Slippage is possible
Backlash	None	None	Yes	May be large
Guard Vacuum	Required only for hazards	Frequently required	No	No
Bakeout to 150°C	Only if water-cooled	OK for Viton® O-rings	Yes	Yes
Internal Drag	Moderate to high	Low	Low	Low

CHAPTER 4.4

Viewports

Chuck Kraft
Larson Electronic Glass

4.4.1 MATERIALS

The three most common viewport materials used in high- and ultra-high-vacuum systems are glass, quartz, and sapphire. The choice of which to use often depends on cost, transmission characteristics, bakeout temperature, style, and materials used for the viewport seal.

Glass viewports, the most widely used and least expensive of the three choices, are available in a wide variety of diameters. The transmission is approximately 90% between 0.32 and 2.6 microns. Glass viewports are bakeable to 400° Celsius (centigrade). Zero-length viewports, where the surface of the glass is below the surface of the flange, are Kovar®[1] sealed. Also available are nonmagnetic viewports with the glass sealed to the end of a stainless steel tube. The viewport is normally mounted with the tube extending out from the vacuum chamber. In diameters up to 4 inches, the tube can be inverted to place the glass close to an object within the chamber. Tubular Kovar® viewports can be made the same way.

Quartz viewports have a transmission of approximately 90% in the range of 0.3 to 2.5 microns and 0.2 to 2.0 microns for UV-grade materials. They are available in both zero-length and standard-length. A standard-length quartz viewport has the quartz disk mounted on a tube. Quartz viewports are sealed to stainless steel

[1] Kovar® is a registered trademark of Carpenter Technology. Kovar is an iron, nickel, and cobalt alloy used for glass sealing.

ISBN 0-12-325065-7
$25.00

frames with a low-temperature brazing alloy. The bakeout temperature is limited to 200° Celsius.

Sapphire viewports have the broadest transmission range. The transmission is above 80% at 0.3 to 4.0 microns and 0.25 to 4.5 microns for UV-grade sapphire. These viewports are sealed with a high-temperature braze material and can withstand bakeout temperatures as high as 450° Celsius. The frame is Kovar®.

4.4.2 MOUNTING SYSTEMS AND PRECAUTIONS

Viewports of all three types are available with a variety of mounting flanges, and, with the exception of the glass and quartz zero-length types, without flanges. In use, it is important not to induce stress into the seal, which could cause cracking or breakage. The two types of stress are mechanical and thermal. Mechanical stress is created from uneven pressure applied to the viewport frame. Flange bolts should be tightened in a cross pattern, and foreign objects should not be allowed to make contact with the frame or the glass. Thermal stress is the result of uneven heat being applied to the frame and/or the glass during welding, bakeout, or in actual use on the system. Viewports are designed for vacuum use. The zero-length glass viewports should be mounted in the normal configuration. Nonmagnetic glass viewports up to 4 inches in diameter, and quartz and sapphire viewports can be mounted with vacuum on either side. Precautions should be taken to avoid pressure differentials greater than 1 atmosphere. In some applications, safety pressure valves or shielding should be considered to protect operators and equipment.

CHAPTER 4.5

Construction Materials

Janda K. Panitz
Sandia National Laboratories

4.5.1 PROPERTIES DEFINING MATERIAL PERFORMANCE

The gas load in a vacuum system critically depends on the materials used to build the system and how they are prepared and maintained. Materials add to gas load in two ways: (1) evaporation, sorption and outgassing, and (2) diffusion and permeation

4.5.1.1 Evaporation, Adsorption, Absorption, and Outgassing

EVAPORATION

Vapor pressure can be defined as the pressure at which a material's evaporation or sublimation rate equals its condensation rate. If you pump a system below a material's vapor pressure, the material sublimes or evaporates. As a general rule, vapor pressure increases with temperature. At a given temperature and pressure, the sublimation or evaporation rate can be predicted from vapor pressure using thermodynamic theory [1].

It is desirable to use low-vapor-pressure materials, e.g., stainless steel, glass, and polyimide, as opposed to high-vapor-pressure materials, e.g., zinc, ice cubes, and camphor, at practical operating temperatures. Selected information is pre-

ISBN 0-12-325065-7
$25.00

sented in Tables 1 to 4. These data can be used for guidance in selecting or rejecting metals, alloys, and selected compounds.

ADSORPTION, ABSORPTION, AND OUTGASSING

Adsorption can be defined as the condensation of molecularly thin films on a solid surface. Absorption is condensation that extends into the near-surface region of a solid. Outgassing occurs when high-vapor-pressure-adsorbed or -absorbed species evaporate as the vacuum system is pumped down and/or heated. Sometimes, high-vapor-pressure solid, liquid, or gas films chemically bond to a surface ("chemisorption"). If this occurs, the pumpdown rate to the high-vacuum regime and below can be substantially slowed as the film is slowly wrested from its substrate.

Water, which readily chemically bonds to many metal and ceramic surfaces when vacuum system components are exposed to lab air, is a particularly ubiquitous and persistent absorbed species. Water is an especially troublesome species when it outgasses because low partial pressures of water strongly interfere with many processes and measurements performed in practical vacuum systems. The problem can be reduced by venting a system with dry nitrogen to blanket surfaces with a gas species that adheres less tenaciously than chemisorbed water during pumpdown.

Using smooth, nonporous, relatively chemically inert materials that are easy to clean and adapting appropriate cleaning techniques will reduce the gas load added by outgassing. A material's performance depends not only on its intrinsic properties but also on on how well it is cleaned. Section 4.9, "Preparation and Cleaning of Vacuum Systems," is critically linked to this section on vacuum material selection. Depending on whether high, ultra-high, and extra-ultra-high vacuums are desired, progressively aggressive cleaning and system maintenance procedures can be used to remove high-vapor-pressure contaminants. (Tools, including gloved hands, coming into contact with vacuum materials must also be clean.)

4.5.1.2 Solubility, Permeation, and Diffusion

Solubility, diffusion, and permeation relate to a material's chemical properties, lattice spacing, and microstructure. Gases and certain other high-vapor-pressure species readily dissolve in many solids. If these materials are used in a vacuum system, dissolved high-vapor-pressure species can diffuse to the surface and add to the gas load. If these materials are used to build a vacuum chamber, soluble gasses can permeate through the walls and add to the gas load [2]. For example, large amounts of hydrogen can be dissolved in palladium, platinum, and nickel. Significant amounts of hydrogen can diffuse out of vacuum fixturing made from

Table I

Selected Properties of Polymeric Materials Discussed in This Section

Material	Cost	Maximum Operating Temperature	Fabrication Ease	Outgas Rate (20–25°C)	Vapor Pressure at Room Temperature	Comments
Silicone Rubber hr pumping	Medium	150°C	Fair	$10^{-7}\frac{\text{T liter}}{\text{cm}^2\,\text{sec}}$ @ 4	—	• Avoid formulations with acetic acid & acetic anhydride • High thermal conductivity (for a polymer)
PTFE (Teflon®)	Medium	150–200°C	Good	—	10^{-5}–10^{-7} T	• Chemically inert • Low friction coefficient • Useful dielectric • Vapor pressure depends on thermal history
APIEZON W	Medium	80°C	Good	—	$<10^{-7}$ T	• Brittle, fractures easily
Glyptal® Varnish	Low	80°C	Good	—	2×10^{-4} T	• Time-honored standard
Polymide	Medium	150–200°C	Good	$5 \times 10^{-8}\frac{\text{T liter}}{\text{cm}^2\,\text{sec}}$ @ 4 hr pumping	—	• Properties depend on thermal history—outgassing can be comparable to stainless steel <200°C • Useful dielectric, adhesive • Absorbs 2–3% water
Polymethyl Methacrylate (PMMA, Lucite, Plexiglas®)	Low	80°C	Good	$120 \times 10^{-8}\frac{\text{T liter}}{\text{cm}^2\,\text{sec}}$ @ 3 hr pumping	—	• Releases water
Polyvinyl Chloride (rigid) (PVC)	Low	80°C	Good	$40 \times 10^{-8}\frac{\text{T liter}}{\text{cm}^2\,\text{sec}}$ @ 3 hr pumping	—	• Requires vacuum-compatible adhesives • Outgasses water

Table 2

Selected Properties of Ceramics Discussed in This Section

Material	Cost	Maximum Operating Temperature	Fabrication Ease	Temperature Where $V.P. = 10^{-6}$ T[a]	Comments
SiO_2 Silica, Glass, Quartz	Low–Medium	500°C (Soda Glass) 1600°C (Quartz)	Special Training	~800°C	
Al_2O_3 • Alumina • Alpha Phase (Sapphire or Corundum)	High	2000°C	Needs special techniques	~1000°C	• Withstands thermal shock • Good thermal conductor (for a ceramic) • Hard, chemically inert, transparent into i.r.
• Anodic aluminum oxide (gamma phase)	Medium	200°C	—	Outgasses $6 \times 10^{-6} \frac{\text{T liter}}{\text{cm}^2 \text{ sec}}$ (after 3 hr pumping @ 20°C)	• Porous—releases trapped water, etc. • Properties vary depending on forming process
ZrO_2 Zirconia	Medium	2700°C	—	~1100°C	
BeO Beryllia	Medium	2500°C	—	~1350°C	• *Dust poisonous*

[a] *V.P.* is vapor pressure.

raw stock that contains dissolved hydrogen. Considerable amounts of hydrogen can pass from lab air into the vacuum system when these materials are incorporated into a vacuum chamber's walls. Similarly, significant amounts of helium can pass though glass and vitreous quartz. Water and air permeate through most polymers.

Permeation and diffusion rates increase with temperature. Permeation occurs more readily through thin-walled chambers. Fick's laws predict how the diffusion and permeation rates of a material change as a function of pressure differentials, temperature, and wall thickness [3].

Porous, open-celled materials, such as many cast metals and hot-pressed or sintered ceramics, allow gas to travel into a vacuum, although this is not permeation as the word is conventionally defined by vacuum technologists.

Table 3

Selected Properties of Metals Discussed in This Section

Material	Cost	Maximum Operating Temperature	Fabrication Ease	Temperature Where $V.P. = 10^{-6}$ T[a]	Comments
Aluminum	Low	660°C	Moderate	812°C	• Metal soft and gummy — more useful as alloy
Cadmium	Low	320°C	Soft	119°C	• Poisonous, alloying constituent, avoid — bad vacuum material
Chromium[b]	Moderate	~1850°C	Difficult	977°C	• Hard, brittle, inert
Copper[c,d]	Moderate	1083°C	Moderate	852°C	
Gold	High	1064°C	Easy	947°C	• Inert, soft, good gasket material, good tribological couple with hard surfaces
Iron[b]	Low	1535°C	Moderate	1032°C	• Excellent thermal conductor at low temperatures • Rusts — bad
Lead	Low	328°C	Moderate	429°C	• Poisonous oxides, dust
Magnesium	Moderate	649°C	Moderate	246°C	• Used in aluminum alloys • Dust flammable
Molybdenum[b]	Moderate	2610°C	Moderate	1822°C	• Soft and gummy
Nickel[b]	Low	1455°C	Moderate	1072°C	• Work hardens readily
Rhenium	High	3180°C	Moderate	2217°C	
Silver	Moderate	962°C	Moderate	685°C	• Good tribological couple with hard surfaces • Beware of chemical reactions
Tantalum	Moderate	2996°C	Moderate	2237°C	
Tin[d]	Low	232°C	Moderate	807°C	• Soft — good low-temperature gasket material
Titanium	Moderate	~1660°C	Hard	1227°C	• Hard!
Tungsten	Moderate	3410°C	Hard	2407°C	• Brittle • Beward of oxide
Zinc[c]	Low	420°C	Easy	177°C	• Bad high-vacuum material

[a] *V.P.* is vapor pressure.
[b] Alloying constituents — steel.
[c] Alloying constituents — brass.
[d] Alloying constituents — bronze.

Table 4

Selected Properties of Miscellaneous Materials Frequently Used as Alloying Constituents

Material	Cost	Maximum Temperature	Fabrication Ease	Temperature Where $V.P. = 10^{-6}$ T[a]	Comments
Carbon	Graphite—low Diamond—high	— —	Graphite—easy Diamond—hard	1860°C	• ~1200°C diamond turns to graphite • Carbon sublimes
Phosphorous	—	—	—	88°C	• Alloying constituent • *Avoid*
Silicon	Moderate	1410°C	Needs special techniques	1147°C	

[a] *V.P.* is vapor pressure.

4.5.2 VACUUM CHAMBER MATERIALS

Glass, quartz, sapphire, stainless steel, and aluminum are commonly used for high, ultra-high, and extra-ultra-high-vacuum system chambers. More cost-effective, more easily machined and manipulated materials such as copper, brass, certain polymers, and somewhat porous ceramics (Macor®, and sintered, anodic, or plasma-sprayed aluminum oxide, zirconium oxide, etc.) are strategically employed in medium- or low-vacuum environments, such as roughing and forelines as well as high vacuums present in many practical physical and chemical vapor deposition systems, microelectronics fabrication tools used for plasma-assisted etching, freeze-dryers, etc.

4.5.2.1 Glass

During the first half of the twentieth century, well into the 1960s, the majority of medium-, high-, and ultra-high-vacuum systems were built using glass. Scientific glass blowing was part of the education of most experimentalists [4]. Valves and feedthroughs consisted of ground glass joints with stopcock grease and sometimes included O-ring seals or creative use of Glyptal® varnish and Apiezon "black wax" [5]. At least one company currently markets a fairly extensive line of glass vacuum components. Glass systems offer a number of advantages. They are relatively inexpensive (especially if the owner is able to do glass blowing). There is a clear view of processes occurring within the chamber, e.g., vacuum deposition, chemical synthesis, or freeze drying. Glass is smooth, relatively chemically inert, and easy to clean for vacuum applications.

There are several possible problems with glass systems. Glass is brittle and therefore fragile. Although glass is strong under compression, it is weak under tension. (A well-designed glass system will ensure that areas subjected to high levels of mechanical stress will be in compression. For example, glass-to-metal seals should contain layers of glasses with gradated thermal coefficients of expansion. The sequence of glasses selected should vary depending on whether the part is heated or cooled during use.) Soft, soda lime glass and, to some degree, borosilicate glass (Pyrex®) are vulnerable to thermal shock and soften at elevated temperatures in excess of 500 to 600°C. The sodium in soft glass and the boron in Pyrex® may chemically contaminate certain processes — especially processes used for fabricating silicon microelectronic devices. (Quartz withstands thermal shock more gracefully, softens at higher temperatures, and does not contaminate silicon devices; manipulating quartz requires advanced glass-blowing skills because it has a strong tendency to devitrify if it is not annealed properly.) If a glass or quartz system is not baked under vacuum for an extended amount of time, interstitial and adsorbed water vapor outgassing limits ultimate pressure. Ultimately, helium diffusion and permeation, especially for thin-walled systems at elevated temperatures, limits background pressure.

4.5.2.2 Steel

Commercially available, nonmagnetic, 304, 316, or 317 stainless steel vacuum system components are marketed by a number of companies. (Type 303 stainless steel is not suitable for use in high vacuum because high-vapor-pressure materials — sulfur, phosphorus, and selenium — are included as alloying constituents to improve machinability.)

Stainless steel systems offer a considerable number of advantages. Stainless steel is strong and resistant to many chemicals. There are documented practices and proprietary processes for cleaning and polishing stainless steel (see Chapter 4.9). Appropriate joining techniques allow a choice of tungsten inert gas (TIG) welding, spot welding, high-temperature brazing, and silver soldering. When soft copper or gold gaskets are used in valves and seals, steel systems can be baked to desorb contaminants and reduce the ultimate pressure. Large-scale manufacture of stock items and competition have reduced costs. Because of the wide variety of stainless steel vacuum components that are commercially available, customized chambers can frequently be assembled by simply bolting together stock items selected from catalogues. Components not commercially available can be fabricated by machinists and welders with some specialized training and care. A judicious choice of stock components, careful cleaning and assembly, and regular baking at elevated temperatures will result in a system capable of operating at ultimate pressures in the ultra-high and extra-ultra-high-vacuum range.

Several possible problems are associated with stainless steel. The raw stock is

inherently expensive and heavy. Nonmagnetic stainless steels require more sophisticated machining skills than other metals such as nickel and certain aluminum alloys. (If stainless steel is not machined at just the right speed with sufficient cutting fluid, it work-hardens.) As a result of phase transformations, nonmagnetic stainless steels deform at temperatures in excess of 500 to 600°C. Stainless steel is attacked by certain chemically aggressive gases and plasmas — fluorine and chlorine, for example [6]. Heavy metal by-products eroded or sputtered from stainless steel may introduce unacceptable contamination. (The silicon microelectronics fabrication community and the magnetically confined fusion community are especially concerned with trace heavy metal contaminants introduced by corroded or sputtered stainless steel fixtures.) Hydrogen diffuses within and permeates through stainless steel and limits ultimate pressure. Heating accelerates diffusion and permeation rate. The diffusion rate also depends on composition. Hydrogen permeates through chromium steel at $\frac{1}{10}$ to $\frac{1}{100}$ the rate at which it travels through pure iron. Adding carbon to iron to make mild steel increases the hydrogen diffusion rate. Carbon also generates a carbon monoxide background.

If hydrogen and carbon monoxide background pressures do not cause problems, then "mild" (low-carbon, cold-rolled) steel is a fine vacuum material if the surface is free of rust. Mild steels can be plated with copper, then nickel, or copper and nickel followed by chromium to avoid rust problems (and increase surface hardness). Mild steel represents a cost-effective alternative to stainless steel for constructing large vacuum systems with conservative performance. Problems arise if porous cast raw stock or raw stock with a high sulfur content is inadvertently procured. Of course, mild steel is magnetic and therefore inappropriate for systems where strong magnetic fields are present.

4.5.2.3 Aluminum and Its Alloys; Anodic Coatings

Chapter 4.8, "Aluminum-Based Vacuum Systems," provides detailed information regarding the use of aluminum as a vacuum material. Aluminum's popularity as a vacuum chamber material has increased in recent years. Aluminum vacuum system components offer an attractive alternative to the helium diffusion problems associated with glass and the hydrogen diffusion problems associated with steel. (The rate at which hydrogen diffuses through aluminum is three to four orders of magnitude lower than the diffusion rate through stainless steel.) As a general rule, aluminum raw stock is less expensive than stainless steel. Aluminum can be anodized to increase its corrosion resistance to some chemically reactive working gases. Anodic aluminum oxide coatings have useful dielectric properties for electrical insulation or shaping plasmas. Certain types of anodic coatings are extremely hard.

If aluminum components are machined in-house, and precise tolerances must be maintained and low background pressures are desired, attention must be paid

to the selection of raw stock. Exposing this metal to a chemically corrosive environment or specifying that piece-parts should be coated with a thick, homogeneous, cosmetically attractive anodic coating imposes additional restrictions on aluminum alloy selection. Cast aluminum alloys and cold-rolled plate stock may be porous and unacceptable for use as vacuum materials. The 7000 series aluminum alloys contain a high-vapor-pressure alloying constituent — zinc; zinc also interferes with homogeneous anodic film formation. The 2000 series aluminum alloys contain copper as an alloying constituent. Copper may cause staining if this alloy is anodized, or initiate localized corrosion in chemically aggressive ambients; frequently these tendencies are not cause for concern, and the 2000 series alloys are good ultra-high-vacuum materials. Extra-ultra-high vacuum is obtained most reliably in aluminum systems built from extruded 1000 series and 6000 series alloys. Machining these alloys in a dry, 95% argon, 5% air ambient significantly decreases outgassing in state-of-the-art vacuum systems. Commercially pure, 1000 series aluminum (and also the 5000 series alloys with magnesium) is soft, exceedingly ductile, and difficult to machine to precise tolerances. The 6000 series alloys, especially 6061-T6, are relatively hard, machinable, and relatively easy to weld or braze. Work-hardened 6061-T6 is hard enough to hold a knife edge that forms a vacuum-tight seal with pure and/or annealed aluminum gaskets. The 6000 series alloys, the 1000 series alloys, and the 5000 series alloys share the advantage that they can readily be anodized.

The composition and microstructure of an anodic aluminum oxide coating can vary widely, depending on the alloy selected and the anodization process parameters [7]. Certain anodization processes, such as those performed around 0°C in sulfuric or sulfuric plus oxalic acid solutions, form relatively dense, black, Martin® or Sanford® "hard coats" that exceed chromium steel in surface hardness. As a general rule, anodic coatings with extremely large pores form in phosphoric acid. Coatings with pores around 10 to 30 nm in diameter form in room temperature sulfuric and oxalic acid. Anodic coatings with extremely fine pores form in chromic acid. In addition to the acid selected for the electrolytic bath, porosity also depends on how the process is operated. As a general rule, increasing acid concentration, electrolyte temperature, and cell voltage results in more porous coatings. Drawing constant, direct current as opposed to rippled or pulsed current leads to more porous coatings. Thicker coatings tend to be more porous. Some anodic coatings are dyed for customers interested in cosmetic appearance and may not be desirable for use in vacuum. Often coatings are sealed by immersion in boiling water/salt solutions to improve corrosion resistance, and cosmetic appearance. Each of these process variations can substantially alter the spectrum of residual contaminants introduced into a vacuum system. Water and entrained electrolyte by-products that desorb from these porous, anodized aluminum may affect processes performed in these systems, especially at elevated temperatures.

Sometimes relatively thin, porous, drab brownish-yellow or green, phosphoric acid plus chromic acid chemical conversion coatings, e.g., Alodyne®, are applied

to aluminum as a low-cost alternative to anodization. Such coatings probably detract from vacuum performance.

Several problems may arise if aluminum alloy vacuum fixturing and chambers are selected. Options and opportunities for avoiding or designing around some of these problems are described in detail in Chapter 4.8. The crystal structure of many aluminum alloys changes irreversibly at relatively low temperatures (~400°C). Some alloys warp irreversibly at elevated temperatures. Aluminum is not a strong as steel and softens substantially at elevated temperatures (starting above 250°C). A layer of dense, hard, relatively chemically inert, air-grown aluminum oxide that is 3 to 30 nm thick depending on thermal history is almost always present. If the thin air oxide layer is breached or damaged, aluminum can be chemically reactive. The air oxide layer limits the number of techniques that can be used to bond aluminum. Aluminum solders that contain aluminum, zinc, and indium as a wetting agent are commercially available, but inappropriate for vacuum applications. Good-quality joints can be formed by inert gas welding.

4.5.3 SPECIAL-PURPOSE MATERIALS

Scientists and technologists frequently require electrical conduction or insulation, thermal conduction or insulation, and lubrication in practical vacuum systems. Sometimes people are concerned with costs—especially when they are building large, unique vacuum systems. This section provides guidelines for selecting materials for specific applications.

4.5.3.1 Electrical Conductors, Electrical Insulators, Thermal Conductors, and Thermal Insulators

ELECTRICAL CONDUCTORS

Oxygen-free, high-conductivity (OFHC) copper is an excellent electrical conductor. Smooth, clean OFHC copper is an excellent vacuum material. Unfortunately copper is chemically active and oxidizes readily when exposed to air, water, or carbon dioxide at elevated temperatures. Large areas of copper with somewhat rough oxidized surfaces (and adsorbed water) will slow pumping substantially into the ultra-high and extra-ultra-high vacuum range. If ultra-high and extra-ultra-high vacuum is required, less chemically reactive metals such as gold, chromium, and/or nickel can be plated on the copper surface. If electrodeposited chromium or gold coatings are deposited, the vacuum technologist must ensure that the electroplater will deposit dense, smooth coatings that are relatively free of organic bright-

ening and leveling additives with high vapor pressures. If the plating facility is careless and does not adequately control the plating process, their chromium or gold coatings may be rough, porous, and loaded with entrained hydrogen. Ultra-high-purity gold coatings (deposited from a cyanide bath with no brighteners or leveling agents) have a matte surface finish with a large surface area that may outgas excessive amounts of sorbed species. The best compromise may be to deposit a thin, smooth, dense "bright" gold or chromium coatings from a bath that contains some small concentration of organic additives, then vacuum bake the electroplated fixtures to decompose and outgas entrained organics and improve vacuum performance. If the plated part will be used at elevated temperatures over an extended amount of time, then a nickel undercoating should be deposited on copper fixtures. This nickel coating serves as a diffusion barrier. It should also be noted that both copper and gold sputter readily; copper and gold surfaces must be protected from plasmas and particle beams.

Silver is an excellent electrical conductor. Unfortunately silver can be a chemically unstable material (see Section 4.5.3.4). Silver can be gold-plated to avoid chemistry problems.

If only modest conductivity is required, then tantalum, rhenium, titanium, molybdenum, tungsten, nickel, or aluminum may suffice. Each of these metals has both good and bad properties. A wise vacuum technologist will use different metals for different components to exploit good properties and avoid potential problems. These metals have electrical resistivities that are higher than silver, gold, and copper. Nickel, aluminum, molybdenum, titanium, rhenium, tantalum, and tungsten sputter less readily than gold and copper. Molybdenum and tungsten are refractory metals with relatively low thermal expansion coefficients and an affinity for oxygen. They seal well with alumina and glass in electrical feedthroughs. However, reactions between oxygen and tantalum, titanium, molybdenum, and tungsten may be troublesome. These metals have a tendency to dissolve large amounts of oxygen or form thick porous oxide layers in partial pressures of oxygen at elevated temperatures. Hydrogen dissolves into titanium and tantalum, especially at elevated temperatures and causes serious embrittlement problems. Tantalum and molybdenum are extremely ductile and somewhat unpleasant to machine. Tungsten tends to be brittle and therefore somewhat difficult to manipulate. Titanium's extreme hardness presents problems to unskilled machinists with inferior tools. Rhenium wire is readily available but other forms of raw stock are less common. Nickel and aluminum alloys, such as the 6000 series, are good, cost-effective vacuum materials that are easily cleaned, shaped, and machined and do not form thick porous oxide layers that outgas excessively or shed particles. However, it should be noted that large pieces of nickel (and other magnetic materials) rob substantial amounts of energy from RF plasmas (in magnetron sputtering systems, for example) and heat-up. Thermal limitations associated with aluminum are noted in the preceding section and in Chapter 4.8.

ELECTRICAL INSULATORS, THERMAL CONDUCTORS, AND THERMAL INSULATORS

As a general rule, the plastic insulation and enamels on commercially available copper wire are not suitable for use in vacuum and must be removed. (Additional information regarding the use and behavior of organic compounds is included in Sections 4.6.1 and 4.7.) If electrical insulation is required on a bare metal wire, polytetrafluoroethylene (PTFE), the Teflon® formulation with the lowest vapor pressure and the highest softening temperature is an appropriate insulator for use down to 10^{-6} torr. Alternatively, polyimide can be used [8]. Polyimide (Kapton® and Vespel®) is an excellent vacuum material, with a vapor pressure comparable to stainless steel at low temperatures. Resins containing polyamic acid (the Pyre ML® and Pyralin® product lines from DuPont, for example) are available that can be painted on vacuum fixtures and then dried and air cured to form a polyimide coating. Not only can these resins be used to form electrically insulating coatings, they can be also used as adhesives and binders. Unfortunately, polyimide is notorious for generating water vapor when its temperature is allowed to rise above 150 to 200°C (depending on the curing temperature originally used to polymerize the resin and subsequent thermal history). Fixtures operating at high temperatures that require electrical insulation can be sheathed with glass, quartz, or ceramic beads.

Many ceramics are good electrical insulators [9]. Many 100% dense, refractory ceramics are excellent vacuum materials. In practice, ceramic stock is frequently formed by sintering or hot-pressing powder stock and may be somewhat porous. Porosity equates to large surface areas (with sorbed species) and virtual leaks. It is a good practice to question potential ceramics suppliers about the density of their product, especially open-celled materials, with tortuous passageways and large expanses of internal surface area. Many suppliers of sintered or hot-pressed ceramics and sputtering targets are willing to work with the vacuum technologist in selecting a range of powder particle sizes and bonding parameters to optimize the material's vacuum performance. It is desirable to use dense, close-celled ceramic stock. If this is not practical, the second-best strategy may be to obtain stock with an open-celled structure formed from relatively large-sized primary particles, with a minimum amount of internal surface area and large openings that pump down easily.

Frequently, vacuum users are concerned with building heaters and need a good thermal conductor that electrically insulates the object being heated from resistive graphite, refractory metal or nickel-chromium alloy heating elements [10]. Diamond, with a thermal conductivity approximately equal to 40 times that of copper at room temperature, is an excellent thermal conductor/electrical insulator followed by beryllium oxide, with a thermal conductivity somewhat less than copper, and aluminum oxide (preferably α-aluminum oxide—"corundum" or "sap-

phire"). Beryllium oxide particles are extremely toxic, and extreme care must be exercised when working with this material. Boron nitride is used as an electrical insulator on some commercially available heaters that are suitable for a wide variety of vacuum applications. Unfortunately, boron nitride has a tendency to adsorb water.

Cubic zirconia is an excellent thermal insulator. It withstands extremely high temperatures and tolerates thermal shock well. Cubic zirconia spacers are particularly suitable for thermally isolating hot stages from supporting platforms to minimize thermal mass.

4.5.3.2 Lubricants

Lubricants are used in a vacuum system for three reasons:

- In a ground glass joint or elastomer O-ring sealed system, the joints and O-rings can be lightly greased so they move, seat, and seal better. (Many vacuum technologists argue against greasing O-rings.)
- Lubricants can be used between moving surfaces in the vacuum to reduce the friction coefficient, wear rate, and particulate generation.
- Lubricants are used as anti-seizing agents. (Mating surfaces of identical metals, e.g., stainless steel nuts and bolts baked at high temperatures in vacuum, are especially likely to seize.)

Three types of lubricants may be suitable for vacuum applications:

- Wet, organic, or silicone-compound oils and greases
- Dry lubricants, e.g., PTFE (Teflon®) and metal dichalcogenide compounds, e.g., molybdenum disulfide and tungsten diselenide
- Metal-on-metal combinations

The choice of a lubricant to use depends on a number of considerations relating to the specific application, including

1. Operating temperature and vapor pressure
2. Sliding motion or rolling motion
3. Reactive gas, electron beam, or plasma ambient
4. Heavy or light loading
5. Extensive repetitive usage or moderate use

OILS AND GREASES

Wet lubricant oils and greases that are specially formulated for vacuum use [11] are appropriate for providing lubrication on ground glass joints and elastomer O-ring seals.

Situations involving repetitive, rolling motion under heavy loads may demand the use of a wet lubricant. Because oils and greases flow, they replenish areas experiencing heavy wear, and are good lubricants for *severe* tribological situations. Because oils and greases flow, they tend to creep over all surfaces in a vacuum system. Taking care to design moving parts so a dry film lubricant or a metal-on-metal couple will suffice is strongly advisable.

DRY LUBRICANTS

Dry lubricants are especially suitable for situations involving sliding motion within the vacuum chamber with light loads and moderate usage. It may be appropriate to machine selected moving parts from polytetrafluoroethylene. Proprietary preparations are available that contain a suspension of PTFE particles [12]. For light loads and moderate use, it might be sufficient to drive off the relatively high-vapor-pressure vehicles and binders in these preparations and operate with the particles that remain.

Alternatively, metal dichalcogenides compounds, in particular, molybdenum disulfide and tungsten diselenide have a long and well-documented history of use as vacuum lubricants. Compared to molybdenum disulfide, tungsten diselenide has a lower vapor pressure, withstands higher temperatures, and is considerably more expensive. Most proprietary preparations with molybdenum disulfide also contain copious amounts of high-vapor-pressure organic dispersants and binders. The temperatures required to drive off these organics in a feasible amount of time are dangerously close to temperature where molybdenum disulfide degrades and decomposes (150 to 200°C). Many of these preparations also contain large amounts of graphite, which loses its lubricant properties in vacuum. Adherent coatings of pure molybdenum disulfide with good lubricant properties can be deposited by sputtering or by electrophoretic deposition [13].

If antiseize coatings on metal hardware are required, reagent-grade molybdenum disulfide, tungsten disulfide, niobium diselenide, or tungsten diselenide powder can be mixed with absolute ethanol or methanol and painted on screw threads, etc. After the alcohol evaporates, the fixtures can be assembled.

METAL-ON-METAL COUPLES

A judicious selection of different metals that are in moving contact with one another may provide sufficiently low friction coefficients and wear rates so wet or dry film lubricants are not needed. As a general rule, two metals that are in moving contact should be as different from one another as possible so adhesion and seizing do not occur. One effective strategy is to deposit a soft, ductile metal coating such as tin, indium, silver, or gold on the part that is easiest to maintain or replace. A hard metal or ceramic is then used for the mating part. Martensitic hard-

facing, nitrided steels such as nitronic 60, carburized steels, titanium nitride coatings, hard-anodized aluminum, etc., can be selectively employed to fabricate moving contacts that do not require lubricants as traditionally defined.

4.5.3.3 Unusual Vacuum Materials

Frequently, materials not traditionally used for vacuum work are evaluated or used — especially for large, unique systems where cost is important.

Certain organic materials are potentially useful vacuum materials. Polyvinyl chloride (PVC) is reported to have a vapor pressure as low as stainless steel at room temperature after a few hours pumping to remove water vapor. One commercial supplier has developed a system for forming vacuum-tight seals for this polymer [14]. (Trouble rapidly ensues if a misguided technologist attempts to use commercially available PVC adhesives sold at hardware stores and plumbing supply houses to bond together PVC vacuum parts.) Polymethylmetacrylate (also commonly called PMMA, "acrylic," "Lucite®," or "Plexiglass®") has been used to construct a vacuum chamber for a large, pulsed power system at Sandia National Laboratories in Albuquerque, New Mexico where cost, dielectric integrity, structural strength, and weight are considerations. A high partial pressure of outgassed water was a major problem until this system was retrofitted with sufficiently large cryopumps. Polystyrene divinylbenzene (Rexolite®) has a higher breakdown strength, useful structural strength, and a substantially lower water absorption/desorption rate than PMMA. It also costs substantially more.

Several years ago, there was intense interest in using concrete as a vacuum material in the large tunnels and shafts used for nuclear weapons tests at the Nevada Test Site and other explosive test chambers. Unfortunately, despite diligent attention to formulating vacuum-grade concrete, the best efforts could only be pumped down to a pressure of 10 torr. Ultimately, steel liners were implemented to maintain a useful vacuum in the Nevada Test Site tunnels and shafts.

4.5.3.4 Materials to Avoid

In general, it is a good policy to consider all materials as unsuitable for use in vacuum until information to the contrary becomes apparent.

HIGH-VAPOR-PRESSURE MATERIALS

As a general rule, the vacuum technologist should be wary of organic materials, especially in high-vacuum systems or at elevated temperatures. The vapor pres-

sures of many practical polymers vary depending on their thermal history and the degree of polymerization and cross-linking that has occurred. As discussed in Section 4.6.1, problems can arise as a result of high-vapor-pressure additives that may vary from batch run to batch run — plasticizers, fillers, and binders — not the polymer itself.

Certain metals are unacceptable for high-vacuum work, e.g., lead, zinc, and cadmium. This includes soft solders that contain lead; hard solders that contain cadmium, cadmium-corrosion-resistant coatings on nuts, bolts, and screws, and zinc, which is an alloying constituent in aluminum solders, brass, some bronzes, and the 7000 series aluminum alloys. Zinc is used as a corrosion-resistant coating on galvanized steel.

Vapor pressure data suggest that brass and lead-tin solder are acceptable materials at foreline pressures. Brass is a good choice for foreline and roughing manifolds and fixturing. It is tempting to use lead–tin solder to join brass or copper roughing and foreline manifolds and fixturing. Unfortunately, lead–tin solders are not very strong. Continued use and vibration from a mechanical pump can cause lead–tin solder joints to crack and leak. Cadmium-free silver solder and indium-based solders are preferable to lead–tin "soft" solder.

Selenium, sulfur, and phosphorous have unacceptably high vapor pressures. Selenium and sulfur are frequently used in metal machining fluids. The vacuum technologist must take care that his or her machine shop avoids the use of such products when machining vacuum fixtures. A copious stream of ethanol or isopropanol serves as a good cutting fluid for light machining. Vegetable shortenings such as Crisco® expedite heavier machining operations and clean up readily in soap and water. (Additional, proprietary selenium and sulfur-free machining fluids are recommended in Chapter 4.8.) Phosphorous can inadvertently be introduced into a vacuum system in the form of electroless nickel coatings; hard, electroplated phosphorous nickel coatings; certain types of amorphous Metglas® metals; phosphorous bronze and oxygen-free copper where phosphorous is used to scavenge oxygen. (This is not the same as oxygen-free high-conductivity copper).

CHEMICALLY UNSTABLE MATERIALS

Materials that chemically react at low temperatures should be regarded with skepticism, especially for use in systems seeing elevated temperatures or plasmas. For example, vapor pressure data suggest that silver metal is a fine vacuum material. Unfortunately, silver readily oxidizes in lab air, and silver oxide reduces and liberates oxygen in vacuum at temperatures slightly above 200°C. Silver is even more likely to react with pollutants in air, e.g., sulfur, sulfur oxides, or hydrogen sulfide, than with oxygen. These decomposition products are even more of a nuisance than oxygen.

REFERENCES

1. A. Roth, *Vacuum Technology*, 3rd ed. (Elsevier Science, New York, 1990), Section 4.1.2.
2. There are applications where the ability of certain gases to permeate certain materials is employed beneficially to build calibrated leaks and to purify gases.
3. A. Roth, *Vacuum Technology*, 3rd ed. (Elsevier Science, New York, 1990), Section 4.2.1.
4. J. Strong, H. V. Neher, and A. E. Whitford, *Procedures in Experimental Physics* (Prentice Hall, Englewood Cliffs, NJ, 1938).
5. The Apiezon Company, Philadelphia, PA.
6. Vacuum system materials for handling partial pressures of highly chemically reactive gases are beyond the scope of this chapter. The American Vacuum Society has published recommended practices for designing systems to handle hazardous gases—John F. O'Hanlon and David B. Fraser, *J. Vac. Sci. Technol.*, **A6**(3) (May–June 1988) 1226–1254.
7. *The Metal Finishing Guidebook*, published at regular intervals by Metals and Plastics Publications, Inc., Hackensack, NJ, has both generic instructions for electrochemical processing plus advertisements for proprietary baths and subcontract processing facilities.
8. B. R. F. Kendall and M. F. Zabielsky, *J. Vac. Sci. Technol.*, **3**(3) (1966) 114. The DuPont Corp., Wilmington, DE, offers extensive information about a wide variety of Teflon® products including PTFE and two polyimide resin formulations it has developed.
9. Some ceramics, e.g. tin oxide, manganous oxide, cuprous oxide, lead oxide, titanium nitride, and substoichiometric compounds, e.g., TiO_x are conductors or semiconductors.
10. If nickel–chromium, Nichrom® heating elements are being considered, vapor pressure data should be considered. Chromium sublimes from this alloy at moderately high temperatures in medium to high vacuums. It should also be noted that many refractory metals, especially those with body-centered cubic microstructures, are vulnerable to hydrogen or oxygen embrittlement at elevated temperatures.
11. The Apiezon Company in Philadelphia, PA, markets a variety of silicone greases for vacuum applications. Silicone-based diffusion pump oils, such as those manufactured by Dow-Corning may be useful as light lubricating oils in vacuum. Unfortunately, silicones tend to creep and slowly cover all warm surfaces.
12. Lubricants with suspensions of Teflon® particles, for example Vydex®, are commercially available from DuPont Corp., Wilmington, DE.
13. Hohmans Inc. in Cleveland, OH can be subcontracted to sputter MoS_2 lubricant coatings. Alternatively, Panitz et al., *J. Vac. Sci. Technol.*, **A11**(4) 1993 pages 1441–1446 describe a simple process for electrochemically depositing MoS_2. This paper has extensive references to research and development studies of sputtered MoS_2.
14. Nuvac®,—U.S. Patent No. 5,341,567, assigned to Galiso, Inc. Montrose, CO.

CHAPTER 4.6

Demountable Seals for Flanges and Valves

R. N. Peacock, Vice President for Development (Retired)
HPS Division of MKS Instruments, Inc.

4.6.1 SEALING OVERVIEW: POLYMER AND METAL SEALS

Construction of conventional vacuum systems would be impossible without leak-free means of demountably joining components. Valves used for the isolation of sections of a system or the control of process gas require similar seals, as do some motion feedthroughs. Seals that normally remain fixed in use are known as *static seals*; seals similar to those in the gate or nosepiece of a valve or a motion feedthrough are called *dynamic seals*.

Seals may use either polymers or metals as the gasket material, depending on the application. Both polymer and metal seals are discussed in this chapter.

4.6.1.1 Considerations Limiting the Use of Polymer Seals

The allowed operating temperature range for polymer seal materials is limited at both high and low values. For example, the commonly used elastomer Viton® [1] becomes hard and seals using it may fail below 0°C. To avoid decomposition and compression set the maximum temperature should be limited to about 150°C for long-term operation or system bakeouts.

ISBN 0-12-325065-7
$25.00

In most well-designed systems using elastomer O-ring seals, the seals are an important source of the system gas load, and often determine the ultimate pressure. A numerical example in Section 4.6.2.2 illustrates this.

In addition, polymers may suffer from chemical attack, aging, and radiation damage.

4.6.1.2 Applications Requiring Metal Seals

Metal seals are preferable in the following situations:

1. When the baking or operating temperature is too high or low for polymers
2. In some chemical environments
3. When gas from the outgassing of polymers or the permeation of atmospheric gasses through polymers is not acceptable
4. When nuclear or ionizing radiation would damage elastomer seals

The cost of building with metal seals is usually much higher than with polymer seals. Metal seals require greater forces, so that components must be much stronger and heavier. Greater precision and specific surface finishes are also needed.

4.6.2 THE ELASTOMERIC AND NONELASTOMERIC POLYMERS USED IN VACUUM SEALING

4.6.2.1 General Chemical and Mechanical Properties of Common Seal Polymers

Table 1 describes some polymers commonly used in vacuum sealing, along with some characteristics of each material. Most materials listed are elastomers. Elastomers are rubberlike substances that tend to return to their original shape after deformation, although the recovery is imperfect. PTFE, PCTFE, and polyimide are not elastomers, but they are used in many small valves for nosepiece and body seals.

Unfortunately, naming the material found in a polymer seal is not sufficient description to fully characterize it. Consider the fluoroelastomer material Viton®. The term is often used generically for the finished elastomer, not properly as the trademark of a resin. A seal manufacturer has many proprietary formulas for O-rings in order to control characteristics such as compression set, hardness, and sometimes color. To make a generic "viton" O-ring the manufacturer may buy Viton® resin from Du Pont Co., or the similar Flourel® [2] resin from 3M Co. A

Table I

Some Frequently Used Polymer Seal Materials with Maximum Long-time Operating Temperature, Expected Installed Life, and Tear Strength/Abrasion Resistance

Generic Name of Polymer	Trade Name(s) Commonly Used	Max. Temperature (°C)	Installed Life (yrs)[a]	Tear Strength/ Abrasion Resistance[b]
Nitrile (Buna-N)		85	1 to 5	Good
Butyl		85	5 to 10	Good
Ethylene propylene		100	5 to 10	Good
Polyurethane		90	5 to 10	Excellent
Fluoroelastomer	Viton®	150	15 to 20	Good
Perfluoroclastomer	Kalrez® Chemraz®	250	15 to 20	Excellent
Silicone	Silastic®	230	15 to 20	Poor
PTFE[c]	Teflon®	280	Long	Excellent
PCTFE	Kel-F®	200	Long	Excellent
Polyimide	Vespel® Kapton® Envex®	250	Long	Excellent

Note: Trademark Ownership:

Viton, Kalrez, Teflon, Vespel, and Kapton are registered trademarks of DuPont Co., Wilmington, DE 19898.

Fluorel and Kel-F are registered trademarks of 3M Co., St. Paul, MN 55144.

Chemraz is a registered trademark of Greene, Tweed & Co., Inc., Kulpsville, PA 19443.

Silastic is a registered trademark of Dow-Corning Corp., Midland, MI 48686.

Envex is a registered trademark of Rogers Corporation, Lithonia, GA 30058.

[a]"Parker O-ring Handbook," (Parker Seal Group, Lexington, KY, 1992), p. A3–14.

[b]"Parker O-ring Handbook," (Parker Seal Group, Lexington, KY, 1992), p. A3–32.

[c]PTFE is subject to moderate cold flow at room temperature, and it becomes a serious problem above 200°C.

Group II metal oxide is added to scavenge HF produced during the polymerization, a filler such as carbon black or diatomaceous earth to give the desired hardness and compression set characteristics, and a curing agent to cause polymerization. Because the compounding by each supplier may differ, so may the properties of the seals [3]. To complicate this situation further, there are several different Viton® polymers. The one generally used today because of its low compression set is Viton-E60C.

Selection of a polymer seal material for a specific application should consider:

1. Temperature requirements, including those for system bakeout
2. Installed life specification for reliable sealing

3. Special chemical exposure, internal or external
4. Abrasion resistance/tear strength requirements
5. Compression set characteristics
6. Permeation and outgassing characteristics
7. Possibility of damage by nuclear radiation
8. Availability and cost

The information in Table 1 is helpful in selecting elastomers for seals on the basis of maximum temperature, proposed life for reliable sealing, and tear strength/abrasion resistance. In many vacuum applications, chemical resistance to process gas is not the most critical factor in elastomer selection. When it is important, there are sources of information [3,4].

Atmospheric ozone is often a cause of O-ring cracking. Nitrile rubber should never be used around high-voltage equipment or UV light, and even the small quantities of ozone in the atmosphere due to sunlight can cause cracking of stressed nitrile in less than one year. The fluoro, perfluoro, urethane, and silicone elastomers resist ozone damage.

Compression and compression set may be defined with the help of Figure 1. The compression, C, from Figures 1(a) and 1(b) is

$$C = (d_1 - d_2)/d_1 \times 100\% \tag{1}$$

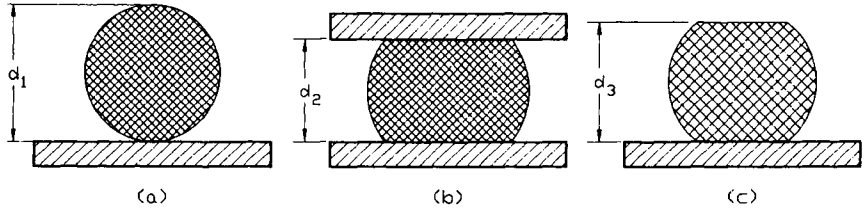
And the compression set, S, is defined as

$$S = (d_1 - d_3)/(d_1 - d_2) \times 100\% \tag{2}$$

The procedure for measuring compression set is defined by ASTM D395-89 Method B [5]. This standard requires that tests of compression set use a 2-214

Fig. 1.

Definition of compression and compression set.



(a) shows an unconstrained O-ring seal of minor diameter d_1. In (b) this O-ring is compressed to a thickness d_2. In (c) the compression has been removed and the O-ring has partially recovered, to a thickness of d_3. The compression set expressed in percent is defined as

$$\text{Compression set} = (d_1 - d_3)/(d_1 - d_2) \times 100\%$$

Measurements for tabulated compression sets are usually made with a compression of 25%, for 22 or 70 hours at a specified temperature. Compression set usually increases with temperature and confinement time

Table 2

Compression Set Given as the Percentage of Initial Compression for Some Elastomers and for Several Test Temperatures (the tests were made in accordance with ASTM D-395 B using 2-214 O-rings measuring 1.0" × 0.139" and a 25% initial compression)

		Temperature (°C)				
Elastomer	Time (hrs)	20	100	150	200	250
Nitrile[a]	22	—	24	42	100	100
	70					
Ethylene propylene[a]	22	—	8	22	47	100
	70					
Silicone[a]	22	—	4	6	25	80
	70					
Fluoroelastomer (Viton® E-60C)[b]	70	—	—	6	25	70
	70					
Perfluoroelastomer						
(Kalrez® 4079)[c]	70	12	—	—	32	—
(Chemraz® 505)[a]	70	19	—	—	25	—

[a] "O-ring Handbook" (Precision Rubber Products Corp., Lebanon, TN, January 1984), pp. 2–9.
[b] "Viton E-60C Data Sheet V-D-3-301" (DuPont Corp., Wilmington, DE, no date).
[c] "Kalrez Data Sheet 216111B" (DuPont Co., Wilmington, DE, April 1992).

O-ring (1-inch diameter × 0.139-inch cross section), and a time of 22 or 70 hours at a specified temperature and a compression of 25%. Compression set usually increases with temperature and confinement time.

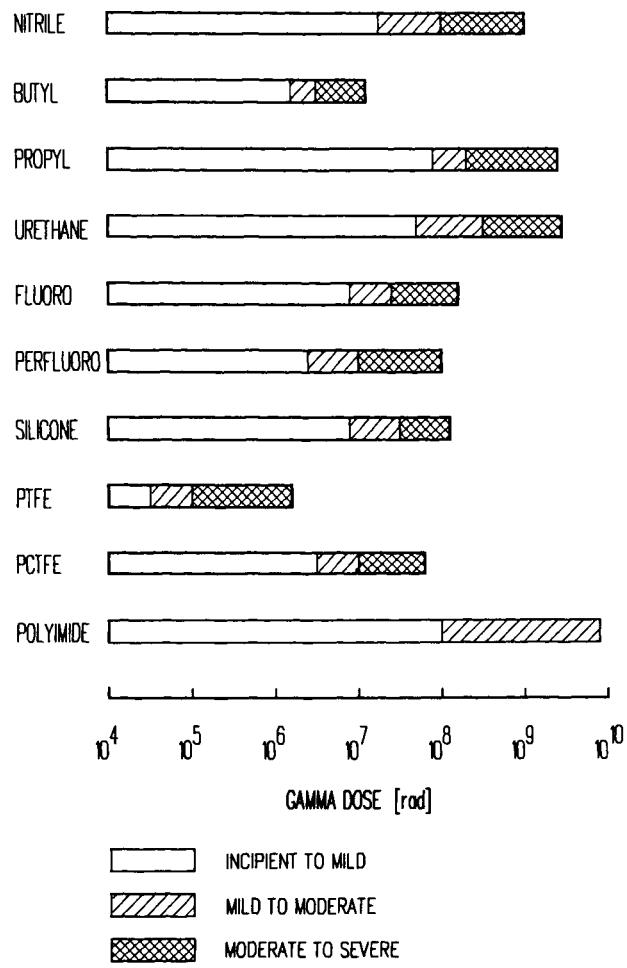
In most joints with elastomers, the force to maintain the seal is provided by the elastic nature of the material. When it is set, this force is reduced. Although such a joint may remain leak free if undisturbed, motion due to vibration or temperature changes may cause it to leak.

Table 2 compares the compression set characteristics of several materials. Short-time compression set is not a problem when these elastomers are used below the maximum temperatures given in Table 1. However, compression set may be severe when a seal is confined for a long time at temperatures allowed by Table 1.

It is generally believed that compression set is greater as the compression increases. This means that it is desirable to use the minimum compression that produces reliable sealing.

Exposure to an amount of nuclear radiation greater than some threshold value causes the hardness and compression set to increase, and the elongation to decrease. Figure 2, using data from Greene, Tweed [6] for the perfluoroelastomer, and from Van de Voorde and Restat [7] for the other materials, shows the effects of gamma radiation on a group of polymers. The butyl, fluoroelastomer, and PTFE materials are notably poorer when subject to radiation damage. Polyurethane and

Fig. 2.



Radiation damage for some polymers as a function of gamma ray dose.

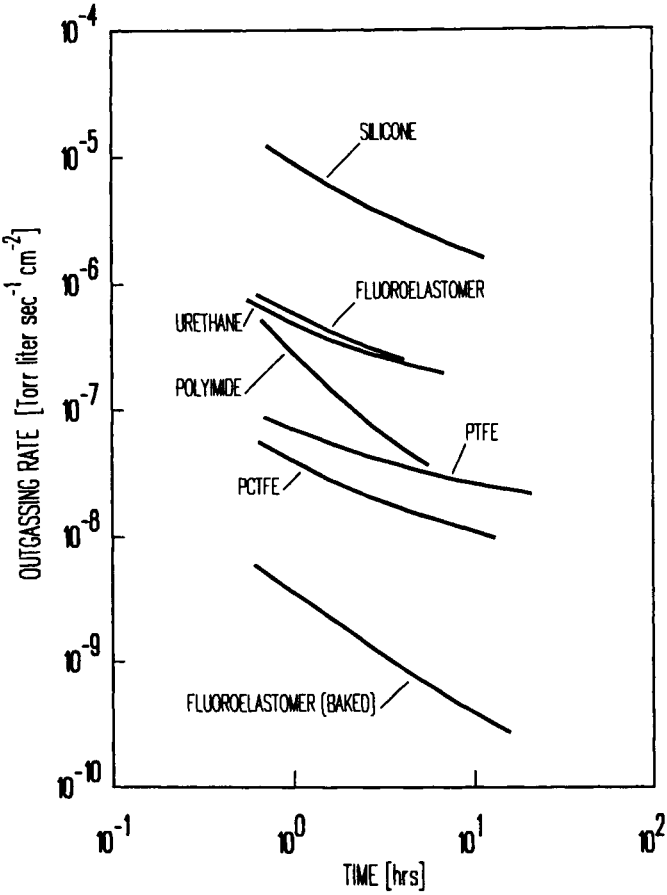
ethylene propylene are the best of the elastomers, although the difference is only about a decade between the fluoroelastomer (poor) and polyurethane (better).

4.6.2.2 Outgassing and Permeation of Polymers Used in Sealing

O-rings are a significant source of gas in vacuum systems. Part of the gas is due to outgassing, and this contribution may be reduced by baking. Figure 3 shows the

Fig. 3.

Room temperature outgassing rates for some polymers commonly used in vacuum sealing.



Data plotted were selected from the literature. The polymer samples were initially outgassed in vacuum, then exposed to room air. The curves show the outgassing rates as a function of time during a subsequent pumpdown. The rates decrease between one and two orders of magnitude during the interval from 1 to 10 hours, but moderate baking causes a three-decade decrease as comparison of the two fluoroelastomer curves shows.

outgassing rate as a function of time after start of pumpdown for several common polymers. The data are from Dayton [8], Hait [9], and Thieme [10]. The other source of gas is permeation through the seal. The quantity of gas per unit time permeating a slab of material of area A and thickness d, when the pressure of the gas is P_1 outside, and P_2 inside, is given by

$$Q = \mathrm{K}\, A\, (P_1 - P_2)/d \tag{3}$$

Table 3

Permeation Constants, K, for Common Seal Polymers and Several Gases for Temperatures in the Range 20–30°C. Units are std. $cm^3\ s^{-1}\ cm^{-2}\ torr^{-1}$

	$K \times 10^{10}$			
Polymer	He	N_2	CO_2	H_2O
Nitrile (Buna-N)[a,b]	1.0	0.024	0.75	100
Butyl[a,b,c]	0.86	0.032	0.52	4–20
Ethylene propylene[a,b,d,e]	2.6–3.9	0.44	0.92	7–70
Polyurethane[a,b]	0.47	0.049	1.4–4.0	35–1,250
Fluoroelastomer[a,b,d,f,g]	1.2–2.3	0.03–0.07	0.3–0.8	5.2
Perfluoroelastomer				
Kalrez[h]	11.2	0.30	2.5	—
Chemraz[i]	14.3	0.88	—	—
Silicone[a,b,d,j]	31–33	10–16	60–300	400–1,000
PTFE (Teflon)[a,b,k]	6.8	0.14–0.32	1.2	3.6
PCTFE (Kel-F)[l]	0.22	0.0005	0.014	0.01
Polyimide (Kapton)[g]	0.25	0.0039	0.26	—

[a] "Parker O-Ring Handbook," (Parker Seal Group, Lexington, KY, 1992), pp. A2–5.
[b] A. Lebovitz, *Modern Plastics*, **43** (1966), 139.
[c] J. E. Ayer, D. R. Schmitt, and R. M. Mayfield, *J. Polymer Sci.*, **3** (1960) 1.
[d] H.-P Weise, K.-H. Ecker, H. Kowalewsky, and Th. Wolk, *Vuoto*, **20** (1990) 225.
[e] A. W. Myers, J. A. Myer, C. E. Rogers, V. Stannett, and M. Szwarc, *Tappi*, **44** (1961) 58.
[f] L. Laurenson and T. M. Dennis, *J. Vac. Sci. Technol.*, **A3** (1985) 1707.
[g] W. G. Perkins, *J. Vac. Sci. Technol.*, **10** (1973) 543.
[h] "KALREZ Data Sheet 213509C," DuPont Corp., Wilmington, DE (September 1994).
[i] "CHEMRAZ Data Sheet WSP 789," Greene, Tweed & Co., Kulpsville, PA (July 1989).
[j] J. A. Barrie, in *Diffusion in Polymers*, J. Crank and G. S. Park, eds. (Academic Press, New York, 1968), pp. 259–313.
[k] R. A. Pasternak, M. V. Christensen, and J. Heller, *Macromolecules*, **3** (1970) 366.
[l] "KEL-F 81 Plastic Data Sheet," 3M Industrial Products Div., St. Paul, MN (1989).

where K is the permeation constant. If A is in cm^2, d in cm, pressure in torr, time in seconds, and Q in $cm^3\ s^{-1}$ then K will have the units $cm^3\ s^{-1}\ cm^{-2}\ cm\ torr^{-1}$.

Table 3 contains values of K for a group of elastomers and for several common gases. At many points in the table a range of values is given because the data found in the literature are not consistent. The values may scatter because of experimental error, or because samples differed. However, the table is adequate to permit comparison of permeation for these polymers. Permeation is moderate for the fluoro and butyl elastomers, and much larger for nitrile, and especially for silicone rubbers. The advantage of PCTFE for seals in small valves used for gas handling is clear.

The permeation of atmospheric gases often determines the base pressure of a

vacuum system using polymer seals. As an example showing the importance of permeation consider a stainless steel system of internal surface area 1 m^2 having five fluoroelastomer O-rings of 8-inch diameter, and nominal $\frac{1}{8}$-inch cross section. Let the effective pumping speed for the residual gases be 100 liter/s. An order of magnitude calculation will yield an estimate of the pressure 10 hours after a 150°C bake. The outgassing rate of a fluoroelastomer used as a seal to atmosphere under these assumptions is about 10^{-8} torr liter s^{-1} cm^{-2} according to de Csernatony [11] and Dayton [8,12]. The outgassing rate of the baked stainless steel may be taken as 10^{-12} torr liter s^{-1} cm^{-2} after Young [13]. The surface area of the O-rings is about 3.5×10^2 cm^2, so that the total gas load due to them is 3.5×10^{-6} torr liter s^{-1}. That from the stainless steel is 1×10^{-8} torr liter s^{-1}. The equilibrium pressure is given by $P = Q/S$. The contribution of the seals to the total pressure is 3.5×10^{-8} torr, while that of the entire chamber surface is only 1×10^{-10} torr. If the system had been built using metal seals the gas load due to the seals would be negligible, and the ultimate pressure would be about 1×10^{-10} torr. It is clear that elastomer seals are often the limiting factor in determining the ultimate pressure of a vacuum system.

Permeation is important in another specific instance: when a mass spectrometer leak detector is used with helium search gas to check joints with polymer seals, permeation can cause false signals. After an O-ring is exposed to helium for about a minute, a small signal is seen that rapidly increases. Another order of magnitude calculation will verify that a helium permeation signal should be expected. Consider a fluoroelastomer O-ring of 100 mm diameter and 3.5 mm (0.139 in) cross section. Assume the effective thickness to be 3.0 mm, and that $\frac{1}{4}$ of the area is exposed to helium at atmospheric pressure. From Table 3 the permeation constant for helium is about 2×10^{-10} cm^3 s^{-1} cm^{-2} cm $torr^{-1}$. Then the equilibrium permeation rate is 4×10^{-6} std. cm^3 s^{-1}. This is a large leak signal.

Once saturated with helium a polymer seal continues to outgas for a long time, causing a changing high background that makes further leak testing difficult. For this reason, it is desirable to design leak detectors without elastomer seals in critical areas where the mass analyzer would be exposed to evolved gas.

Particulate contamination is a serious problem in integrated circuit fabrication. O-rings subject to friction in dynamic seals can produce particles. Perhaps these particles are from the fillers used in compounding. Means of reducing particulates from seals are discussed by De Laurentis, Wu, Salfelder, and Uritsky [14].

4.6.2.3 Review of O-Ring Material Selection

O-rings made of nitrile rubber are limited to applications in unbaked vacuum systems because the maximum allowed temperature is 85°C. It is prone to cracking

in the presence of ozone from high-voltage equipment or natural sources, especially when compressed. Table 3 shows that the permeation constant for water through nitrile is several times as high as for a fluoroelastomer. However, nitrile O-rings are widely used because they are inexpensive.

Butyl rubber is only moderately permeable to atmospheric gases, including water. It is sometimes used as a low-cost seal material for unbaked systems.

Polyurethane may be chosen for seals because of its exceptional toughness and wear resistance, specific chemical resistance, or its slightly better resistance to radiation damage as compared to other elastomers. Ethylene propylene is average in most respects, but similar to polyurethane in radiation resistance.

Silicone rubber is not used for general vacuum sealing. Its mechanical properties are poor, and the permeation of atmospheric gases is much higher than for other elastomers. However, it is used in some applications because it will withstand short-term exposure to temperatures as high as 300°C without complete failure. Perfluoroelastomers have good mechanical properties, including tear strength, abrasion resistance, and lubricity. They can be used to higher temperatures than the other elastomers, and are average as regards permeation. Chemical resistance is excellent. However, perfluoroelastomer seals are so expensive that they are used only when their special advantages justify the cost.

PTFE and PCTFE are not elastomers and have little application in static seals. O-rings with an outer layer of PTFE and an elastomer core are available [15]. PCTFE is commonly used for the nosepiece and body seals in small valves for gas handling. PCTFE is a semicrystalline material, and has low gas permeation. Both have excellent chemical resistance.

Polyimide has been used in vacuum valves intended for baking to 300°C. It is very hard and requires large seal forces. It is also hygroscopic. Hait [9] discussed the merits of polyimide for UHV seals. In the form of thin sheet, polyimide has been used for sealing flanges [16].

It is clear why a fluoroelastomer is so commonly used for vacuum sealing. It withstands temperatures adequate for the bakeouts required to remove adsorbed water from system walls, the life is good, the compression set moderate, and it is among the better materials as concerns permeation. It is, however, easily damaged by nuclear radiation.

4.6.2.4 Design Considerations When Using Elastomer Seals

Proper finish of contacting sealing surfaces and defining the compression of the seal are just as important as selection of the material. Most static seals are in the form of O-rings made of elastomer materials. Sessink and Verster [17] studied the conditions necessary for sealing with elastomer O-rings and found that the

important consideration was a contact pressure at the metal to polymer interface of 15 kg/cm^2. The proper compression for a fluoro or nitrile elastomer O-ring should fall in the range of about 20–25%, including extreme cases arising from the tolerances of the minor diameter of the O-ring and of the compression determining feature. Greater compression decreases seal reliability by aggravating compression set and, particularly for nitrile, cracking. The compression recommendations in the handbook of one of the large O-ring manufacturers are not necessarily correct [18]. The seal force per unit length required to produce 20 to 25% compression is dependent on the minor diameter of the O-ring and the hardness of the material. For an O-ring of $\frac{1}{8}$-inch nominal minor diameter, hardness 75 Durometer, the required compressive force is about 15 lb per linear inch of seal. The manufacturers' handbooks have some force vs. compression data for other cases [19,20]. The optimum seal contact surface on a flange or valve seat is not mirror smooth. Sealing is more reliable against a surface with fine concentric grooves. Good results are obtained using a face-turned surface cut with a $\frac{1}{32}$-inch radius insert and a feed of about 0.003 inch. A surface made this way will have approximately a 32-microinch finish measured across the grooves, and will be free of roughness in the direction of the grooves. It is not desirable to use abrasives to finish seal surfaces because of the smearing of the surface, which may embed abrasive particles, and because the finish marks are not as easily controlled as are the marks from turning.

An O-ring must not be twisted or cut, either when inserted into a groove, or during use. Grooves should be made with radiused edges to prevent seal damage. In a "radial seal" the confinement is radial, and the sealing action is on the I.D. and O.D. of the O-ring. Assembly of the components of a radial seal is difficult. Improper assembly may shear off part of the O-ring. Poor design of dynamic seals may result in short seal life. For example, when an O-ring is used as a valve nosepiece seal seating against a conical female seat the O-ring tends to rotate nonuniformly during closure. Repeated cycling leads to twisting and early seal failure.

Gas trapped in a hidden volume sealed to the outside, but with a slow leak to the system interior, causes what is known as a "virtual leak." The symptoms include slow pumpdown and pressure bursts. To avoid virtual leaks, O-ring grooves should be vented by a narrow slot extending from groove to flange interior.

In most cases O-rings will produce leak-free seals without the use of grease. Grease is often a harmful contaminant, and it aggravates virtual leak problems, particularly when used excessively. When grease is permitted, just enough grease to make the O-ring surface shiny is adequate for static seals. However, when O-rings are used in rotary or linear motion feedthroughs they must be well lubricated to prevent damage. O-ring motion feedthroughs should be restricted to applications above about 10^{-6} torr because of the exposed lubricated surfaces. Desorp-

tion characteristics of some vacuum greases have been described by Laurenson [21]. The Fomblin® perfluorinated polyether oils and greases [22] have very low vapor pressures and are chemically stable.

When fluoroelastomer O-rings are intended for use at pressures below 10^{-8} torr they may be prebaked in vacuum at about 150–200°C to reduce the outgassing of water and other volatile material contained within the elastomer [23]. Occasionally O-rings are cleaned with ethyl alcohol or fluorinated solvents. Baking is then necessary to remove the solvents completely. Metal seals are recommended when lower pressures and cleaner systems are needed.

4.6.3 METAL SEALS

4.6.3.1 History of Metal Seals for Vacuum Joining

The history of metal seals is closely related to that of UHV. The first UHV-compatible metal-sealed valves were designed by Alpert [24] at the Westinghouse Research Laboratories. The tubulations were small, about 12 mm, and the diameter of the valve orifice 6.3 mm. Bills and Allen [25] developed a small metal-sealed valve that had a negligible closed conductance. These valves were available commercially for many years [26].

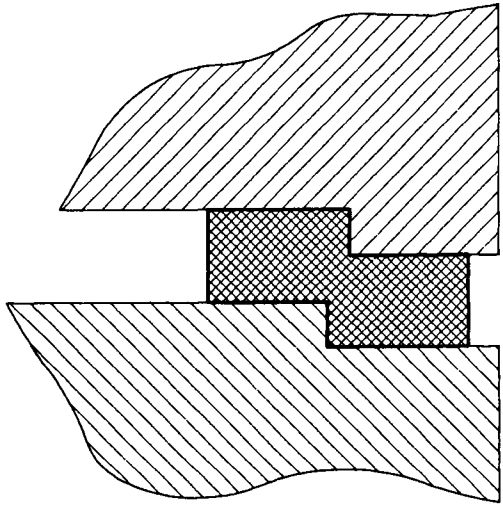
In the late 1950s the demand for large all-metal vacuum systems for space and plasma physics research led to the development of suitable components. In the United States, NASA supported extensive studies of seals and sealing surfaces [27,28,29].

The Stellarators for fusion research at Princeton Plasma Laboratory operated under UHV conditions using gold wire seals [30], and Souchet discussed gold wire seals used at CERN [31]. Gold wire seals are reliable, but the cost of the gold is significant, and in large sizes the fully annealed, approximately 1-mm-diameter gold wire is hard to handle.

Shear seal flanges using flat copper gaskets were first described by Lange and Alpert [32], and were made commercially until about 1962. The shear seal principle as illustrated in Figure 4 performed well, but the flanges were inconvenient because they were sexed. When the knife-edge flange design developed by Wheeler and Carlson [33,34] was released by Varian Associates as the ConFlat® flange [35], it quickly replaced all prior metal-sealed flanges. The seal detail is shown in Figure 5. Varian Asociates held a patent [36] on the particular knife-edge configuration used in the ConFlat®.

Fig. 4.

Seal detail of shear-seal union first described by Lange and Alpert.



A copper gasket is partially sheared between male and female flanges.

Fig. 5.

Seal detail of CF flange union showing a copper gasket captured between knife edges.

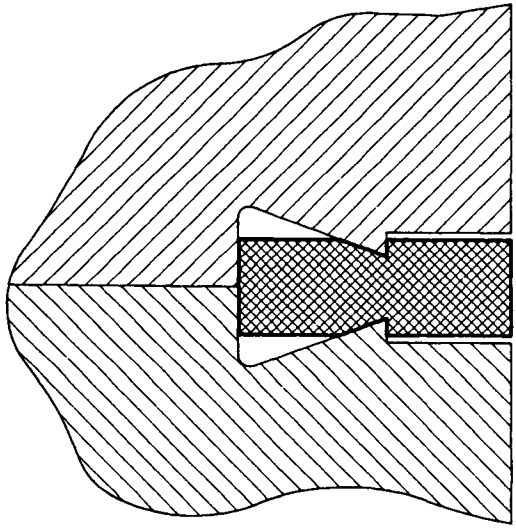
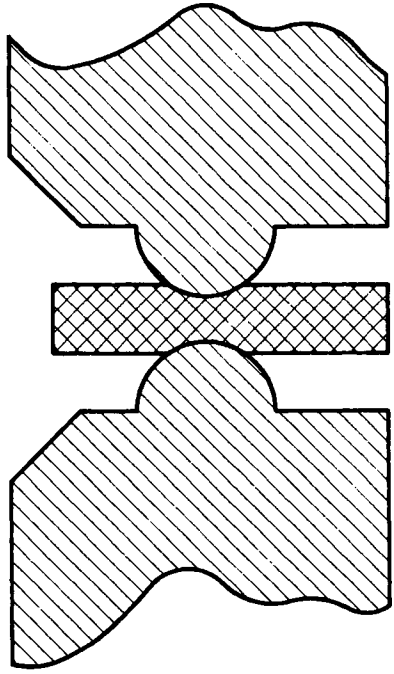


Fig. 6.
Seal detail of a small face seal fitting showing a copper gasket indented by radiused knife edges.



To escape this patent other manufacturers made flanges with the same bolt pattern, outside diameter, and knife-edge diameter, but with radiused knife-edges. They did not seal as reliably as the ConFlat®, and soon disappeared from the market. The radiused knife-edge design, sketched in Figure 6, remains in use on small face seal fittings available from various sources.

Although the patent has expired, the word ConFlat® remains a trademark. To avoid using this name, flanges of ConFlat® type are often called "CF flanges."

4.6.3.2 Stored Energy

If a union consisting of a flange pair and seal is to remain leak tight when subject to the differential motions caused by changing temperatures with the inevitable temperature gradients, it is necessary that some element of the union store energy elastically. In the case of seals using elastomers, the elastomer itself serves this function. In CF flanges the deformation of the flange pair helps to maintain the

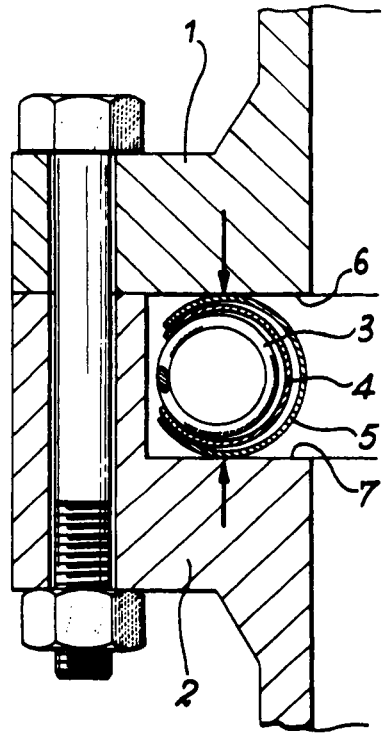
seal [37,38]. The problem of seal integrity is more difficult as CF flanges grow larger. Sixteen-inch CF flanges may not remain leak free after baking.

Slight changes are sometimes made in the details of CF flanges. In Europe, the vertical face of the knife-edge is often made with a small positive angle.

It is possible to make a metal seal assembly that is itself elastic. The Helicoflex® seal developed by CEFILAC [39,40] of St. Etienne, France, is an example of one such seal. Illustrated in Figure 7, it is a form of metallic O-ring consisting of a helical garter spring made endless by welding, a layer of hard metal to distribute the force from the discrete turns of the spring, and an outer layer of a softer seal metal such as aluminum, copper, or silver. If the cross section of the Helicoflex® seal is circular, the seal-to-flange contact area is larger than for a CF and a correspondingly greater seal force is needed for the same seal metal. To reduce

Fig. 7.

Helicoflex® seal developed by CEFILAC, St. Etienne, France.



The figure is from the patent drawing. In the figure (1) and (2) are the flanges, with seal surfaces (6) and (7). The components of the metal O-ring are the garter spring (3), a liner (4), and an outer layer of seal metal (5).

the seal force requirement, Helicoflex® "delta seals" are available with modified surfaces providing a reduced contact area [41]. One advantage of the Helicoflex® seal is that it can be made in shapes for use with non-circular flanges. Very large sizes are also possible. The outer layer of seal metal can be chosen to suit the application. A large number of Helicoflex® seals with silver outer layer were used on racetrack-shaped ports 0.22 × 1.6 m on the TFTR (Toroidal Fusion Test Reactor) Tokamak at the Princeton Plasma Laboratory [42].

Welch and co-workers [43] evaluated several types of metal seals for possible use in the AGS (Alternating Gradient Synchrotron) accelerator at Brookhaven National Laboratories, and chose the Helicoflex® "delta seal." Unfortunately, the Helicoflex® seal is not an ideal elastic element. The displacement vs. force curve (hysteresis curve) for compression/decompression of a seal has a very open loop [39].

4.6.3.3 Other Metal Seals

Many other metal seals are available commercially. Aluminum seals of diamond cross section for installation between flat-faced flanges as shown in Figure 8 are available commercially. Unterlerchner [44] reported the results of a study at CERN that used a large number of seals of this type. Sealing was reliable.

When CF flange knife-edges are damaged by careless handling, they may not seal with the standard gaskets. Fend [45] dscribed a seal for use with CF flanges as sketched in Figure 9. Seals are available in aluminum, copper, and nickel, and it is claimed that these gaskets require only about 50% of the seal force of a standard CF gasket of the same material.

Fig. 8.
Seal detail showing diamond cross section aluminum seal between flat-faced flanges.

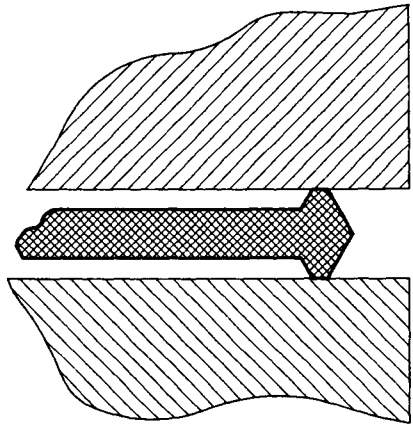
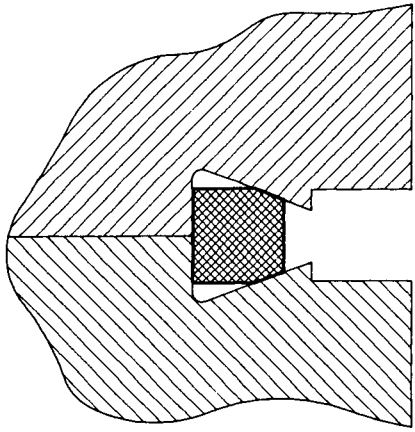


Fig. 9.
Seal detail of a union between standard CF flanges, but using a corner-crush gasket.



The firm of VAT VAKUUMVENTILE AG [46] sells a family of reusable seals with indium as the seal metal. They are intended for use with the ISO-KF flange system. Bakeout temperature is limited to about 75°C, but they can be a very effective way to eliminate elastomers from an existing system with KF flanges.

4.6.3.4 Theory of Metal Seals

To obtain a metal seal between a hard metal and a soft metal, it is essential that the soft metal be deformed plastically at the interface. Roth [47,48] investigated the mechanism of hard/soft sealing. Armand, Lapujoulade, and Paigne [49] compared theory and experiment for the leakage conductance of contacting surfaces with microirregularities.

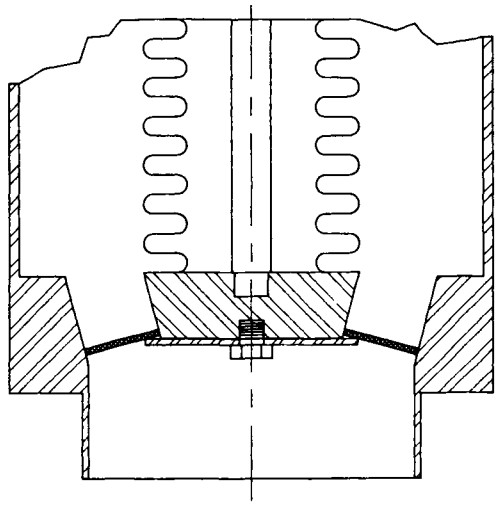
In addition, including shear in the plastic deformation is conducive to effective sealing. There are several ways to obtain shear of a seal. Spreading of a gasket during clamping between flat surfaces causes some shear of seal metal with respect to the flange surface. A slippage like displacement of a flange with respect to the seal causes shear. Perhaps most important, as the knife-edge of a CF flange enters the copper gasket there is considerable shear of the copper on the near-vertical surface. Small imperfections such as scratches on the copper gasket disappear in the steep portion of the imprint. This may be seen by intentionally scratching a gasket with a scribe, compressing it fully between a pair of CF flanges, removing the gasket, and inspecting it with a microscope [33,50]. The scratch will be visible on the shallow-angle portion of the imprint, but will disappear on the vertical. The disappearance of surface faults in the imprint is thought

to be a requirement for sealing. Exactly where the sealing occurs with a CF flange remains an interesting question. CF flanges also seal reliably when the flange faces and knife-edges have been glass bead blasted.

The situation is different when a radiused knife-edge is used. Today, such knife-edges are found in face seal fittings from several manufacturers intended to be used with tubing from $\frac{1}{4}$ to 1 inch in diameter. Making up a seal using a scratched gasket may not cause the scratch to fully disappear anywhere in the knife-edge indentation. One illustration of the poorer sealing properties of the radiused knife-edge is the fact that these fittings will not seal when the stainless surfaces are glass bead blasted. A good way to make a leak of about 10^{-6} to 10^{-5} atm cc/s is to bead-blast a pair of these fittings and assemble the joint.

Seals between hard metal pairs were considered theoretically by Inbar [51]. An interesting hard metal seal for use between flanges was described by Verheyden and Kline [52]. It is doubtful that two clean hard metal surfaces can be brought into such intimate contact with only elastic deformation that a leak-free seal would be the result. Figure 10 represents metal-sealed angle and gate valves [46] where the seal is said to be between hard–hard pairs without plastic deformation. However, one element of this seal is plated with a very thin film of silver to prevent cold welding. Perhaps this silver film serves to fill the minute voids remaining

Fig. 10.
Sketch of hard–hard seal construction as used by VAT VAKUUMVENTILE AG in metal-sealed valves.



The tapered seat, the nosepiece, and the conical washer are made of hard stainless steels. Note that the washer must seal on both its I.D. and O.D.

when the metal surfaces are brought into contact under high loading. Whatever the mechanism, sealing of these valves is reliable over a long cycle life.

A few molecular layers of adsorbed oil or water between the interfaces of a metal seal can cause an imperfect seal to appear to be leak free. This is known as a "liquid seal," and the effect is particularly annoying when encountered in precision leak valves where it causes the conductance to jump abruptly as the valve is opened. Because this is a common effect, and atmospheric water vapor is enough to cause it, a metal seal should be considered tight only when it has been baked and tested in a humidity-free environment.

4.6.3.5 Seal Metals

Many metals are used for vacuum sealing. In order of increasing seal load requirement, the most common include indium, lead, gold, aluminum, silver, copper, and nickel. Indium has a very low vapor pressure, and it even seals successfully to ceramics and glass. But the maximum operating temperature is limited to about 70°C, and indium is subject to cold flow. If a seal is to last, the indium must be contained. Lead has been used as wire or as plated coatings. Aluminum is used for several types of seals, but is not a perfect performer. Perhaps the oxide film prevents consistent metal to metal contact. That aluminum seals were not as reliable as copper was mentioned by van Heerden [50].

Silver is an excellent seal metal, but seals using it are expensive, and gold is even more so. Copper gaskets are commonly used with CF flanges. When copper gaskets are baked at 400°C, oxidation of the copper limits the number of bakeouts to 20 to 30. Wikberg [53] tested silver-plated copper gaskets at CERN and found that they would withstand at least ten 300°C bakeout cycles of 8 h duration without tarnishing or peeling of the silver. Copper is not acceptable in some chemical environments present in semiconductor device fabrication. Nickel gaskets are sometimes substituted in these cases. Even when nickel is fully annealed, the seal force requirement is so great that it is difficult to avoid breaking the CF flange bolts. A dry film coating of molybdenum disulphide is helpful on bolts used with CF flanges to obtain maximum clamping with minimum torque.

4.6.3.6 Review of Metal Sealing

Metal seals are required when the temperatures are high or low, when the outgassing and permeation that occur with polymer seals are not allowed, or in the presence of sufficient nuclear radiation to damage polymers.

CF-type flanges used with copper gaskets are satisfactory for most circular

flange purposes for diameters up to 12 inches. For larger sizes and noncircular shapes, Helicoflex® seals are about the only choice. However, they are expensive, may have long delivery times if made to order, and should be evaluated by use to gain experience before making a final decision to use them in a large project.

It is probable that development of metal seals will continue, driven by the trend to ever lower pressures for semiconductor device fabrication, and by the construction of large storage ring accelerators where both extremely low pressures and high reliability are important.

REFERENCES

1. Viton® is a registered trademark of DuPont Co., Wilmington, DE 19898.
2. Flourel® is a registered trademark of 3M Co., St. Paul, MN 55144.
3. *VITON Fluid Resistance Guide VT-500.1 (R2) 191281E* (DuPont, Wilmington, DE, December 1994).
4. *The General Chemical Resistance of Various Elastomers* (Los Angeles Rubber Group, Inc., Los Angeles, 1970).
5. *Standard Test Methods for Rubber Property—Compression Set,* ASTM D395-89 (American Society for Testing and Materials, Philadelphia 1989).
6. CHEMRAZ DATA SHEET HC 789 (Greene, Tweed & Co., Inc., Kulpsville, PA 1989).
7. M. H. Van de Voorde and C. Restat, *Selection Guide to Organic Materials for Nuclear Engineering*, CERN Report 72-7 (CERN, Geneva 17 May 1972).
8. B. B. Dayton, in *1959 Sixth Nat. Symp. Vac. Technol. Trans.* (Pergamon Press, New York, 1960), pp. 101–119.
9. P. W. Hait, *Vacuum*, **17** (1967) 547.
10. G. Thieme, *Vacuum*, **13** (1963) 137.
11. L. de Csernatony, *Vacuum*, **16** (1966) 129.
12. B. B. Dayton, *J. Vac. Sci. Technol.*, **A13** (1995) 451.
13. J. R. Young, *J. Vac. Sci. Technol.*, **6** (1969) 398.
14. E. DeLaurentis, S. Wu, J. Salfelder, and Y. Uritsky, *Microcontamination*, **3** (1994) 45.
15. A. W. Chesterton Co., Stoneham, MA 02180.
16. T. J. Edwards, J. R. Budge, and W. Hauptli, *J. Vac. Sci. Technol.*, **14** (1977) 740.
17. B. W. L. M. Sessink and N. F. Verster, *Vacuum*, **23** (1973) 319.
18. *Parker O-Ring Handbook* (Parker Seal Group, Lexington, KY, 1992), Sec. A2, p. 15.
19. *Parker O-Ring Handbook* (Parker Seal Group, Lexington, KY, 1992), Sec. A4, pp. 8–9.
20. *O-Ring Handbook* (Precision Rubber Products Corp., Lebanon, TN, January 1984), Sec. 2, p. 6.
21. L. Laurenson, *Vacuum* **27** (1976) 431.
22. Fomblin is a registered trademark of Montedison USA, New York.
23. L. de Csernatony, *Vacuum*, **16** (1966) 13.
24. D. Alpert, *Rev. Sci. Instru.*, **22** (1951) 536.
25. D. G. Bills and F. G. Allen, *Rev. Sci. Instru.*, **26** (1955) 654.
26. Granville-Phillips Co., Boulder, CO 80301.
27. F. O. Rathbun, Jr., ed., *Sealing Action at the Seal Interface*, NASA Report N63-18.159 (1963).
28. L. G. Gitzendanner and F. O. Rathbun, Jr., *Design Criteria for Zero Leakage Connectors for Launch Vehicles*, NASA Report N65-30841 (1965).

29. J. A. Bain and J. P. Laniewski, *Zero Leakage Design for Ducts and Tube Connections for Deep Space Travel*, NASA Report 68-20974 (1967).
30. D. J. Grove, *1958 Fifth Nat. Symp. Vac. Technol. Trans.* (Pergamon Press, New York, 1959), p. 9.
31. R. Souchet, *Le Vide*, No. 123 (1966) 235.
32. W. J. Lange and D. Alpert, *Rev. Sci. Instru.*, **28** (1957) 726.
33. W. R. Wheeler and M. Carlson, *1961 Trans. Eighth Nat. Vac. Symp.* (Pergamon Press, New York, 1962), p. 1309.
34. W. R. Wheeler, *1963 Trans. Tenth Nat. Vac. Symp.* (Macmillan, New York, 1963), p. 159.
35. Conflat® is a registered trademark of Varian Vacuum Products, Lexington, MA 02173.
36. M. A. Carlson and W. R. Wheeler, U.S. Patent 3,208,758 issued September 28, 1965.
37. T. Wikberg and E. Dodelin, *Vacuum*, **41** (1990) 2082.
38. K. Kitamura, K. Itoh, T. Uchida, M. Kondoh, K.Obara, K. Nakamura, and Y. Murakami, *J. Vac. Sci. Technol.*, **A12** (1994) 3217.
39. Helicoflex® is a registered trademark of CEFILAC S.A., St. Etienne, France.
40. C. Abbes, S. Bianchi, R. Chevallereau, M. Moreau, J. Rogement, R. Roques, C. Rouaud, and H. Royer, U.S. patent 3,820,799, issued June 28, 1974.
41. R. de Villepoix, J. Fages, R. Forges, C. Abbes, and C. Rouaud, U.S. Patent 4,561,662, issued 31 Dec. 1985.
42. R. B. Fleming, R. W. Brocker, D. H. Mullaney, and C. A. Knapp, *J. Vac. Sci. Technol.*, **17** (1980) 337.
43. K. M. Welch, G. T. McIntyre, J. E. Tuozzolo, R. Skelton, D. J. Pate, and S. M. Gill, *Vacuum*, **41** (1990) 1924.
44. W. Unterlerchner, *Vacuum*, **41** (1990) 1920.
45. H. Fend, *Vacuum*, **47** (1996) 527.
46. VAT VAKUUMVENTILE AG, Haag/SG, Switzerland.
47. A. Roth, *J. Vac. Sci. Technol.*, **9** (1972) 14.
48. A. Roth, *J. Vac. Sci. Technol.*, **A1** (1983) 211.
49. G. Armand, J. Lapujoulade, and J. Paigne, *Vacuum*, **14** (1964) 53.
50. P. J. van Heerden, *Rev. Sci. Instru.*, **26** (1955) 1130.
51. A. Inbar, *Vacuum*, **22** (1972) 187.
52. L. Verheyden and K. Klein, *Vacuum*, **21** (1971) 401.
53. T. Wikberg, *J. Vac. Sci. Technol.*, **5** (1968) 205.

CHAPTER 4.7

Outgassing of Materials

Benjamin B. Dayton
Consultant

4.7.1 RELATIONSHIPS AMONG SYSTEM PRESSURE, PUMPING SPEED, AND OUTGASSING

4.7.1.1 System Parameters That Must Be Known or Measured

NET PUMPING SPEED IN VACUUM CHAMBER

A typical vacuum plant or system consists of a main vacuum chamber, of known volume V and exposed wall area A of known material and surface condition, in which a process is to be carried out at a measured low pressure p maintained by a vacuum pump of rated speed S_p. If the pump is an oil diffusion pump, it may be provided with a suitable baffle or cold trap of known conductance, C_b, between pump and chamber to prevent vapor diffusing back into the chamber from the pump which also usually has internal means of reducing backmigration of oil vapor. If the pump is a turbomolecular pump or cryopump, no trap is necessary, but the turbo pump may have a protective screen at the inlet. A high-vacuum valve of known conductance, C_v, may be included to isolate the pump from the chamber during venting to atmosphere. In addition, it is sometimes necessary to take into consideration the added impedance, $1/C_p$, to gas flow caused by short lengths of pipe between the chamber and the valve housing, particularly when this pipe connection includes a right-angle bend. The net pumping speed S_n available at the connection to the chamber is then given by

$$1/S_n = 1/C_b + 1/C_v + 1/C_p + 1/S_p \qquad (1)$$

ISBN 0-12-325065-7
$25.00

where the conductances C_b, C_v, and C_p are assumed to have been measured or calculated without entrance or exit corrections, because the diameter of the pipe connections and the valve and baffle housings is usually about the same as the diameter of the intake opening to the pump. If the pump speed, S_p, has been measured with a test dome according to AVS (American Vacuum Society) or ISO (International Standards Organization) standard recommended practice, it will include the only entrance correction that need be made, because the location of the vacuum gauge is chosen so that the measured speed simulates that for gas flowing from a large entrance chamber across the pump inlet [1]. Of course, if no baffle or valve is used, the corresponding impedance terms are omitted from Equation (1). Pump speeds are usually expressed in liters per second and conductances in torr-liters per second per torr pressure drop.

4.7.1.2 Intrinsic or Free Outgassing Rate

The intrinsic or "free outgassing rate" (outgassing rate at pressures so low that readsorption is negligible), K_{fh}, of the chamber walls and other surfaces exposed inside the chamber at room temperature can be estimated from [2,3,4]

$$K_{fh} = K_{fl}/t_h^{\alpha} \tag{2}$$

where t_h is the pumping time in hours up to a limiting time, t_{hm}, (usually several hours) at which this simple equation cannot be used, α is usually constant and about equal to 1 for metal surfaces up to a time t_{hm}, which depends on the thickness of the oxide layer on the metal and about equal to 0.5 for elastomers and plastics up to a time t_{hm}, which depends on the thickness of the material, and K_{fl} is the free outgassing rate (in torr · liter $s^{-1}cm^{-2}$) after one hour of pumping, which is independent of the ratio of net pumping speed, S_n, to exposed outgassing area, A_m.

To convert outgassing rate in torr · liter $s^{-1}cm^{-2}$ to Pa m s^{-1} multiply by 1333.22. To convert from torr · liter $s^{-1}cm^{-2}$ at 25°C to molecules $s^{-1}cm^{-2}$, multiply by 3.24×10^{19}.

PROPERTIES OF THE MATERIALS EXPOSED TO THE VACUUM

The material of the chamber wall is usually stainless steel or aluminum, but may be mild steel or copper, and for these metals the value of K_{fl} depends primarily on the thickness and condition of the thin oxide film that always exists on the surface and on the partial pressure, P_e, of water vapor in the air to which the surface was exposed for a given time t_e prior to pump down. Values of K_{fl} as calculated from measurements of the outgassing rate of stainless steel and aluminum for various exposures to water vapor are given in Table 1.

Table I

Measured Outgassing Rate Constants for Stainless Steel and for Aluminum Alloy, and Calculated Free Outgassing Rate

Material and Surface Treatment	A_m/S_w	p_e	t_e	$10^9 K_{m1}$	α_1	$10^9 K_{f1}$	R N
A. Stainless steel							
1. 305 baked 24h 150°C	2410	18	0.33	0.7	1.3	623	[
2. 304 baked 48h 150°C							
a. Electropolished	717	15	1.0	8.3	1.11	1410	[:
b. Electropolished +He glow discharge	717	15	1.0	4.1	0.92	694	[:
c. Vacuum remelt/detergent	717	15	1.0	4.6	1.17	779	[:
d. Mill finish/detergent	717	15	1.0	4.9	1.14	830	[:
3. 304 first pumpdown							
a. Belt-polished	790	10?	>24	1.6	0.9	153	[:
b. Buff-polished	790	10?	>24	1.0	1.0	96	[:
4. 304 baked 24h 150°C, immersed in water	1198	23	24	0.8	1.2	88	[·
5. SS Kh18N9T							
a. Untreated	134	10?	>24	28	1.4	486	[:
b. Cord-brushed	134	10?	0.5	7.0	1.2	372	[:
c. Ultrasonic	134	10?	0.5	3.1	1.1	165	[:
d. Mechanical polished	134	10?	0.5	2.1	1.0	117	[:
e. Chemical polished	134	10?	0.5	1.8	1.0	96	[:
6. 304 highly polished, pumped to 8×10^{-10} torr 12°C							
a. First exposure 12°C	14	10	0.25	8.0	1.1	103	[(
b. Second exposure 12°C	14	1	0.25	4.3	1.2	119	[(
c. Third exposure 12°C	14	0.1	0.25	1.9	1.3	114	[(
d. Fourth exposure 12°C	14	0.01	0.25	0.8	1.35	103	[(
7. 304 No heat treatment, pumped to $<10^{-7}$ torr							
a. First exposure	2.36	18	1/60	40	1.33	873	['
b. Second exposure	2.36	18	1/12	80	1.2	783	['
c. Third exposure	2.36	18	1/2	160	1.1	655	['
d. Fourth exposure	2.36	18	3	400	1.0	761	['
e. Fifth exposure	2.36	18	19	600	0.8	805	['
f. Sixth exposure	2.36	18	48	600	0.8	799	['
8. 304 baked at 180°C, pumped to ultimate pressure							
a. First exposure	19	4×10^{-5}	0.5	0.80	0.8	558	[8
b. Second exposure	19	4×10^{-5}	0.17	0.34	0.8	399	[8
B. Aluminum alloy							
1. A6063-EX bak. 24h 150°C	2410	18	0.33	10	1.1	8890	[.
2. 6061 baked 48h 150°C							
a. Inert extrude (EX)TIG	717	15	1.0	7.5	1.12	1270	[:
b. Mirror finish	717	15	1.0	6.4	1.12	1080	[:
c. Mill fin./detergent	717	15	1.0	7.8	1.19	1320	[:

(continue

Table 1—Continued

Material and Surface Treatment	A_m/S_w	p_e	t_e	$10^9 K_{m1}$	α_1	$10^9 K_{f1}$	Ref. No.
3. A6063-EX immersed H_2O	2256	23	24	13	1.2	2670	[4]
4. A6063-EX argon glow discharge, chemical etch							
a. First exposure	2×10^4	9	0.25	0.47	1.1	4980	[9]
b. Second exposure	2×10^4	0.01	0.25	0.08	1.2	8180	[9]

Note:

A_m = exposed geometric area of outgassing material (cm^2)
S_w = pumping speed for water vapor at 25°C (L/s)
K_{m1} = measured outgassing rate at 1 h (torr $L/s^{-1} cm^{-2}$)
α_1 = absolute value of slope of log–log graph of measured outgassing rate vs. time after 1 h of pumping
p_e = partial pressure (in torr) of water vapor during exposure of preconditioned vacuum system for time t_e (in h) before pumpdown at room temperature
K_{f1} = calculated free outgassing rate at 1 h (torr $L/s^{-1} cm^{-2}$)

References

1. J. Chen et al., *J. Vac. Sci. Technol.*, **A5** (1987) 3422.
2. H. F. Dylla, D. M. Manos, and P. H. LaMarche, *J. Vac. Sci. Technol.*, **A11** (1993) 2623.
3. N. Yoshimura and H. Hirano, *J. Vac. Sci. Technol.*, **A7** (1987) 3351.
4. J. Chen and Y. Liu, *J. Vac. Sci. Technol.*, **A5** (1987) 262.
5. V. Zhilnin, L. Zhilnina, and A. Kuzmin, *Proc. 4th Internat. Vac. Congress, Manchester 1968*, Pt. 2 (Inst. Physics and Physical Soc., London, 1968), pp. 801–808.
6. H. J. Schuetze and K. E. Hennings—See B. Dayton, *J. Vac. Sci. Technol.*, **A13** (1995) 451.
7. H. Galron, *Vacuum*, **23** (1973) 177.
8. R. Strehlow and D. Richardson, *Proc. 4th Internat. Vac. Congress, 1968*, Pt. 1 (Inst. Physics and Physical Soc., London, 1968) pp. 175–178.
9. J. Chen et al., *J. Vac. Sci. Technol.*, **A12** (1994) 1750.

If the chamber wall is baked at temperatures of 150 to 200°C for a sufficient time during the pumpdown, most water adsorbed in the oxide layer can be removed, and the outgassing rate of the chamber wall then depends on the content of hydrogen, oxygen, and other gases that are dissolved in the underlying metal and that diffuse to the exposed surface, as well as on the desorption of residual water molecules chemisorbed in the oxide layer. Under these conditions, the exponent α is usually somewhat less than 1. To obtain ultra-high-vacuum pressures, it is necessary to clean the surface thoroughly and to remove most of the dissolved hydrogen by baking under vacuum at 300 to 500°C for several hours [5,6].

In addition to outgassing from the walls of the chamber, baffle housing, and vacuum valve, there will be outgassing from any elastomer gaskets used to seal the flange connections and the valve plate. The elastomer material must be known and the total exposed area estimated. For elastomers and plastics, the exponent α is usually about 0.5 except for silicone elastomer, for which α is about 1. Values of K_{f1} for elastomers and plastics are given by K_{m1} in Table 2.

Table 2

Outgassing Rate Constants for Elastomers and Plastics Calculated from Experimental Data

Material	A_m	S_a	$10^7 K_{m1}$	α_1	Ref. No.
Araldite B	30	0.7	18	0.4	[1]
Araldite D	30	0.7	19	0.3	[1]
Araldite F	30	0.7	15	0.5	[1]
Convaseal	65	3.7	14	0.5	[7]
Epicote	82	0.7	25	0.5	[2]
Eppon, Shell Oil Co.	41	0.4	40	1.2	[5]
Epoxy Resin (CIBA)	39	1	27	0.4	[14]
Hypalon	25	(0.1)	18	0.5	[8]
Kel-F	65	1.5	0.4	0.57	[5]
Methylmethacrylate	174	0.7	42	0.9	[2]
Natural rubber	25	(0.1)	63	0.5	[8]
Neoprene	65	3.7	330	0.5	[7]
Neoprene	25	(0.1)	31	0.5	[8]
Neoprene	—	1	70	0.3	[3]
Nylon	—	1	120	0.5	[3]
Perbunan	65	3.7	35	0.5	[7]
Perbunan	25	(0.1)	28	0.5	[8]
Pertinax	22.7	0.6	60	0.31	[6]
Permafil (G.E)	57	0.4	400	1.8	[5]
Plexiglass M222 (Röhm and Haas)	12	0.10	19	0.43	[3]
Plexiglass (Alsthom)	30	0.7	31	0.4	[1]
Plexiglass	8	0.6	7	0.43	[6]
Polyacrylate	25	(0.1)	132	0.5	[8]
Polyamid (Bayer)	12	0.12	46	0.5	[4]
Polyethylene (BASF)	12	0.10	2.4	0.5	[4]
Polyethylene (Dynamit-AG)	12	0.23	2.6	0.5	[4]
Polystyrol (Dynamit-AG)	12	0.14	15	0.5	[4]
Polystyrol (BASF)	12	0.14	6	0.5	[4]
Polystyrol	110	0.6	6	0.59	[6]
Polyurethane (Bayer)	12	0.85	5	0.5	[4]
Polyvinylcarbazol (BASF)	12	0.23	16	0.5	[4]
PTFE	—	1	3	0.45	[3]
Silastic	—	1	30	1.3	[3]
Silicone (74)	25	(0.1)	66	1.00	[8]
Silicone (54)	25	(0.1)	83	0.63	[8]
Silicone (R 60, 1.7 mm thick)	14	0.18	20	0.75	[10]
Silicone (R 30, 1.4 mm thick)	14	0.18	40	1.0	[10]
Silicone (R 30, 4 mm thick)	14	0.18	80	1.0	[10]
Teflon, Dupont	12	0.23	4	0.7	[4]
Teflon, Dupont	65	1.5	5	0.68	[5]
Teflon	32	0.6	0.7	0.48	[6]
Teflon	25	(0.1)	3.3	0.68	[8]
Ultramid (BASF)	12	0.1	17	0.5	[4]
Viton	25	(0.1)	0.13	0.51	[8]
Viton A	10^4	300	0.19	0.5	[9]

(continues)

Table 2—Continued

Note:

A_m = exposed geometric area of outgassing material (cm^2)
S_a = pumping speed for air at 25°C (L/s)
K_{m1} = measured outgassing rate after 1 h of pumping (torr $L/s^{-1}cm^{-2}$)
α_1 = absolute value of slope of log–log graph of measured outgassing rate vs. time after 1 h of pumping

References

1. R. Geller, *Le Vide*, **13** (74) (1958) 71.
2. J. C. Boulassier, *Le Vide*, **14** (80) (1959) 39.
3. B. D. Power and D. J. Crawley, *Trans. 1st Internat. Vac. Congress, Namur*, Vol. 1 (Pergamon Press, New York, 1960), pp. 206–211.
4. R. Jaeckel and F. J. Schittko, *Forschungsberichte des Wirtschafts- und Verkehrsministerium Nordrhein-Westfalen* Nr. 369, Westdeutscher Verlag, Cologne, 1957.
5. B. B. Dayton, *Trans. 6th Nat. Vac. Symp.* (Pergamon Press, New York, 1960), pp. 101–119.
6. G. Thieme, *Vacuum*, **13** (1963) 137.
7. J. Blears, E. Greer, and J. Nightingale, *Trans. 1st Internat. Vac. Congr., Namur*, Vol. 2 (Pergamon Press, New York, 1960), pp. 473–479.
8. W. Beckmann, *Vacuum*, **13** (1963) 349.
9. B. B. Dayton, *J. Vac. Sci. Technol.*, **A13** (1995) 451.
10. H. von Münchausen and F. Schittko, *Vacuum*, **13** (1963) 549.
11. A. Schram, *Le Vide*, **18** (103), (1963) p. 55.
12. F. J. Schittko, *Vacuum*, **13** (1963) 525.
13. R. Jaeckel, *Trans. 8th Nat. Vac. Symp. and 2nd Internat. Vac. Congress*, Vol. 1 (Pergamon Press, New York, 1962), pp. 17–26.
14. B. H. Colwell, *Vacuum*, **20** (1970) 481.

There may also be some permeation of gases in the atmosphere through elastomer seals between flanges as well as permeation of hydrogen through some metals due to reaction of the outer surface with water, particularly for metals (such as iron) that do not form a protective coating [7].

The rate of permeation in torr · liters per s per cm^2 of exposed wall or gasket area of gas at the partial pressure P_n and absolute temperature T outside the enclosure through an elastomer gasket or through a glass or metal wall of thickness w_m (in mm) into a vacuum is given by

$$K_u = (760/273) \times 10^{-3}\, T(U_{nm}/w_m)P_n^{1/j}, \tag{3}$$

where U_{nm} is the permeability coefficient, which varies with the absolute temperature T_m of the mth material according to

$$U_{nm} = a_{nm} \exp(-10^3\, E_{nm}/jRT_m) \tag{4}$$

where the preexponential factor, a_{nm}, is nearly constant over a wide temperature range and has the permeability units of cm^3(STP)/cm^2 per torr pressure differ-

ential per mm thickness of a sheet of the material, E_{nm} is the activation energy for permeation in kcal/mol, j is the dissociation number ($j = 1$ for most gases through plastics and elastomers and $j = 2$ for diatomic molecules permeating metal), and R is the gas constant in calories per degree K per mole ($R = 1.987$). (STP is the standard temperature of 273 K and pressure of 760 torr.)

In the absence of leaks in the welds and residual pump fluid vapor in the system, the ultimate pressure is determined by permeation through gaskets and walls according to

$$p_u = K_u A_u / S \tag{5}$$

where A_u is the effective area through which permeation occurs and S is the net pumping speed for the permeating gases.

Table 3 lists the constants for permeation and diffusion of some gases through nonmetallic materials. Table 4 lists values of the permeability coefficient, U_{nm}, for gases through nonmetallic materials at room temperatures. Additional information can be found in the references cited.

Table 3

Constants for Permeation and Diffusion of Gases Through Nonmetallic Materials

Material	Gas	E_{nm}	a_{nm}	H_{nm}	d_{nm}	Ref. No.
Rubber (natural)	H_2	6.9	6.0×10^{-4}	5.9	0.23	[1]
Rubber (natural)	O_2	6.6	1.1×10^{-3}	7.5	0.57	[1]
Rubber (natural)	N_2	9.3	6.2×10^{-3}	8.7	2.9	[1]
Rubber (natural)	He	6.5	1.8×10^{-4}	—	—	[1]
Neoprene G	H_2	8.1	1.2×10^{-3}	6.6	0.28	[1]
Neoprene G	O_2	9.9	7.1×10^{-3}	9.4	3.1	[1]
Neoprene G	N_2	10.6	7.4×10^{-3}	10.3	9.3	[1]
Perbunan	N_2	11.0	1.4×10^{-2}	10.2	7.2	[1]
Perbunan	He	7.0	1.6×10^{-4}	—	—	[1]
Pyrex	He	6.2	2.6×10^{-8}	6.5	4.8×10^{-4}	[2]

Note:

E_{nm} = activation energy for permeation in kcal/mol
H_{nm} = activation energy for diffusion in kcal/mol
a_{nm} = permeation preexponential factor in cm^3(STP)/cm^2 per torr pressure differential per mm of material thickness
d_{nm} = diffusion preexponential factor in cm^2/s
Temperature range: 10–100°C

References

1. G. J. van Amerongen, *J. Appl. Phys.*, **17** (1946) 972.
2. W. A. Rogers, R. S. Buritz and D. Alpert, *J. Appl. Phys.*, **25** (1954) 868.

Table 4

Permeability Coefficient, U_{nm}, for Gases Through Nonmetallic Materials in $cm^3(STP)s^{-1}cm^{-2}$ for a Pressure Gradient of 1 torr/mm

Material	Gas	Temp. °C	U_{nm}	Ref. No.
Buna-S	N_2	20–30	6.35×10^{-10}	[1]
Buna-S	O_2	20–39	1.72×10^{-9}	[1]
Buna-S	H_2O	20–30	2.4×10^{-7}	[1]
Butyl rubber	N_2	20–30	3.2×10^{-11}	[1]
Butyl rubber	O_2	20–30	1.3×10^{-10}	[1]
Butyl rubber	H_2O	20–30	4×10^{-9}–2×10^{-8}	[1]
Hycar	N_2	20–30	2.4×10^{-11}–2.5×10^{-10}	[1]
Hycar	O_2	20–30	9.6×10^{-11}–8.2×10^{-10}	[1]
Hycar	H_2O	20–30	1×10^{-7}	[1]
Kel-F	N_2	20–30	9×10^{-13}–1.3×10^{-11}	[1]
Kel-F	O_2	20–30	2.5×10^{-12}–5.4×10^{-11}	[1]
Kel-F	H_2O	20–30	3×10^{-11}–3.6×10^{-9}	[1]
Mylar	H_2O	20–30	1.3×10^{-8}–2.3×10^{-8}	[1]
Neoprene	N_2	20–30	1.18×10^{-10}	[1]
Neoprene	N_2	27.1	1.27×10^{-10}	[2]
Neoprene	O_2	20–30	4.0×10^{-10}	[1]
Neoprene	H_2O	20–30	1.8×10^{-7}	[1]
Neoprene	H_2	18.2	9.0×10^{-10}	[2]
Pyrex	He	20	5.6×10^{-13}	[2]
Rubber	air	25	4.2×10^{-10}	[2]
Rubber	H_2O	25	2.5×10^{-7}	[2]
Teflon FEP	N_2	20–30	2.2×10^{-10}	[1]
Teflon FEP	O_2	20–30	5.9×10^{-10}	[1]
Teflon FEP	H_2O	20–30	5.0×10^{-9}	[1]
Viton-A	N_2	20–30	4.4×10^{-11}	[1]
Viton-A	O_2	20–30	1.5×10^{-10}	[1]
Viton-A	H_2O	20–30	5.2×10^{-9}	[1]

References

1. A. Lebovits, *Modern Plastics*, **43** (March 1966) 139.
2. R. M. Barrer, *Diffusion in and Through Solids* (Cambridge University Press, London, 1941).

4.7.1.2 Relationships Among System Pressure, Pumping Speed, Outgassing Rates, and Chamber Dimensions

GENERAL EQUATION FOR THE PRESSURE IN THE CHAMBER

Assuming that there are n different volatile constituents that can outgas from the m different materials exposed in the vacuum chamber, the principle of material

balance results in the equation

$$p_n S_n + p_n \Sigma f_n e_{nm} A_m = \Sigma K_{nm} A_m - V(dp_n/dt) \tag{6}$$

where p_n is the partial pressure of the nth gas in torr, S_n is the net pumping speed for this gas in liters/sec, A_m is the exposed area in cm^2 of the mth material, K_{nm} is the "free" or intrinsic outgassing rate in torr · liter $s^{-1} cm^{-2}$ of this material for the nth gas, defined as the outgassing rate when readsorption is negligible, t is the time in seconds, V is the chamber volume in liters, e_{nm} is a sticking or sorption coefficient for the nth gas on the mth material, and

$$f_n = (R_0 T/2\pi M_n)^{1/2} \cdot 10^{-3} \tag{7}$$

is the "effusion law" factor in units of liters $s^{-1} cm^{-2}$ at the absolute temperature T for the nth gas of molecular weight M_n and R_0 is the molar gas constant ($R_0 = 8.31 \times 10^7$ ergs $deg^{-1} K\ mol^{-1}$).

The total pressure p in the vacuum chamber at the pumping time t in seconds is given by [2]

$$p = \sum_n p_n = \sum_n \frac{\sum_m K_{nm} A_m - V(dp_n/dt)}{\sum_m f_n e_{nm} A_m + S_n} \tag{8}$$

where it is assumed that the partial pressure, p_n, of each gas species is known by measurement with a mass spectrometer or residual gas analyzer (RGA). Since outgassing rate measurements on individual materials are often made with total pressure gauges, such as an ionization gauge, using the air or nitrogen calibration factor, in practice the preceding equation can be simplified to

$$p = \frac{\sum_m K_m A_m - V(dp/dt)}{\sum_m f_a e_m A_m + S} \tag{9}$$

where p is the total pressure as measured with the air calibration factor, K_m is the air- (or nitrogen-) equivalent free outgassing rate at the time t, e_m is the air-equivalent sorption coefficient, f_a is the value of f_n when the molecular weight M_n is taken as 29 (or 28), and S is the net pumping speed for air. When the chamber is empty of material to be processed and the outgassing from elastomer gaskets can be neglected, this equation simplifies further to

$$p = \frac{K_{fh} A}{f_a \sigma A + S} - \frac{V}{f_a \sigma A + S} \frac{dp}{dt} \tag{10}$$

where σ is the effective air-equivalent sorption coefficient for the wall material (of geometric surface area A), which may be somewhat larger than the ordinary sticking coefficient because of the porosity of the oxide layer and will depend on the fraction of adsorption sites that are occupied at time t; A is the chamber wall area; and

$$K_{fh} = K_{fl}/(t/3600)^{\alpha} \tag{11}$$

where K_{fl} is the free outgassing rate of the wall material after one hour of pumping and t is in seconds. This equation contains the hidden factor (one hour)$^{\alpha} = 1$, which allows for values of α other than 1.

RELATIONSHIP BETWEEN REQUIRED SPEED AND CHAMBER DIMENSIONS

The volume V in liters of a cylindrical vacuum chamber of length L (cm) and radius R (cm) with dished heads can be estimated from

$$V = 10^{-3}\,\pi R^2(L + 0.4R) \tag{12}$$

where L is usually about $3R$ corresponding to $V = R^3/100$. The area (in cm^2) of the walls of such a chamber can be estimated from

$$A = 2\pi R(L + 1.1R) \tag{13}$$

and for $L = 3R$ the area

$$A = 26\,R^2 \tag{14}$$

For typical industrial vacuum systems filled with material to be processed (such as vacuum coating) the net air speed, S_d, (in liters/sec) of the high-vacuum pump (diffusion pump or turbomolecular pump) should be about 2.5 times the volume V of the empty chamber, or

$$S_d = 2.5\,V = R^3/40 \tag{15}$$

and the roughing pump or backing pump speed, S_r, is usually about $S_d/100$ or

$$S_r = V/40 = R^3/4000 \tag{16}$$

For typical industrial vacuum systems with chambers filled with material to be processed,

$$S_d/A = R/1000 \tag{17}$$

where A is the wall area of the empty chamber.

4.7.2 INITIAL PUMPDOWN FROM ATMOSPHERIC PRESSURE

During the initial pumpdown from atmosphere of the empty chamber with a mechanical pump (roughing pump) of net speed S_r, the outgassing and sorption terms can be neglected until the pressure produced by the roughing pump alone falls below a value given by the semiempirical formula

$$p_v = 300\, K_1\, A/V \tag{18}$$

where K_1 is the net outgassing rate (including readsorption) at one hour and the numerical coefficient has the dimensions of time. This pressure p_v is typically in the range from 10^{-3} to 10^{-2} torr when the ultimate pressure of the roughing pump is less than 10^{-3} torr. In the region from atmospheric pressure, p_0, (about 750 torr) to p_v integration of Equation (10), neglecting the outgassing and sorption terms, with a roughing pump of maximum speed S_r and ultimate pressure p_{ru}, gives

$$p = p_{ru} + p_0 \exp(-S_r t/V) \tag{19}$$

and the time (in s) to reach the pressure p will be

$$t_v = 2.3\, (V/S_r) \log_{10}[(p_0 - p_{ru})/(p - p_{ru})] \tag{20}$$

Then the time to reach the pressure p_v will be approximately

$$t_v = 12\, V/S_r \tag{21}$$

For $L = 3R$ the ratio A/V in Equation (18) equals $2600/R$ so that

$$p_v = 7.8 \times 10^5\, K_1/R \tag{22}$$

Since K_1 is often of the order of 10^{-7} torr · liter $s^{-1} cm^{-2}$, then for R of about 30 cm (1 foot) p_v is about 3×10^{-3} torr. Since oil diffusion pumps can operate when the forepressure is less than a few tenths of 1 torr, there is no need to wait until the pressure falls to p_v to switch the high-vacuum pump into operation by opening the high-vacuum valve. From Equation (20) and Equation (16), the roughing time, or time to reach the switchover pressure, p_s, for diffusion-pumped systems of about 0.2 torr will typically be

$$t_v = 8\, V/S_r = 320\, s = 5.3 \text{ minutes} \tag{23}$$

For cryopumped systems, the safe switchover pressure depends on the volume of the chamber, and the recommendations of the manufacturer should be followed.

4.7.3

PRESSURE VS. TIME DURING OUTGASSING

4.7.3.1 Pumpdown with Outgassing

After switching to operation with a high-vacuum pump of constant speed $S_d = 100\ S_r$, the time to reduce the pressure from p_s to $p_d = 10^{-2}p_s$ (neglecting outgassing) will be

$$t_{vd} = 4.6\ V/S_d \tag{24}$$

or, since these systems typically have a time constant $\tau = V/S_d = 0.4$ s, the pressure will drop from $p_s = 0.2$ torr to $p_d = 2 \times 10^{-3}$ torr in about 2 s. When the pressure reaches $p_v S_r/S_d = 3 \times 10^{-5}$ torr, the pressure then decreases more slowly as outgassing, and inleakage become the limiting processes. The term $-[V/(f_a \sigma A + S)](dp/dt)$ can usually be neglected when the pressure decreases by another factor of 3, or typically when the pressure with the high-vacuum pump operating is less than 1×10^{-5} torr. The pressure is then given by

$$p = K_{fh}A/(f_a \sigma A + S_d) + p_u = (K_{fl}/t_h^{\alpha})/(f_a \sigma + S_d/A) + p_u \tag{25}$$

where p_u is the "ultimate pressure" (lower pressure limit) due to leaks and permeation from the atmosphere, t_h is the pumping time in hours, and σ is the sticking probability based on the geometric surface area A. At room temperature (298 K) the value of f_a (for air) is 11.7 liters $s^{-1}cm^{-2}$. For empty chambers with metal walls that have not been baked, the gas at pressures below 10^{-5} torr is about 90% water vapor. The effusion law factor, f_w, for water vapor at 298 K is 14.8 liters $s^{-1}cm^{-2}$.

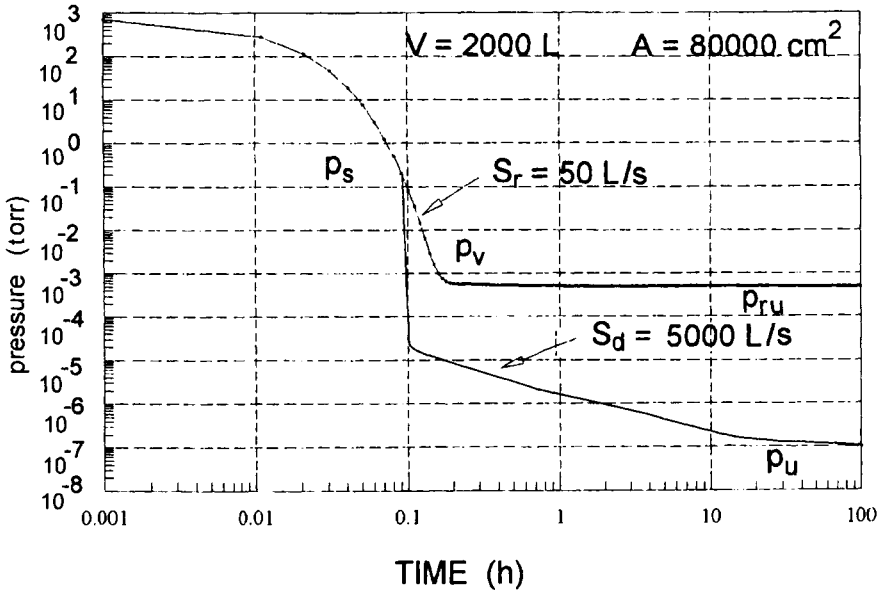
For stainless steel, as shown below, it can be assumed that σ is about 6×10^{-3} after one hour of pumping, so that for chambers with $S_d/A < 1$ (or $R < 1000$ cm) the readsorption term $f_w \sigma = 0.089$ for water vapor cannot be neglected in comparison to S_d/A.

The preceding relationships are illustrated by Figure 1 for the pumpdown curve of a typical vacuum system with $V = 2000$ liters, $S_r = 50$ liters/sec, A = 80,000 cm^2, $S_d = 5000$ liters/sec, $K_{fl} = 1 \times 10^{-7}$ torr · liter $s^{-1}cm^{-2}$, $p_0 = 750$ torr, $p_s = 0.2$ torr, $p_v = 10^{-3}$ torr, and ultimate pressure $p_u = 1 \times 10^{-7}$ torr with the diffusion pump operating as shown by the dashed line.

4.7.3.2 Rate of Rise of Pressure in Closed Chamber

If the high-vacuum valve is closed after the time t_0 (in seconds), then $S = 0$ and, using Equation (11), the differential equation, Equation (10), for the pressure

Fig. 1.



Curves of pressure vs. time during pumpdown in presence of outgassing.

becomes

$$dp/dt + cp = b/(t + t_0)^{\alpha} \tag{26}$$

where

$$c = f_a \sigma A/V \tag{27}$$

$$b = (3600)^{\alpha} K_{fl} A/V \tag{28}$$

and t is the time in seconds after closing the valve. Assuming σ independent of pressure and time, this equation can be integrated to give

$$p = p_0 e^{-ct} + be^{-ct} \int_0^t \frac{e^{ct'}}{(t' + t_0)^{\alpha}} dt' \tag{29}$$

where the integral must be evaluated numerically and

$$p_0 = (K_{fl} A/t_{h0}{}^{\alpha})/(S + f_a \sigma A) \tag{30}$$

in which t_{h0} is the total pumping time in hours at which the valve was closed. When the pressure rise time in seconds $t << t_0$,

$$p = p_0 e^{-ct} + (b/ct_0{}^{\alpha})(1 - e^{-ct}) \tag{31}$$

where

$$b/c = (3600)^{\alpha} K_{f1}/f_a\sigma \tag{32}$$

From Equation (31) the rate-of-rise equation is

$$dp/dt = [(b/t_0^{\alpha}) - cp_0]e^{-ct} \tag{33}$$

and the initial ($t = 0$) rate-of-rise times the chamber volume V divided by the chamber wall area A is

$$(V/A)\, dp/dt = (K_{f1}/t_{h0}^{\alpha})/(1 + f_a\sigma A/S) \tag{34}$$

which is just the net (measured) outgassing rate (including readsorption) at the pumping time (in h) t_{h0} when the valve was closed. When $S/A >> f_a\sigma$, one obtains the free outgassing rate at t_{h0}. This is one method of measuring the outgassing rate. The equations are based on the assumption that the time exponent α is constant over the times involved in the measurements, which is usually approximately true for typical values of α and times up to t_{hm} as defined in Section 4.7.1.2 (p. 485) and Section 4.7.4.1 (p. 500). Of course, when the system contains layers of oil or solids with volatile plasticizers, the evolution of vapors from these sources will not follow Equation (31).

4.7.4 THE OUTGASSING RATE OF ELASTOMERS AND PLASTICS

4.7.4.1 Outgassing by Diffusion of Absorbed Gas to Exposed Surface

It has been found experimentally [2] that the outgassing rate of elastomers and plastics, once the volatile organic matter (plasticizer) has been removed, is determined mainly by the diffusion of absorbed gas and water molecules from the bulk of the material to the exposed surface, where the molecules can escape directly from open pores or move by surface diffusion to adsorption sites on the exposed surface from which they soon evaporate. Nitrogen and oxygen molecules initially adsorbed on the directly exposed (geometric) surface are desorbed in less than a second at room temperature because the energy of adsorption on elastomers and plastics is less than 10 kcal/mol. Similarly, water molecules on the exposed surface are desorbed quickly from most elastomers and plastics (but not from silicone elastomers). However, air and water molecules that have dissolved in the bulk material by diffusion through capillary pores before evacuation are only removed slowly at room temperature because of the small diameter of the capillar-

ies and the "random walk process," which is the nature of diffusion along a capillary pore.

The diffusion of gas into or out of a solid is controlled by Fick's law [8]

$$\frac{\partial}{\partial x}\left(D_1\frac{\partial C}{\partial x}\right)+\frac{\partial}{\partial y}\left(D_2\frac{\partial C}{\partial y}\right)+\frac{\partial}{\partial z}\left(D_3\frac{\partial C}{\partial z}\right)=\frac{\partial C}{\partial t} \tag{35}$$

where C is the concentration (in molecules cm^{-3}) of gas in the solid at the location (x, y, z), D_1, D_2, and D_3 are diffusion coefficients (in $cm^2 s^{-1}$), which may vary with the location (x, y, z) and the direction of the concentration gradient, and t is the time (in s). For material in sheet form of sufficient thickness, w_m, that it can be treated as a semiinfinite solid with a uniform concentration C_0 at time zero and a constant diffusion coefficient D_{nm} for the nth gas, when the exposed surface is suddenly exposed to a very low pressure by fast pumping of the vacuum chamber containing the material, the concentration distribution along the x axis perpendicular to the surface at the time t will be given by

$$C(x, t) = C_0 \operatorname{erf}[x/2(D_{nm}t)^{1/2}] \tag{36}$$

where erf refers to the error function, which is tabulated in various mathematical handbooks [9].

For gas molecules that are adsorbed on the walls of capillary pores in a solid with adsorption lifetime, τ_a, and that diffuse along the capillary by desorption followed by collision with the wall and readsorption, the diffusion coefficient is given approximately by [2]

$$D_{nm} = (\lambda^2/4)/[(\lambda/u_{nm}) + \tau_\alpha] \tag{37}$$

where λ is the mean distance (in cm) in the x direction traveled by a molecule bewteen collisions with the capillary wall and u_{nm} is the average molecular velocity at the temperature T_m of the solid.

The free outgassing rate in molecules $s^{-1} cm^{-2}$ is then given by

$$N(t) = D_{nm}[\partial C(x, t)/\partial x]_{x=0} \tag{38}$$

where the concentration gradient is evaluated at $x = 0$ corresponding to the exposed surface and $t > \pi\lambda^2/16D_{nm}$. Substituting Equation (36) in Equation (38) gives

$$N(t) = C_0(D_{nm}/\pi t)^{1/2} \tag{39}$$

and the free outgassing rate in torr · liter $s^{-1} cm^{-2}$ is then

$$K_{fh} = (R_v T/60N_a)C_0(D_{nm}/\pi t_h)^{1/2} \tag{40}$$

where $R_v = 62.36$ torr · liter K^{-1} is the molar gas constant in vacuum engineering units, T is the absolute temperature of the gas in the vacuum chamber, $N_a = 6.02 \times 10^{23}$ is Avogadro's number (the number of gas molecules in one mole),

Table 5

Diffusion Coefficients (in cm^2/s) for Gases Through Nonmetallic Materials

Material	Gas	Temp. °C	D_{nm}	Ref. No.
Neoprene	H_2	27	1.8×10^{-6}	[1]
Neoprene	N_2	27	1.9×10^{-7}	[1]
Perbunan	H_2	25	4.1×10^{-6}	[2]
Perbunan	N_2	25	2.5×10^{-7}	[2]
Pyrex	He	20	4.5×10^{-11}	[1]
Rubber (vulcanized)	H_2	25	8.5×10^{-6}	[1]
Rubber (vulcanized)	N_2	25	1.5×10^{-6}	[1]
Rubber (vulcanized)	O_2	25	2.1×10^{-6}	[1]
Rubber (vulcanized)	CO_2	25	1.1×10^{-6}	[1]

References

1. R. M. Barrer, *Diffusion in and Through Solids* (Cambridge University Press, London, 1941).
2. G. J. van Amerongen, *J. Appl. Phys.*, **17** (1946) 972.

and t_h is the pumping time in hours but cannot be less than $(\pi\lambda^2/16D_{nm})/3600$. Thus, for elastomers and plastics where the outgassing rate is controlled by diffusion of a single molecular species with constant adsorption lifetime τ_a, the value of α in Equation (2) is about $\frac{1}{2}$ and the free outgassing rate at one hour depends on the initial molecular concentration C_0 according to

$$K_{fl} = (R_v T/60N_a)(D_{nm}/\pi)^{1/2} C_0 \tag{41}$$

where the diffusion coefficient varies with the absolute temperature, T_m, of the solid according to

$$D_{nm} = d_{nm} \exp(-H_{nm}/jRT_m) \tag{42}$$

Values of the diffusion coefficient, D_{nm}, for gases through nonmetallic materials are given in Table 5 and values of the constants d_{nm} and H_{nm} are listed in Table 3. The solubility coefficient (in cm^3(STP)/cm^3 per torr) for the nth gas in the mth material at the temperature T_m is given by

$$s_{nm} = 0.1(a_{nm}/d_{nm})\exp[-(E_{nm} - H_{nm})/jRT_m] \tag{43}$$

because it can be shown that

$$U_{nm} = 10 s_{nm} D_{nm} \tag{44}$$

The initial gas concentration C_0 (in molecules/cm^3) in Equation (41) will be a sum of the concentrations due to the solution in the material of gases in the atmosphere during exposure. If the exposure time is long so that equilibrium is

reached at the prevailing temperature, the concentration of the nth gas will be

$$C_{n0} = (760/273) \times 10^{-3}(N_a/R_v)s_{n0}P_{n0}^{1/j} \quad (45)$$

where s_{n0} is the value of the solubility coefficient s_{nm} for the nth gas at the initial pressure P_{n0} and temperature $T_m = T_0$. Of course, in addition to the atmospheric gases dissolved in the solid elastomer or plastic there will usually be some residual plasticizer, but the vapor pressure of the plasticizer may be so low that the outgassing rate is determined mainly by the dissolved gas content. Table 6 lists values of the solubility coefficients for common gases in typical solids.

Since experiments have shown [2] that the sorption probability, e_{nm}, for the common gases colliding with the exposed surface of an elastomer and diffusing some distance into the bulk through the entrance to an open capillary pore, is less than 1×10^{-7} so that $f_n e_{nm} A_m$ is negligible compared to S_n when $S_n/A_m > 1 \times 10^{-4}$, then K_{f1} is approximately equal to the net outgassing rate from elastomers in most systems.

Equation (39) should be valid until $t_h > t_{hm}$, where

$$t_{hm} = [\pi(w_m/20)^2/16D_{nm}]/7200 \quad (46)$$

in which w_m is the thickness (in mm) of a sheet of elastomer exposed to the vacuum on both sides [2]. When $t_h > t_{hm}$, the outgassing rate should begin to decay more or less exponentially with pumping time as the gas content of the elastomer becomes more than half depleted.

Table 6

Solubility Coefficients for Common Gases in Typical Materials (in units of cm^3(STP)/cm^3 per torr)

Material	Gas	Temp. °C	s_{nm}	Ref. No.
Iron (α)	H_2	300	2.9×10^{-4}	[3]
Neoprene G	H_2	25	3.8×10^{-5}	[2]
Neoprene G	N_2	25	4.7×10^{-5}	[2]
Neoprene G	O_2	25	9.9×10^{-5}	[2]
Neoprene	A	36	2.0×10^{-4}	[1]
Perbunan	H_2	25	3.7×10^{-5}	[2]
Perbunan	N_2	25	4.6×10^{-5}	[2]
Perbunan	O_2	25	1.0×10^{-4}	[2]
Pyrex	He	500	1.1×10^{-5}	[1]
Rubber (natural)	N_2	25	6.8×10^{-5}	[2]
Rubber (natural)	O_2	25	1.3×10^{-4}	[2]

References

1. R. M. Barrer, *Diffusion in and Through Solids* (Cambridge University Press, London, 1941).
2. G. J. van Amerongen, *J. Appl. Phys.*, **17** (1946) 972.
3. C. J. Smithells (ed.), "Metals Reference Book," 5th ed. (Butterworths, London, 1976).

4.7.4.2 Experimental Data on Outgassing of Elastomers and Plastics

Measured values of the outgassing rate K_{m1} and the exponent α_1 after one hour of pumping are listed in Table 2 for common elastomers and plastics. Also listed are the area A_m (in cm^2) of the outgassing material and the net pumping speed, S_a, for air at 25°C in the chamber containing the outgassing material. Although the ratio S_a/A_m was smaller than 0.01 in some experiments, the values of K_{m1} can be considered to be the free outgassing rate because the sorption coefficient for solution of gas in the bulk is sufficiently low, as noted earlier.

Data are also available on the percentage of different gases evolved obtained by mass spectrometer measurement for some materials [10,11], but in general the main constituent is water vapor, together with some nitrogen, oxygen, and carbon dioxide, when the material has been exposed for some time to the atmosphere and not baked under vacuum. Carbon monoxide sometimes shows up in the mass spectrometer, but this may be due to interaction of H_2O with carbon-contaminated hot filaments in the gauges [12].

4.7.5 THE OUTGASSING RATE OF METALS AND CERAMICS

4.7.5.1 Experimental Data on Outgassing of Metals and Ceramics

Measured values of the outgassing rate K_{m1} and the exponent α_1 after one hour of pumping are listed in Table 7 for common metals and ceramics. Also listed is the area A_m (in cm^2) of the outgassing material and the net pumping speed, S_a, for air at 25°C (usually limited by an orifice) in the test chamber containing the outgassing material. Since the outgassing of unbaked metals is mainly due to evolution of water vapor from the oxide layer on the surface, and the sorption probability for water vapor in the oxide layer in general varies from about 10^{-4} to 10^{-2} depending on the fraction of adsorption sites occupied and surface heterogeneity, the listed values of K_{m1} cannot be considered as free outgassing rate constants but may depend on the ratio A_m/S_w, where S_w is the net pumping speed for water vapor at 25°C, as well as the partial pressure, p_e, of water vapor in the atmosphere and the time of exposure, t_e, before pumpdown.

Table 1 gives reported outgassing rates at one hour, K_{m1}, for aluminum alloy and stainless steel for various values of A_m/S_w, p_e and t_e as well as the free outgassing rate (where readsorption would be zero) calculated from the semiempiri-

Table 7

Outgassing Rate Constants for Metals and Ceramics Calculated from Experimental Data

Material	A_m	S_a	$10^9 K_{m1}$	α_1	Ref. No.
Alundum (901 Norton)	39	1	400	1	[14]
Aluminum (spray-coated mild steel)	65	0.4	60	0.75	[5]
Copper	12	0.08	20	1	[4]
Degussite (Al_2O_3)	12	0.1	75	1.1	[13]
Graphite wool (Morganite)	39	1	1000	1	[14]
Iron	12	0.07	5	1	[4]
Mild steel (slightly rusty)	—	1	600	3.1	[3]
Molybdenum	12	0.12	14	1	[4]
Mullite (800, United Fireclay)	39	1	800	1	[14]
Mullite (Morganite)	39	1	620	1	[14]
Nickel	12	0.08	10	1	[4]
Porcelain, glazed	30	0.7	6.5	0.5	[1]
Pyrex	10^3	1	7.2	1.1	[11]
Pyrophyllite	12	0.1	20	1.2	[13]
Pyrophyllite (fired)	39	1	1500	0.8	[14]
Sillimanite (Zirconal)	39	1	500	0.6	[14]
Silver	12	0.07	50	1	[4]
Stainless steel (ICN sanded)	10^3	1	8.3	1.2	[11]
Stainless steel (ICN fresh)	10^3	1	13.5	0.9	[11]
Stainless steel (NS22S)	10^3	1	14.4	1.3	[11]
Stainless steel (NS22S electropolish)	10^3	1	4.3	1.0	[11]
Steatite (Al_2O_3)	30	0.7	90	1	[1]
Steel (chromium plated, fresh)	10^3	1	7.1	1	[11]
Steel (nickel plate)	10^3	1	2.8	1.1	[11]
Steel (niclel plated, fresh)	10^3	1	4.2	0.9	[11]
Steel (chromium plated, polished)	10^3	1	9.1	1	[11]
Tantalum	12	0.07	9	1	[4]
Tungsten	12	0.07	7	1	[4]
Zirconia (Zirconal Ltd.)	39	1	1200	1	[14]
Zirconium	12	0.15	8	1	[4]

Note:

A_m = exposed geometric area of outgassing material (cm^2)
S_a = pumping speed for air at 25°C (L/s)
K_{m1} = measured outgassing rate after 1 h of pumping (torr $L/s^{-1} cm^{-2}$)
α_1 = absolute value of slope of log–log graph of measured outgassing rate vs. time after 1 h of pumping

References

1. R. Geller, *Le Vide*, **13**(74) (1958) 71.
2. J. C. Boulassier, *Le Vide*, **14**(80) (1959) 39.
3. B. D. Power and D. J. Crawley, *Trans. 1st Internat. Vac. Congress, Namur*, Vol. 1 (Pergamon Press, New York, 1960), pp. 206–211.

(*continues*)

Table 7—Continued

4. R. Jaeckel and F. J. Schittko, *Forschungsberichte des Wirtschafts- und Verkehrsministerium Nordrhein-Westfalen*, Nr. 369. Westdeutscher Verlag, Cologne, 1957.
5. B. B. Dayton, *Trans. 6th Nat. Vac. Symp.* (Pergamon Press, New York, 1960), pp. 101–119.
6. G. Thieme, *Vacuum*, **13** (1963) 137.
7. J. Blears, E. Greer, and J. Nightingale, *Advances in Vacuum Science and Technology*, vol. 2, edited by E. Thomas (Pergamon Press, New York, 1960), pp. 473–479.
8. W. Beckmann, *Vacuum*, **13** (1963) 349.
9. B. B. Dayton, *J. Vac. Sci. Technol.*, **A13** (1995) 451.
10. H. von Münchausen and F. Schittko, *Vacuum*, **13** (1963) 549.
11. A. Schram, *Le Vide*, **18** (103) (1963) 55.
12. F. J. Schittko, *Vacuum*, **13** (1963) 525.
13. R. Jaeckel, *Trans. 8th Nat. Vac. Symp. and 2nd Internat. Vac. Congress*, vol. 1 (Pergamon Press, New York, 1962), pp. 17–26.
14. B. H. Colwell, *Vacuum*, **20** (1970) 481.

cal equations

$$K_{fl} = K_{ml}[1 + (A_m/S_w)f_w\sigma]/(2\theta_{op}f_e) \tag{47}$$

where $f_w = 14.8$ liters $s^{-1}\,cm^{-2}$ for water vapor at 25°C, the sticking probability σ is arbitrarily set equal to 6×10^{-3} as a constant average value [13], and at 25°C

$$\theta_{op} = (5.21 \times 10^{-3}\, p_e)^{.33} \tag{48}$$

is the fraction of the available sites for physisorption of H_2O in a precursor state that would be actually occupied by H_2O in equilibrium, according to a Freundlich isotherm, with the partial pressure p_e at 25°C ($\theta_{op} = 0.5$ at the saturation pressure 23 torr), prior to dissociation and chemisorption, and at 25°C

$$f_e = [1 - \exp(-t_e/4.46)]^{.5} \tag{49}$$

gives the fraction of the saturation coverage θ_{op} attained after an exposure to the partial pressure p_e for t_e hours according to a semiempirical equation based on the data of Galron [14]. The numerical constants in Equations (48) and (49) will be different at temperatures other than 25°C.

When $(A_m/S_w)f_w\sigma >> 1$, the rate of readsorption is so high that the outgassing process is near the state of equilibrium corresponding to an adsorption–desorption isotherm. Since it can be shown that the activation energy for desorption of H_2O from the oxide layer varies with the fractional coverage θ in a manner similar to that for a Temkin isotherm, several authors [15,16,17] have modeled the outgassing process in terms of a quasi-equilibrium approach to a Temkin isotherm, either the simple isotherm for a uniform surface or the more complicated isotherm equation for nonuniform surfaces [18]. These models predict different dependence of outgassing rate on temperature, but only limited experimental data [19] exist with which to determine the correct model.

The measured value K_{m1} for metals exposed to moist atmosphere prior to pumpdown should preferably be reported using the net pumping speed, S_w, for water vapor, but in most cases the reported value appears to be for the air pumping speed, S_a. If the pressure is measured with an ion gauge using the air calibration factor, the indicated pressure, p_{aw}, will be lower than the true pressure, p_w, because the gauge sensitivity for H_2O is about 0.75 times the sensitivity for air (or N_2) and H_2O usually constitutes about 90% of the total pressure due to outgassing from an unbaked sample. Since $S_a/S_w = 0.79$, the product $p_{aw}S_a = 0.6\, p_w S_w$ approximately. The measured outgassing rate $K_{mh} = p_n S_n/A_m$, and the reported rate using air calibration will thus be somewhat less than the true outgassing rate. However, the behavior of a vacuum system in which pressure is measured with ion gauges using the air calibration factor can still be predicted using the outgassing rate tables.

From Table 1 it is evident that the free outgassing rate for aluminum alloy is an order of magnitude greater than that for stainless steel, which is presumably due to the greater thickness of the oxide layer on the former [20]. Although in general the outgassing rate for water vapor is likely to increase with the thickness of the oxide layer, various attempts to reduce the outgassing rate, in systems that cannot be baked, by reducing the surface roughness and oxide layer thickness by polishing and detergent cleaning have not been entirely successful, but *in situ* glow discharge cleaning with He gas reduces the outgassing rate by an order of magnitude [21,22].

4.7.6

OUTGASSING RATE OF PRECONDITIONED VACUUM SYSTEM AFTER SHORT EXPOSURE TO THE ATMOSPHERE

4.7.6.1 Formula for Outgassing Rate During Second Pumpdown

If the outgassing rate varies with time t_I (in h) as

$$K_I = K_1/t_I^{\alpha} \tag{50}$$

during the first pumpdown where α is nearly constant up to a time $t_m > 1$ h, then we can designate the outgassing rate during the second pumpdown of a previously evacuated system by K_{II} and successive pumpdowns by appropriate Roman numerals, while the subscript 1 always identifies the state of a variable after one hour of pumping on the first pumpdown. The dimensions of K_1 may be taken as torr liter $s^{-1} cm^{-2}$ provided it is realized that the preceding equation is

a simplification of the formula

$$K_1 = K_1 t_1^{\alpha}/t_1^{\alpha} \tag{51}$$

where $t_1 = 1$ h, so that in Equation (50) K_1 is actually multiplied by a factor 1 having dimensions $(\text{h})^{\alpha}$.

For many materials that have not been previously degassed, the value of α is between 0.3 and 1.0 when $t_1 < t_m$. For most plastics and elastomers (except silicone rubber) the value of α is approximately 0.5 corresponding to the ordinary law of desorption limited by the rate of diffusion of gas (usually mostly water vapor) through the solid. For silicone rubber, unpolished metals, glass, and ceramics, the initial outgassing (due primarily to water vapor) corresponds to an exponent α that is usually in the neighborhood of 1.0.

When a solid sorbent is not altered by excessive heating or sintering and the gas content is not too high, the process of physical adsorption and absorption is often simply the reverse of desorption. In the absence of externally generated force fields, the diffusion coefficient for a gas through a solid is not changed when the direction of the concentration gradient is reversed. Therefore, it is to be expected that the sorption rate of a partially degassed elastomer for water vapor, nitrogen, and oxygen when exposed to the atmosphere at a fixed relative humidity will vary with time in a manner similar to the variation of the outgassing rate of these gases with time.

When the sorption and desorption rates depend primarily on the diffusion of water vapor and are reversible in the above sense, then it can be shown by an application of Duhamel's theorem [23] that the outgassing rate, K_{II}, during the second pumpdown, for materials that obey Equation (50) on the first pumpdown, is given by the equation

$$K_{II} = \frac{K_1}{(t_{II} + t_c + t_e)^{\alpha}} + \frac{K_{12}}{(t_{II})^{\alpha}} - \frac{K_{12}}{(t_{II} + t_e)^{\alpha}} \tag{52}$$

where t_{II} is the time (in h) from the beginning of the second pumpdown, t_c is the conditioning time (total time in hours of the first pumpdown), t_e is the time in hours during which the surface is exposed to atmosphere having a constant partial pressure, P_1, of water vapor after the first pumpdown, and as a first approximation it is assumed that

$$K_{12} = (P_1/P_0)^n K_1 \tag{53}$$

where n is a constant that depends on the adsorption isotherm, P_0 is the partial pressure of water vapor during exposure before the first pumpdown, and the sorption and desorption of nitrogen and oxygen is neglected. Equations (52) and (53) seem to hold for outgassing from plastics and elastomers where $\alpha = \frac{1}{2}$ and $n = 1$ (Henry's law), but experimental data on the outgassing of water molecules from

the oxide layer on metals (for which α is in the neighborhood of 1 and n is less than 1 corresponding to a Freundlich isotherm) indicate that Equation (52) is not obeyed. The expected outgassing rate and system pressure on a second pumpdown for metal chambers that have been preconditioned by long pumping (with or without baking) and then exposed to the atmosphere for a time t_e can be estimated from

$$K_{m1} = K_{f1}(2\theta_{op} f_e)/[1 + (A_m/S_w) f_w \sigma] \tag{54}$$

using Equations (48) and (49), together with appropriate values of K_{f1} and α_1 from Table 1.

Equation (52) is only valid when the venting and roughing times of the vacuum chamber are small compared to t_e and t_c, the material is maintained at constant temperature, and the time t_{II} is within the limits indicated by $0.01 < t_{II} < t_I < t_m$, where

$$t_I = t_{II} + t_e + t_c \tag{55}$$

When the conditioning time, t_c, is long compared to t_{II} and the exposure time is brief so that $100\, t_e < t_{II} < 10^{-3}\, t_c$, Equation (52) reduces to

$$K_{II} = K_{12}\, \alpha\, t_e/(t_{II})^{\alpha+1} \tag{56}$$

which can only be applied to elastomers and plastics with α about 0.5. For elastomers with $\alpha = \frac{1}{2}$, Equation (55) gives a slope of $-\frac{3}{2}$ when plotted on log–log graph paper as shown by Rogers [24]. When $t_{II} + t_c > t_m$, the slope can increase to values less than -2. Log–log plots of an equation similar to Equation (52) with $\alpha = \frac{1}{2}$ are given in Roger's article.

A general method of computing outgassing rate for elastomers and plastics after several pumpdown and exposure cycles is given in Reference [23].

4.7.7

METHODS OF DECREASING THE OUTGASSING RATE

4.7.7.1 Selection of Gasket Materials

If an elastomer gasket must be used, the preferred materials are fluorinated hydrocarbons such as Kalrez or Viton-A, which if necessary can be baked at 150°C. The exposed surface of gasket material can be minimized by proper design of the gasket grooves in the flanges, but the effective exposed area depends on how the gasket fits in the groove and usually is greater than the inside gap between the flange surfaces. Metal gaskets and suitably designed flanges must be used on

ultra-high-vacuum systems both because of the lower outgassing rate of metals and the ability to bake the system without damaging the gaskets.

4.7.7.2 Recommended Bakeout Temperatures and Times

Systems with metal walls and gaskets and not containing substances such as plastics, brass, cadmium-plated metal, etc., which are damaged by temperatures over 200°C can be baked under vacuum for 24 hours at 200–250°C to remove most of the water vapor adsorbed in the oxide layer on the metal walls [25]. The residual outgassing will then usually be due to hydrogen dissolved in the metal. To lower the outgassing rate for hydrogen appreciably it is necessary to bake at 450–500°C for 12 to 24 hours [5]. The fabrication of the metal chamber from vacuum degassed steel only reduces the outgassing rate by a factor of about 2. Pyrex glass walls can be baked up to 450°C.

4.7.7.3 Cleaning Methods and Surface Treatment for the Interior Walls of the Vacuum Chamber

The oxide layer on the metal walls can be reduced in thickness by glass bead bombardment [26] (which may introduce undesired contaminants on the walls) or mechanical polishing. Organic contaminants and water vapor can be partially removed by vapor degreasing or by the use of glow discharges or electron bombardment under vacuum and by flushing with dry nitrogen [27]. Surface roughness can be reduced somewhat by electropolishing and chemical etching. Nitrile or polyethylene gloves should be used when necessary to handle parts to be processed in the vacuum chamber, and air conditioning to lower the atmospheric humidity is desirable. Venting the vacuum system with dry nitrogen, and using a positive pressure of dry nitrogen inside a covering of plastic sheeting with entry flaps while removing and inserting parts to be processed, will reduce the adsorption of water vapor on the chamber walls. Parts to be processed should be stored in sealed containers with dessicants.

4.7.8 MEASUREMENT OF THE OUTGASSING RATE OF MATERIALS

Various methods for measuring the outgassing rate have been summarized by Elsey [28]. Recommended practices for measuring and reporting outgassing rates

have been published by the American Vacuum Society [29]. Frequently, reported outgassing rates have been much lower than the "free" outgassing rate because of the use of a small ratio of pumping speed to sample area, resulting in appreciable readsorption of gas. The reported rates are thus of little value unless S/A and the sticking coefficient σ are known. The direct molecular beam method of measurement [30] avoids this problem but may give incorrectly high values due to noncosine-law distribution of molecules in the intercepted beam.

REFERENCES

1. B. B. Dayton, *Vacuum*, **15** (1965) 53.
2. B. B. Dayton, *Trans. 6th Nat. Vac. Symp.* (Pergamon Press, New York, 1960), pp. 101–119.
3. B. B. Dayton, *Trans. 8th Nat. Vac. Symp.* (Pergamon Press, New York, 1962), pp. 42–57.
4. T. Kraus, *Trans. 5th Nat. Vac. Symp.* (Pergamon Press, New York, 1959), pp. 38–40; *Trans. 6th Nat. Vac. Symp.* (Pergamon Press, New York, 1960), pp. 204–205.
5. R. Calder and G. Lewin, *Brit. J. Appl. Phys.*, **18** (1967) 1459–1472.
6. R. Barton and R. Govier, *Vacuum*, **20** (1970) 1–6.
7. F. J. Norton, *J. Appl. Phys.*, **11** (1940) 262; *J. Appl. Phys.*, **28** (1957) 34.
8. R. M. Barrer, *Diffusion in and Through Solids* (Cambridge University Press, London, 1941).
9. M. Abramowitz and I. A. Stegun (eds.), *Handbook of Mathematical Functions* (Dover, New York, 1972).
10. R. Jaeckel and F. J. Schittko, *Forschungsberichte des Wirtschafts-und Verkehrsministerium Nordrhein-Westfalen* Nr. 369. Westdeutscher Verlag, Cologne, 1957.
11. G. Thieme, *Vacuum*, **13** (1963) 137.
12. J. R. Young, *J. Appl. Phys.*, **30** (1959) 1671.
13. B. B. Dayton, unpublished paper.
14. H. Galron, *Vacuum*, **23** (1973) 177.
15. P. A. Redhead, *J. Vac. Sci. Technol.*, **A13** (1995) 467.
16. G. Horikoshi, *J. Vac. Sci. Technol.*, **A5** (1987) 2501.
17. K. Kanazawa, *J. Vac. Sci. Technol.*, **A7** (1989) 3361.
18. D. Hayward and B. Trapnell, *Chemisorption*, 2nd ed. (Butterworths, Washington, 1964).
19. A. Glassford, R. Osiecki, and C. Liu, *J. Vac. Sci. Technol.*, **A2** (1984) 1370.
20. N. Yoshimura, H. Hirano, T. Sato, I. Ando, and S. Adachi, *J. Vac. Sci. Technol.*, **A9** (1991) 2326.
21. H. Dylla, D. Manos, and P. LaMarche, *J. Vac. Sci. Technol.*, **A11** (1993) 2623.
22. M. Li and H. F. Dylla, *J. Vac. Sci. Technol.*, **A13** (1995) 71.
23. B. Dayton, *J. Vac. Sci. Technol.*, **A13** (1995) 451.
24. K. W. Rogers, *Trans. 10th Nat. Vac. Symp.* (Macmillan, New York, 1963), pp. 84–87.
25. Y. E. Strausser, *Proc. 4th Internatl. Vac. Congress. (1968)* (Inst. Physics and Physical Soc., London, 1969), p. 469.
26. J. Amoignon and J. P. Couillaud, *Le Vide*, No. 141 (May–June 1969), pp. 181–189.
27. A. Berman, *Vacuum*, **47** (1996) 327.
28. R. J. Elsey, *Vacuum*, **25** (1975) 347.
29. AVS Standards Committee, *J. Vac. Sci. Technol.*, **2** (1964) 314.
30. S. Komiya and Y. Sugiyama, *J. Vac. Sci. Technol.*, **16** (1979) 689.

CHAPTER 4.8

Aluminum-Based Vacuum Systems

James L. Garner
SMC Corporation

The many advantages of aluminum for vacuum applications have driven industry to overcome most of its technical disadvantages, making it a practical alternative to stainless steel in vacuum system design. Since the properties and techniques for stainless steel are well understood, these are used as a departure point for comparing and contrasting with those for aluminum.

The physical properties of aluminum differ from those of stainless steel in many ways. Usually, one is of such overriding importance that it drives the selection of aluminum. Sufficient information is provided here to make such engineering judgments possible.

Except where noted otherwise, the equation variables in this section can be in any consistent unit system. Sample values are given in convenient engineering units but, in some cases, must be converted to a consistent unit system when used with the other variables in the sample calculations.

4.8.1 OUTGASSING

Properly cleaned aluminum surfaces outgas at rates comparable to, or better than, those of stainless steel. Outgassing in unbaked and baked regimes must be discussed separately, since the mechanisms differ.

ISBN 0-12-325065-7
$25.00

4.8.1.1 Unbaked System Outgassing

In unbaked systems, desorbing moisture dominates the gas load and comes off clean aluminum at about the same rate as from stainless steel [1]. A comparison of this performance is shown in Table 1. Outgassing rates decline in approximate inverse proportion to pumping time and, after 100 minutes, remain within about a 3:1 range for both aluminum and stainless steel on all the various surface finishes tested. Indeed, comparisons of published data suggest that moisture sorbs and desorbs from virtually all nonporous surfaces at similar rates [2]. While outgassing is proportional to effective surface area, this area is relatively independent of surface roughness, as defined in conventional engineering terms such as R_a [3]. The thick, porous scale on aluminum received from a mill can create large effective areas and produce high outgassing rates. However, conditioning of such surfaces is relatively easy as described in Section 4.8.3.

Table 1

Outgassing Rates of Stainless Steel and Aluminum with Various Finishes After Exposure to Air for One Hour

Material/Treatment	Q_0	n	$Q(t = 100)$ torr·L/cm^{2}sec
Stainless Steel:			
Literature Average	1.8×10^{-7}	1.0	1.84×10^{-9}
Electropolished (Summa Method® with phosphoric acid commonly used in vacuum components)	7.82×10^{-7}	1.11	4.72×10^{-9}
Compound electropolished (<0.02μm roughness using simultaneous mechanical and electropolishing — Dr. Ohmi, Tohoku University.)	3.73×10^{-7}	1.16	1.71×10^{-9}
Detergent-Cleaned, vacuum remelted SS304 (3× lower C than 304L)	5.49×10^{-7}	1.17	2.55×10^{-9}
Electropolished/high-temperature baked combined process (Proprietary — Quantum Mechanics Corp., Sonoma, CA)	7.88×10^{-7}	1.27	2.29×10^{-9}
Mill Finish (3μm roughness as received) with Alconox® detergent cleaning	5.27×10^{-7}	1.14	2.75×10^{-9}
Aluminum:			
EX-Processed 6063 (extruded in dry Ar/O_2 atmosphere)	7.73×10^{-7}	1.12	4.25×10^{-9}
Mirror Processed 6063 (0.02μm roughness using ethanol-cooled, diamond-tip cutting tool)	6.38×10^{-7}	1.12	3.62×10^{-9}
Mill Finished (3μm roughness as received) with Alconox® detergent cleaning	1.02×10^{-6}	1.19	4.25×10^{-9}

Source: Adapted from H. F. Dylla et al., Correlation of outgassing of stainless steel and aluminum with various surface treatments, *J. Vac. Sci. Technol.*, **A11**(5) (September–October 1993) 2629.

Note: Outgassing rate: $Q = Q_0 t^{-n}$ torr·L/cm^{2}sec, where t = time in minutes.

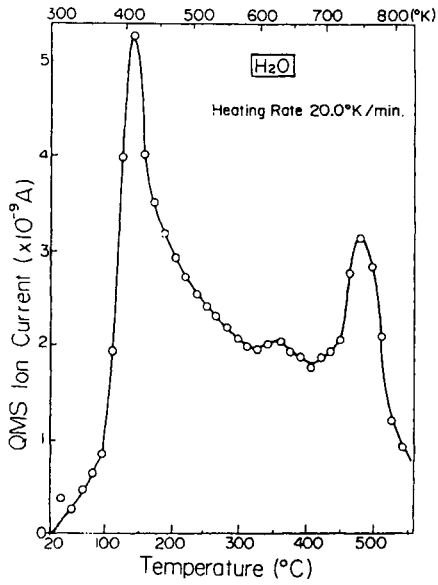
4.8.1.2 Baked System Outgassing

Moisture can be rapidly removed by baking a chamber. The thermal degassing spectrum in Figure 1 shows that substantial degassing starts at about 80°C and is very high at 150°C [4]. It is believed that the moisture desorption rate of aluminum in this temperature range is higher than that from stainless steel at even higher temperatures and adequate for practical systems. Baking temperatures above 150°C should be avoided, because aluminum alloy strength falls and its tempered properties begin to diminish above 170°C. Of course, lower strength and annealed properties can be accounted for in the design if high temperature operation is necessary.

In fully baked-out systems, adsorbed moisture is virtually eliminated, leaving hydrogen as the dominant gas load. The principal mechanism for the release of hydrogen is believed to be permeation from pockets inside the bulk metal and from the exterior surface. At room temperature, hydrogen permeation rate through aluminum is approximately seven orders of magnitude lower than through stainless steel [5]. Therefore, aluminum normally has a much lower UHV outgassing rate than stainless steel, typically on the order of 10^{-13} to 10^{-14} torr·L/sec·cm^2 [6–12]. These rates have been reported approachable with stainless steel only using elaborate vacuum melting and high-temperature baking processes. The sim-

Fig. 1.

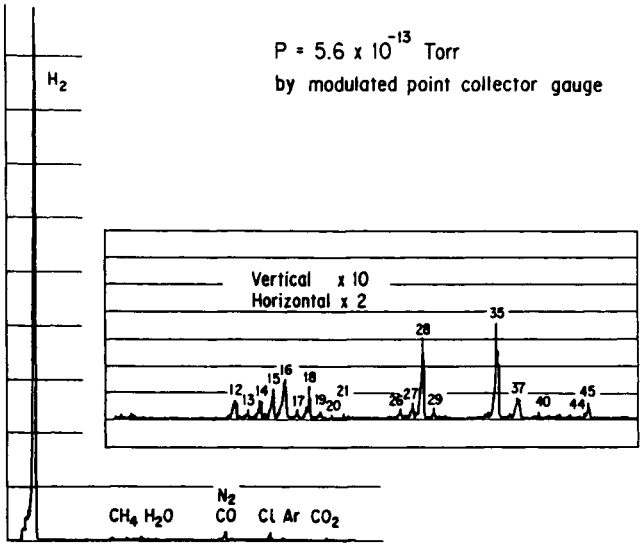
Thermal desorption spectra of moisture on EX processed aluminum 6063.



(Courtesy of M. Mohri et al., "Surface study of Type 6063 aluminum alloys," *Vacuum*, **34**(6) (1984) 644.

Fig. 2.

Residual gas spectrum of an aluminum UHV chamber.



[Courtesy of: H. Ishimaru, "Developments and Applications for All-Aluminum Alloy Vacuum Systems," *MRS Bulletin*, **15**(7) (July 1990).]

ilarity of hydrogen outgassing rates reported for various surface treatments of aluminum [13] suggests that it is determined by the basic characteristics of aluminum rather than any particular surface oxide formation or morphology. Pure aluminum is reported to have a somewhat lower outgassing rate than its alloys [14,15]; however, it must be clad to a high-strength alloy for structural integrity. In the past, voids and vias formed during the manufacturing process sometimes created outgassing problems. However modern, commercially available alloys rarely exhibit these problems.

Hydrogen diffuses through metals in its elemental form [16] which is chemically active and believed to react with carbon and oxides on the surface to produce secondary background gases such as CO and hydrocarbons. In aluminum chambers, these gasses are much less prevalent (Figure 2) [17], probably because there is much less carbon in common alloys of aluminum than in even low-carbon 316L stainless steel.

4.8.2

DEMOUNTABLE SEALS

Elastomer sealing techniques for aluminum chambers are identical to those for stainless steel and are widely used and understood. Metal seals have been more

difficult to implement because of the relative softness of aluminum and its high rate of thermal expansion. However, today a number of commercially available products have practically alleviated these problems. Metal seals are, of course, highly desirable, because O-rings tend to dominate the gas load in a chamber. For example, the outgassing from a single 100 mm × 5.3 mm Viton® O-ring is approximately equal to that from 3 ft^2 of metal chamber surface after 1 hour of pumping and 7 ft^2 after 10 hours. Pressures below 10^{-7} torr with elastomer seals are very costly to achieve, in terms of pumping speed required, unless special measures are taken such as differential pumping or baking of the O-rings [18].

4.8.2.1 Theoretical Model of Seal Effectiveness

The basic sealing mechanism for a metal seal involves forcing soft gasket material into the microscopic grooves on a hard sealing surface through which gas can flow (see Figure 3). As sealing pressure is increased, gasket material is pressed deeper into the grooves reducing the cross section (hence the conductance) of the leak paths. The pressure required to achieve a desired leak rate may be estimated as follows [19]:

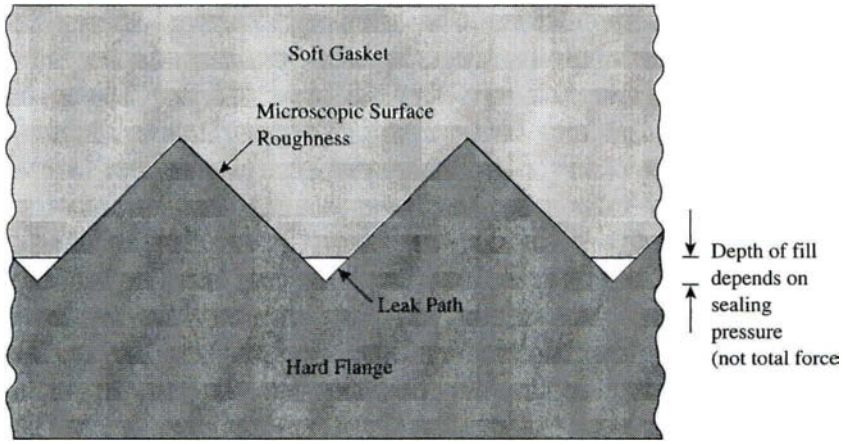
$$P = \frac{R}{3} \ln \left[\frac{0.34\, \Delta p\, Lr^2}{Qw} \right] \tag{1}$$

where P = mechanical pressure between the gasket and sealing surface
R = gasket metal sealing factor: 3Kgf/mm^2 for soft aluminum, 5Kgf/mm^2 for soft copper
Δp = gas pressure across the seal (typically atmospheric)
L = length of the seal
r = peak to peak roughness of the hard seal surface
Q = desired leak rate
w = width of the seal

Using this formula, a 10^{-13} torr·L/sec leak rate can be achieved in a 100-mm-diameter by 1-mm-wide seal with 1μm (0.001 mm) surface roughness using a soft aluminum gasket under 27.4 Kgf/mm^2 sealing pressure. This serves as a convenient benchmark since a factor of 10 variation in any of these parameters changes the required sealing pressure by only 2.3 Kgf/mm^2. This pressure is much higher than the ultimate strength of most soft aluminum alloys but is in the range of their Brinell hardness (such as 26 Kgf/mm^2 for 1060-H14) [20]. Therefore, work-hardening caused by substantial plastic deformation of an aluminum gasket will provide the necessary sealing pressure (see Figure 4).

Fig. 3.

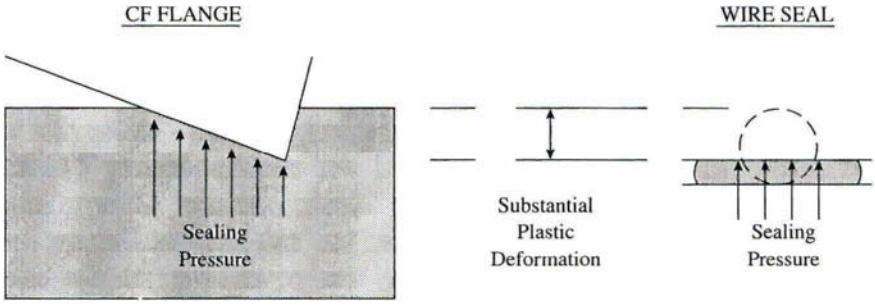
Sealing mechanism of demountable seals.



Leak rate is net conductance of microscopic grooves left unfilled by flow of gasket material.

Fig. 4.

Sealing pressure depends on hardness of work-hardened gasket material.



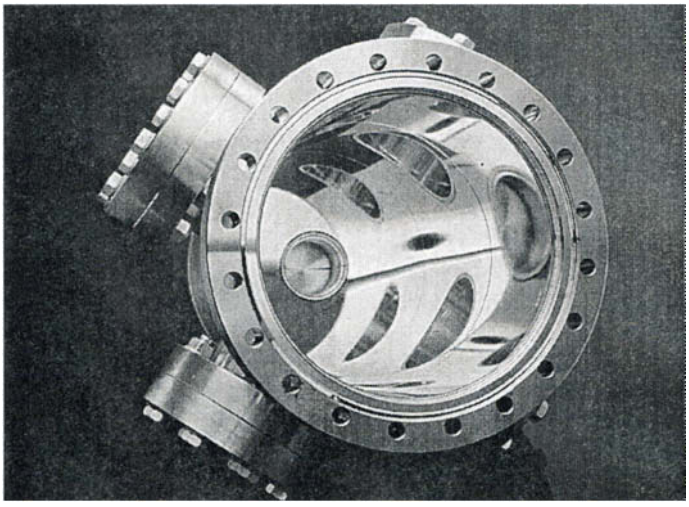
This is relatively uniform since plastic deformation is substantial.

4.8.2.2 CF Flanges

Commercially available aluminum CF flanges (see Figure 5) produce this plastic deformation by pressing the knife-edge deep into a soft aluminum gasket. Plastic deformation begins to occur after only 3 μm of penetration and produces enough work-hardening to achieve a UHV seal long before the flange faces actually touch. Deeper penetration merely widens the area of the seal reacting against the gasket hardness but does not significantly increase the effective sealing pressure. Confinement of the gasket within its flange well stores approximately 5 μm (6 μm for

Fig. 5.

Aluminum CF flanges and chamber.



(Courtesy of SMC Corporation.)

copper) of elastic strain that can compensate for slight changes in seal geometry due to thermal gradients or mechanically induced strains.

4.8.2.3 Wire Seals

Wire seals, reported in the literature, achieve approximately the same sealing pressure by substantially flattening the wire [21,22]. These require less bolting force than CF seals because sealing pressure is concentrated over a much narrower seal width.

4.8.2.4 Metal C-rings and Diamond-Shaped Seals

Commercially available metal C-rings have relatively wide seal widths. Although the average pressure over the contact area is approximately one-third that required above, there is probably a strain concentration near the center that does produce this. The use of a collapsible knife-edge on the C-ring, made of soft gasket material, has also been used to produce high sealing pressure by concentrating sealing force [23]. When practical considerations limit the amount of sealing pressure supportable by a C-ring, the sealing surface can be either polished to reduce the "r" term in Equation (1) or made rougher to create ridges that produce areas of

local high sealing pressure. Experimental data suggests that reliable sealing occurs for surface roughness below 1μm and above 10 μm but not within this range [24].

4.8.2.5 Materials Considerations

For long life, the yield strength of the sealing surface must be substantially higher than the hardness of the gasket material. Commercially available aluminum CF flanges have a yield strength of 40 Kgf/mm^2, which is well above the 26 Kgf/mm^2 gasket hardness. However, the 6061 alloy widely used as a chamber material has a yield strength of only 28 Kgf/mm^2, which may be insufficient for long life [25]. An inert hard coating (such as ion-plated TiC) on the flange face can be useful in preventing cold welding to the gasket. Without this, provisions may be necessary for prying flanges apart when opening. Coatings also prevent the formation of porous oxides that can cause diffusion leak paths across the seal.

4.8.2.6 Potential for Seal Disturbance During Bakeout

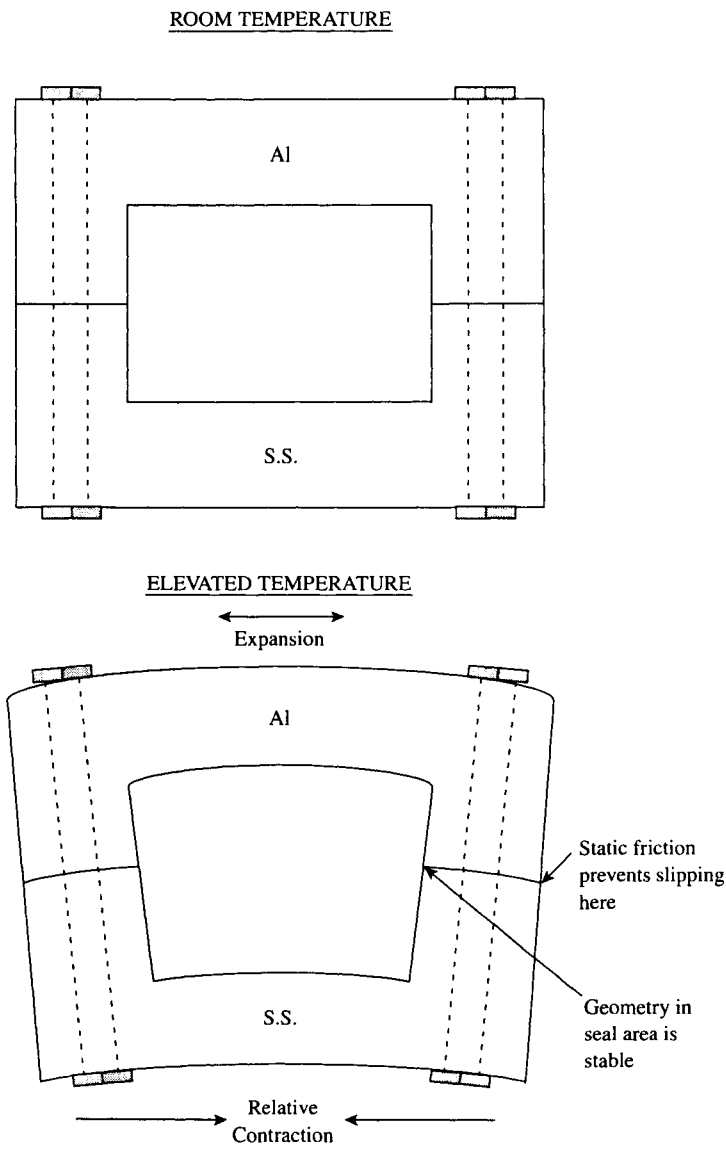
Because of differences in thermal expansion between room temperature and 150°C, unconstrained aluminum and stainless steel flanges develop a 0.1% dimensional difference. This represents a 1 μm radial seal displacement per millimeter of radius. Also, because of changes in bolt length and flange thickness, clamping force falls by 15% when using aluminum bolts or rises by over 30% when using stainless steel bolts (for initial tension at half tensile strength). These effects are further amplified by temperature gradients that can easily occur if heating or cooling is too rapid because of the large difference in thermal conductivity of the two materials. Aluminum-to-stainless steel transition pieces, made by explosion bonding or special brazing, can be used to avoid these problems but are expensive.

4.8.2.7 Constrained Flanges

To minimize the effects of thermal expansion on seal integrity, demountable dissimilar metal flanges are normally constrained to prevent motion near the seal. This is accomplished by tightly clamping them face to face so that static friction between the bolt circles prevents slippage. Differential expansion produces radial tensile stress in the stainless steel flange and compressive stress in the aluminum flange, which result in small elastic strains but leave the seal geometry near the constrained bolt circle effectively undisturbed (Figure 6). High bolt torque, on the order of 300 in·lb for $\frac{5}{16}$-inch bolts, is used with CF flanges. Seal integrity has been demonstrated through 150° bakeouts for standard CF sizes up to 12 inches

Fig. 6.

Constrained flanges.



Elastic strains induced by differential thermal expansion are small especially in the seal area near the bolts

[26]. Similar constraints are used in other demountable flange systems such as those used in CERN's LEP particle accelerator [27] and the metal C-ring seal proposed by KEK [28].

4.8.2.8 Unconstrained Flanges

It is sometimes desirable to leave flange faces unconstrained. For example, CF flanges can be assembled with a small gap between the faces to allow for additional tightening in case of future leaks. In this case, differential expansion causes the gasket to roll so that the inner edge tilts slightly toward the aluminum flange. Experimental data have shown that reliable CF seals can be maintained for up to about 2.1 degrees of roll [29]. Thus, standard CF sizes of up to 6 inches can be baked as high as 150°C without measurable leaks. Since the aluminum gasket material softens at elevated temperatures, bolts should be made of aluminum so that thermal expansion reduces clamping force by a like amount, preventing the knife-edges from biting deeper into the gasket. Wire seals in unconstrained flanges appear to allow seal surface displacement up to at least half the original thickness of the wire [30]. Metal C-rings also allow substantial displacement.

4.8.3 CLEANING AND SURFACE FINISHING

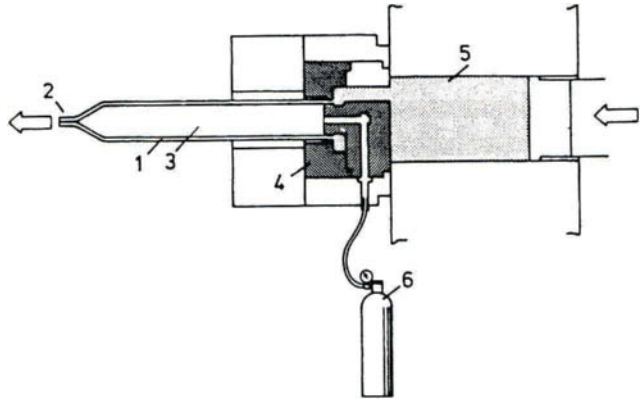
Outgassing from interior chamber surfaces can be minimized by creating a thin, dense, stable passivation layer. This both reduces the amount of gas that can be entrapped on the surface and provides a barrier against gas diffusion from the bulk. The Al_2O_3 that naturally forms on aluminum can be made to have excellent vacuum characteristics and has been widely used. This oxide should be as pure as possible, because hydroxides and impurities, such as magnesium oxide, tend to disrupt the lattice structure, leading to lower density and higher porosity. Suitable finishes are produced by completely removing the scale as received from the mill and growing a new layer under controlled conditions. Scales can be removed either by mechanical cutting (e.g., extrusion or machining) or by chemical etching. High-performance scales can be grown on all standard alloys commonly used in vacuum chambers. However, pure aluminum (1000 series) tends to produce slightly better results, probably because it does not contain alloying elements that disrupt the oxide structure.

4.8.3.1 Mechanical Finishes

Mechanical finishing methods have been designated EX and EL processes [31]. The former involves extruding or machining the surface, without cutting fluid, in an oxygen and argon atmosphere at standard pressure that immediately reoxidizes the newly exposed bare metal (Figures 7 and 8 respectively). Argon makes up

Fig. 7.

Schematic diagram of the EX special extrusion process.

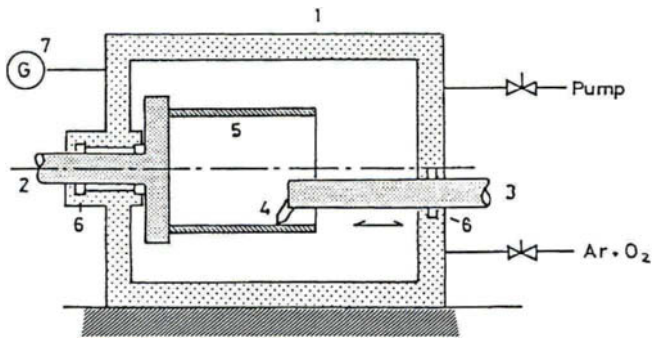


1) Al Pipe, 2) Vacuum tight end, 3) Argon and oxygen mixture, 4) Port-hole die, and 5) Gas Tank. (Courtesy of H. Ishimaru, "Developments and Applications for All-Aluminum Alloys Vacuum Systems," *MRS Bulletin*, Vol. XV, No. 7, July 1990, p. 24.)

about 90% of the atmosphere to dilute the oxygen and prevent explosive "burning" of machining chips and particles. Moisture content must be no more than a few ppm to minimize the formation of hydroxides. The EL process involves machining in a normal air atmosphere but using alcohol (normally ethanol) as a cutting fluid [32]. This results in a similar thin, dense oxide scale; however, the actual chemical mechanism that produces it is unclear. It has been reported that scratches can be repaired by manually scraping with Scotch-Brite® soaked in ethanol.

Fig. 8.

Schematic diagram of EX machining process.



(1) vacuum chamber, 2) axis of the lathe, (3) axis of the tool, (4) tool, (5) work, (6) vacuum seal, and (7) vacuum gauge. (Courtesy of H. Ishimaru, "Developments and Applications for All-Aluminum Alloys Vacuum Systems," *MRS Bulletin*, Vol. XV, No. 7, July 1990, p. 24.)

4.8.3.2 Chemical Finishing — A Well-Documented Recipe

Chemical finishing methods are much easier than mechanical cleaning and offer vacuum performance that is often comparable. Although many successful procedures have been demonstrated, the cleaning procedure developed for the Advanced Photon Source at Argonne National Laboratory is among the simplest and most carefully studied and documented [33,34]. It is designed for 6063 and 2219 alloys and may be summarized as follows [35]:

1. Use water-based cutting fluids and avoid oils during initial manufacturing.
2. Preclean with a high pressure spray of 2% Almeco 18® (Henkel/Parker Amchem P3 Detergent).
3. Clean in an ultrasonic bath of 2% Almeco 18 at 65°C for 10 minutes. When alloy 2219 is present, temperature must be reduced to 50°C to prevent excessive copper buildup on the surface. Copper formed inadvertently can be removed with 2% Citrinox®. The latter is also used instead of Almeco 18 when copper components are present in the structure.
4. Rinse with deionized water at room temperature for 10 minutes. (Elevated temperatures were found to cause the oxide scale to substantially thicken.)
5. Blow-dry with hot, dry nitrogen.

The scale that reforms in the presence of air and moisture is a heterogeneous mix of hydroxides and oxides (γ-$Al(OH)_3$, γ-AlOOH, and γ-Al_2O_3). When heated to 150°C, these transform to γ-Al_2O_3 [36] creating a relatively stable and homogeneous passivation layer. For this reason, at least one bake-out should immediately follow any chemical finishing procedure.

Surfaces heavily contaminated with oil should be precleaned with organic solvents. Other cleaning agents such as Alconox® have been successfully used [37,38]. Ultrasonic agitation should be intense (approximately 8 mW/cm^3) but not enough to cause pitting due to cavitation. Oxides of alloying elements, such as magnesium, that the detergent does not attack, sometimes form a smut that can be removed with an agent such as Wyandotte Alutone® [40]. (This has not been a problem with 6063 cleaned with Almeco 18.) Additional cleaning can be done with strong acids or bases such as Amklene® (principally KOH) [41] or Amchem etchant #33 [42], however the benefits of this appear to be marginal [43]. Several rinsing cycles are often used. Initial rinsing with tap water followed by deionized water has been demonstrated to be effective [44].

4.8.4 MECHANICAL CONSIDERATIONS

Aluminum is commercially available in a large variety of alloys with widely varying characteristics. Most alloys suitable for vacuum systems have a yield strength

comparable to stainless steel 304 and generally require wall thicknesses only slightly greater. An important cost advantage of aluminum over stainless steel is the relative ease with which complex chamber geometries can be fabricated by machining from a single block or by extrusion. Alloy selection usually depends on fabrication issues such as weldability, machineability, extrudability, availability, and cost. Sometimes corrosion resistance and chemical compatibility with the process must also be considered.

4.8.4.1 Alloy Types and Designations

Generally, only wrought aluminum alloys are suitable for high-vacuum chambers, because cast alloys are too porous. Wrought alloys are designated by a four-digit number followed by a hyphenated suffix specifying temper. The first digit specifies the dominant alloying element as follows: 1, >99.00% aluminum; 2, copper; 3, manganese; 4, silicon; 5, magnesium; 6, magnesium and silicon; and 7, zinc. The other three digits follow no particular pattern.

Data on manufacturability, strength, temperature affects, and chemical compositions of selected alloys are given in Tables 2, 3, 4, and 5 respectively [45]. All alloy series have been successfully used in high-vacuum systems except the 7000 series, which is unsuitable because of the high vapor pressure of its principal alloying element, zinc. Alloys in the 2000 series may require special shielding in applications sensitive to copper. Usually, the main body of a chamber is made from 5000 or 6000 series alloys such as 5083, 6061, or 6063. Components requiring high strength can be made from 2000 series, especially 2219, which is the most weldable of this series.

The main challenge to machining aluminum is its tendency to form long chips that tear rather than shear from the surface. Softer alloys tend to build up on the cutting tool unless it is free of scratches and carefully lubricated to inhibit bonding [46]. Diamond tools can be used for especially fine finishes. In some cases, machining can be done without cutting fluid to minimize residual contamination. However, the use of water-based cutting fluids has been shown to provide comparable vacuum performance [47]. Mineral oil should be avoided. The machineability ratings in Table 2 are in increasing order of chip length and finish quality.

4.8.4.2 Chamber Wall Thickness Requirements

Chambers can be considered as being composed of flat plates and cylinders for purposes of determining the wall thickness required to withstand the differential pressure created by an internal vacuum.

Table 2

Comparative Characteristics and Applications of Selected Aluminum Alloys

ALLOY AND TEMPER	RESISTANCE TO CORROSION		Workability (Cold)①	Machinability③	Brazeability④	WELDABILITY⑤			SOME APPLICATIONS OF ALLOYS
	General①	Stress-Corrosion Cracking②				Gas	Arc	Resistance Spot and Seam	
1060-0	A	A	A	E	A	A	A	B	Chemical equipment, railroad tank cars
H12	A	A	A	E	A	A	A	A	
H14	A	A	A	D	A	A	A	A	
H16	A	A	B	D	A	A	A	A	
H18	A	A	B	D	A	A	A	A	
1100-0	A	A	A	E	A	A	A	B	Sheet metal work, spun hollowware, fin stock
H12	A	A	A	E	A	A	A	A	
H14	A	A	A	D	A	A	A	A	
H16	A	A	B	D	A	A	A	A	
H18	A	A	C	D	A	A	A	A	
1350-0	A	A	A	E	A	A	A	B	Electrical conductors
H12, H111	A	A	A	E	A	A	A	A	
H14, H24	A	A	A	D	A	A	A	A	
H16, H26	A	A	B	D	A	A	A	A	
H18	A	A	B	D	A	A	A	A	
2011-T3	D⑥	D	C	A	D	D	D	D	Screw machine products
T4, T451	D⑥	D	B	A	D	D	D	D	
T8	D	B	D	A	D	D	D	D	
2014-0	..	..	..	D	D	D	D	B	Truck frames, aircraft structures
T3, T4, T451	D⑥	C	C	B	D	D	B	B	
T6, T651, T6510, T6511	D	C	D	B	D	D	B	B	
2017-T4, T451	D⑥	C	C	B	D	D	B	B	Screw machine products, fittings
2018-T61	..	..	..	B	D	D	C	B	Aircraft engine cylinders, heads and pistons
2024-0	..	..	..	D	D	D	D	D	Truck wheels, screw machine products, aircraft structures
T4, T3, T351, T3510, T3511	D⑥	C	C	B	D	C	B	B	
T361	D⑥	C	D	B	D	D	C	B	
T6	D	B	C	B	D	D	C	B	
T861, T81, T851, T8510, T8511	D	B	D	B	D	D	C	B	
T72	..	..	..	B	D	D	C	B	
2025-T6	D	C	..	B	D	D	B	B	Forgings, aircraft propellers
2036-T4	C	..	B	C	D	C	B	B	Auto body panel sheet
2117-T4	C	A	B	C	D	D	B	B	Rivets
2124-T851	D	B	D	B	D	D	C	B	Aircraft structures
2218-T61	D	C	..	..	D	D	C	B	Jet engine impellers and rings
T72	D	C	..	B	D	D	C	B	
2219-0	..	..	..	..	D	D	A	B	Structural uses at high temperatures (to 600°F) High strength weldments
T31, T351, T3510, T3511	D⑥	C	C	B	D	A	A	A	
T37	D⑥	C	D	B	D	A	A	A	
T81, T851, T8510, T8511	D	B	D	B	D	A	A	A	
T87	D	B	D	B	D	A	A	A	
2618-T61	D	C	..	B	D	D	C	B	Aircraft engines
3003-0	A	A	A	E	A	A	A	B	Cooking untensils, chemical equipment, pressure vessels, sheet metal work, builder's hardware, storage tanks
H12	A	A	A	E	A	A	A	A	
H14	A	A	B	D	A	A	A	A	
H16	A	A	C	D	A	A	A	A	
H18	A	A	C	D	A	A	A	A	
H25	A	A	B	D	A	A	A	A	
3004-0	A	A	A	D	B	A	A	B	Sheet metal work, storage tanks
H32	A	A	B	D	B	A	A	A	
H34	A	A	B	C	B	A	A	A	
H36	A	A	C	C	B	A	A	A	
H38	A	A	C	C	B	A	A	A	
3105-0	A	A	A	E	A	A	A	B	Residential siding, mobile homes, rain carrying goods, sheet metal work
H12	A	A	B	E	A	A	A	A	
H14	A	A	B	D	A	A	A	A	
H16	A	A	C	D	A	A	A	A	
H18	A	A	C	D	A	A	A	A	
H25	A	A	B	D	A	A	A	A	
4032-T6	C	B	..	B	D	D	B	C	Pistons
5005-0	A	A	A	E	B	A	A	B	Appliances, utensils, architectural, electrical conductor
H12	A	A	A	E	B	A	A	A	
H14	A	A	B	D	B	A	A	A	
H16	A	A	C	D	B	A	A	A	
H18	A	A	C	D	B	A	A	A	

(continues)

Table 2—Continued

ALLOY AND TEMPER	RESISTANCE TO CORROSION		Workability (Cold)[5]	Machinability[3]	Brazeability[6]	WELDABILITY[6]			SOME APPLICATIONS OF ALLOYS
	General[1]	Stress-Corrosion Cracking[2]				Gas	Arc	Resistance Spot and Seam	
5005-0									
H32	A	A	A	E	B	A	A	A	
H34	A	A	B	D	B	A	A	A	
H36	A	A	C	D	B	A	A	A	
H38	A	A	C	D	B	A	A	A	
5050-0	A	A	A	E	B	A	A	B	Builder's hardware, refrigerator trim, coiled tubes
H32	A	A	A	D	B	A	A	A	
H34	A	A	B	D	B	A	A	A	
H36	A	A	C	C	B	A	A	A	
H38	A	A	C	C	B	A	A	A	
5052-0	A	A	A	D	C	A	A	B	Sheet metal work, hydraulic tube, appliances
H32	A	A	B	D	C	A	A	A	
H34	A	A	B	C	C	A	A	A	
H36	A	A	C	C	C	A	A	A	
H38	A	A	C	C	C	A	A	A	
5056-0	A[4]	B[4]	A	D	D	C	A	B	Cable sheathing, rivets for magnesium, screen wire, zipper
H111	A[4]	B[4]	A	D	D	C	A	A	
H12, H32	A[4]	B[4]	B	D	D	C	A	A	
H14, H34	A[4]	B[4]	B	C	D	C	A	A	
H18, H38	A[4]	C[4]	C	C	D	C	A	A	
H192	B[4]	D[4]	D	B	D	C	A	A	
H392	B[4]	D[4]	D	B	D	C	A	A	
5083-0	A[4]	A[4]	B	D	D	C	A	B	Unfired, welded pressure vessles, marine, auto aircraft cryogenics, TV towers, drilling rigs, transportation equipment, missile components
H321, H116	A[4]	A[4]	C	D	D	C	A	A	
H111	A[4]	B[4]	C	D	D	C	A	A	
5086-0	A[4]	A[4]	A	D	D	C	A	B	
H32, H116	A[4]	A[4]	B	D	D	C	A	A	
H34	A[4]	B[4]	B	C	D	C	A	A	
H36	A[4]	B[4]	C	C	D	C	A	A	
H38	A[4]	B[4]	C	C	D	C	A	A	
H111	A[4]	A[4]	B	D	D	C	A	A	
5154-0	A[4]	A[4]	A	D	D	C	A	B	Welded structures, storage tanks, pressure vessels, salt water service
H32	A[4]	A[4]	B	D	D	C	A	A	
H34	A[4]	A[4]	B	C	D	C	A	A	
H36	A[4]	A[4]	C	C	D	C	A	A	
H38	A[4]	A[4]	C	C	D	C	A	A	
5252-H24	A	A	B	D	C	A	A	A	Automotive and appliance trim
H25	A	A	B	C	C	A	A	A	
H28	A	A	C	C	C	A	A	A	
5254-0	A[4]	A[4]	A	D	D	C	A	B	Hydrogen peroxide and chemical storage vessels
H32	A[4]	A[4]	B	D	D	C	A	A	
H34	A[4]	A[4]	B	C	D	C	A	A	
H36	A[4]	A[4]	C	C	D	C	A	A	
H38	A[4]	A[4]	C	C	D	C	A	A	
5454-0	A	A	A	D	D	C	A	B	Welded structures, pressure vessels, marine service
H32	A	A	B	D	D	C	A	A	
H34	A	A	B	C	D	C	A	A	
H111	A	A	B	D	D	C	A	A	
5456-0	A[4]	B[4]	B	D	D	C	A	B	High strength welded structures, pressure vessels, marine applications, storage tanks
H321, H116	A[4]	B[4]	C	D	D	C	A	A	
5457-0	A	A	A	E	B	A	A	B	
5652-0	A	A	A	D	C	A	A	B	Hydrogen peroxide and chemical storage vessels
H32	A	A	B	D	C	A	A	A	
H34	A	A	B	C	C	A	A	A	
H36	A	A	C	C	C	A	A	A	
H38	A	A	C	C	C	A	A	A	
5657-H241	A	A	A	D	B	A	A	A	Anodized auto and appliance trim
H25	A	A	B	D	B	A	A	A	
H26	A	A	B	D	B	A	A	A	
H28	A	A	C	D	B	A	A	A	
6005-T1, T5	..	..	..	..	A	A	A	A	
6053-0	..	..	..	E	B	A	A	B	Wire and rod for rivets
T6, T61	A	A	..	C	B	A	A	A	

(continues)

Table 2—Continued

ALLOY AND TEMPER	RESISTANCE TO CORROSION		Workability (Cold)③	Machinability③	Brazeability④	WELDABILITY⑤			SOME APPLICATIONS OF ALLOYS
	General①	Stress-Corrosion Cracking②				Gas	Arc	Resistance Spot and Seam	
6061-0	B	A	A	D	A	A	A	B	Heavy-duty structures requiring good corrosion resistance, truck and marine, railroad cars, furniture, pipelines
T4, T451, T4510, T4511	B	B	B	C	A	A	A	A	
T6, T651, T652, T6510, T6511	B	A	C	C	A	A	A	A	
6063-T1	A	A	B	D	A	A	A	A	Pipe railing, furniture, architectural extrusions
T4	A	A	B	D	A	A	A	A	
T5, T52	A	A	B	C	A	A	A	A	
T6	A	A	C	C	A	A	A	A	
T83, T831, T832	A	A	C	C	A	A	A	A	
6066-0	C	A	B	D	D	D	B	B	Forgings and extrusion for welded structures
T4, T4510, T4511	C	B	C	C	D	D	B	B	
T6, T6510, T6511	C	B	C	B	D	D	B	B	
6070-T4, T4511	B	B	B	C	D	A	A	A	Heavy duty welded structures, pipelines
T6	B	B	C	C	D	A	A	A	
6101-T6, T63	A	A	C	C	A	A	A	A	High strength bus conductors
T61, T64	A	A	B	D	A	A	A	A	
6151-T6, T652	..	..	..	..	B	..	..	..	Moderate strenght, intricate forgings for machine and auto parts
6201-T81	A	A	..	C	A	A	A	A	High strength electric conductor wire
6262-T6, T651, T6510, T6511	B	A	C	B	B	B	B	A	Screw machine products
T9	B	A	D	B	B	B	B	A	
6351-T1	..	..	C	C	C	B	A	B	Extruded shapes, structurals, pipe and tube
-T4	A	..	C	C	C	B	A	B	
-T5	A	..	C	C	C	B	A	A	
-T6	A	..	C	C	C	B	A	A	
6463-T1	A	A	B	D	A	A	A	A	Extruded architectural and trim sections
-T5	A	A	B	C	A	A	A	A	
-T6	A	A	C	C	A	A	A	A	
6951-T42, T62	..	..	..	..	A	A	A	A	
7005-T53	..	..	..	..	B	C	A	A	
7049-T73, T7352	C	B	D	B	D	D	D	B	Aircraft forgings
7050-T73510, T73511 T74⑦, T74510⑦, T74511⑦, T74520⑦, T7651, T76510, T76511	C	B	D	B	D	D	D	B	Aircraft and other structures
7075-0	..	..	..	D	D	D	D	B	Aircraft and other structures
T6, T651, T652, T6510, T6511,	C③	C	D	B	D	D	D	B	
T73, T7351	C	B	D	B	D	D	D	B	
7175-T74, T7452, T7454	C	B	D	B	D	D	C	B	
7178-0	..	..	..	..	D	D	D	B	Aircraft and other structures
T6, T651, T6510, T6511	C③	C	D	B	D	D	D	B	
7475-O	..	..	..	..	D	D	D	B	Shell Casings
7475-T61, -T651	C	C	D	B	D	D	B	B	Aircraft & Other
7475-T761, T7351	C	B	D	B	D	D	D	B	Structures
8017-H12, H22, H221	A	A	A	D	A	A	A	A	Electrical conductors
8030-H12, H221	A	A	A	E	A	A	A	A	Electrical conductors
8176-H14, H24	A	A	A	D	A	A	A	A	Electrical conductors
8177-H13, H23, H221	A	A	A	E	A	A	A	A	Electrical conductors

Source: Aluminum Association, *Aluminum Standards and Data*, Tenth Edition, 900 19th St., N.W., Washington, D.C., 1990, p. 51, Table 3.3.

Table 3

Typical Mechanical Properties of Aluminum Alloys

ALLOY AND TEMPER	TENSION				HARDNESS	SHEAR	FATIGUE	MODULUS
	STRENGTH ksi		ELONGATION percent in 2 in.		BRINELL NUMBER	ULTIMATE SHEARING STRENGTH	ENDURANCE① LIMIT	MODULUS① OF ELASTICITY
	ULTIMATE	YIELD	1/16 in. Thick Specimen	1/2 in. Diameter Specimen	500 kg load 10 mm ball	ksi	ksi	ksi x 10^3
1060-0	10	4	43	..	19	7	3	10.0
1060-H12	12	11	16	..	23	8	4	10.0
1060-H14	14	13	12	..	26	9	5	10.0
1060-H16	16	15	8	..	30	10	6.5	10.0
1060-H18	19	18	6	..	35	11	6.5	10.0
1100-0	13	5	35	45	23	9	5	10.0
1100-H12	16	15	12	25	28	10	6	10.0
1100-H14	18	17	9	20	32	11	7	10.0
1100-H16	21	20	6	17	38	12	9	10.0
1100-H18	24	22	5	15	44	13	9	10.0
2014-0	27	14	..	18	45	18	13	10.6
2014-T4, T451	62	42	..	20	105	38	20	10.6
2014-T6, T651	70	60	..	13	135	42	18	10.6
2024-0	27	11	20	22	47	18	13	10.6
2024-T3	70	50	18	..	120	41	20	10.6
2024-T4, T351	68	47	20	19	120	41	20	10.6
2024-T361②	72	57	13	..	130	42	18	10.6
2219-0	25	11	18	..	..	..	..	10.6
2219-T42	52	27	20	..	..	..	..	10.6
2219-T31, T351	52	36	17	..	..	..	..	10.6
2219-T37	57	46	11	..	..	..	..	10.6
2219-T62	60	42	10	..	..	..	15	10.6
2219-T81, T851	66	51	10	..	..	..	15	10.6
2219-T87	69	57	10	..	..	..	15	10.6
3003-0	16	6	30	40	28	11	7	10.0
3003-H12	19	18	10	20	35	12	8	10.0
3003-H14	22	21	8	16	40	14	9	10.0
3003-H16	26	25	5	14	47	15	10	10.0
3003-H18	29	27	4	10	55	16	10	10.0
5083-0	42	21	..	22	..	25	..	10.3
5083-H321, H116	46	33	..	16	..	..	23	10.3
5086-0	38	17	22	..	..	23	..	10.3
5086-H32, H116	42	30	12	..	..	..	..	10.3
5086-H34	47	37	10	..	..	27	..	10.3
5086-H112	39	19	14	..	..	..	..	10.3
5154-0	35	17	27	..	58	22	17	10.2
5154-H32	39	30	15	..	67	22	18	10.2
5154-H34	42	33	13	..	73	24	19	10.2
5154-H36	45	36	12	..	78	26	20	10.2
5154-H38	48	39	10	..	80	28	21	10.2
5154-H112	35	17	25	..	63	..	17	10.2
5456-0	45	23	..	24	..	..	..	10.3
5456-H112	45	24	..	22	..	..	..	10.3
5456-H321, H116	51	37	..	16	90	30	..	10.3
6061-0	18	8	25	30	30	12	9	10.0
6061-T4, T451	35	21	22	25	65	24	14	10.0
6061-T6, T651	45	40	12	17	95	30	14	10.0
6063-0	13	7	..	..	25	10	8	10.0
6063-T1	22	13	20	..	42	14	9	10.0
6063-T4	25	13	22	..	..	..	..	10.0
6063-T5	27	21	12	..	60	17	10	10.0
6063-T6	35	31	12	..	73	22	10	10.0
6063-T83	37	35	9	..	82	22	..	10.0
6063-T831	30	27	10	..	70	18	..	10.0
6063-T832	42	39	12	..	95	27	..	10.0

Source: Aluminum Association, *Aluminum Standards and Data*, Tenth Edition, 900 19th St., N.W., Washington, D.C., 1990, p. 30, Table 2.1.

Note: The above values are for comparison purposes only. Actual characteristics vary. To convert ksi values in the table to Kgf/mm^2, multiply by 0.7045.

Table 4

Typical Tensile Properties at Various Temperatures[1]

ALLOY AND TEMPER	TEMP.	TENSILE STRENGTH, ksi		ELONGATION IN 2 IN., PERCENT
	°F	ULTIMATE	YIELD②	
1100-0	−320	25	6	50
	−112	15	5.5	43
	− 18	14	5	40
	75	13	5	40
	212	10	4.6	45
	300	8	4.2	55
	400	6	3.5	65
	500	4	2.6	75
	600	2.9	?	80
	700	2.1	.6	85
1100-H14	−320	30	20	45
	−112	20	18	24
	− 18	19	17	20
	75	18	17	20
	212	16	15	20
	300	14	12	23
	400	10	7.5	26
	500	4	2.6	75
	600	2.9	2	80
	700	2.1	1.6	85
1100-H18	−320	34	26	30
	−112	26	23	16
	− 18	25	23	15
	75	24	22	15
	212	21	19	15
	300	18	14	20
	400	6	3.5	65
	500	4	2.6	75
	600	2.9	2	80
	700	2.1	1.6	85
2011-T3	75	55	43	15
	212	47	34	18
	300	28	19	25
	400	16	11	35
	500	6.5	3.8	45
	600	3.1	1.8	90
	700	2.3	1.4	125
2014-T6, T651	-320	84	72	14
	-112	74	65	13
	18	72	62	13
	75	70	60	13
	212	63	57	15
	300	40	35	20
	400	16	13	38
	500	9.5	7.5	52
	600	6.5	5	65
	700	4.3	3.5	72
2017-T4, T451	−320	80	53	28
	−112	65	42	24
	− 18	64	41	23
	75	62	40	22
	212	57	39	18
	300	40	30	15
	400	16	13	35
	500	9	7.5	45
	600	6	5	65
	700	4.3	3.5	70
2024-T3 (sheet)	−320	85	62	18
	−112	73	52	17
	− 18	72	51	17
	75	70	50	17
	212	66	48	16
	300	55	45	11
	400	27	20	23
	500	11	9	55
	600	7.5	6	75
	700	5	4	100
2024-T4-T351 (plate)	−320	84	61	19
	−112	71	49	19
	18	69	47	19
	75	68	47	19
	212	63	45	19
	300	45	36	17
	400	26	19	27
	500	11	9	55
	600	7.5	6	75
	700	5	4	100
2024-T6, T651	−320	84	68	11
	−112	72	59	10
	− 18	70	58	10
	75	69	57	10
	212	65	54	10
	300	45	36	17
	400	26	19	27
	500	11	9	55
	600	7.5	6	75
	700	5	4	100
2024-T81, T851	−320	85	78	8
	−112	74	69	7
	− 18	73	68	7
	75	70	65	7
	212	66	62	8
	300	55	49	11
	400	27	20	23
	500	11	9	55
	600	7.5	6	75
	700	5	4	100
2024-T861	−320	92	85	5
	−112	81	77	5
	− 18	78	74	5
	75	75	71	5
	212	70	67	6
	300	54	48	11
	400	21	17	28
	500	11	9	55
	600	7.5	6	75
	700	5	4	100

(*continues*)

Table 4—Continued

ALLOY AND TEMPER	TEMP. °F	TENSILE STRENGTH, ksi ULTIMATE	TENSILE STRENGTH, ksi YIELD②	ELONGATION IN 2 IN., PERCENT
2117-T4	-320	56	33	30
	-112	45	25	29
	- 18	44	24	28
	75	43	24	27
	212	36	21	16
	300	30	17	20
	400	16	12	35
	500	7.5	5.5	55
	600	4.7	3.3	80
	700	2.9	2	110
2124-T851	-452	102	90	10
	-320	86	79	9
	-112	76	71	8
	- 18	73	68	8
	75	70	64	9
	212	66	61	9
	300	54	49	13
	400	27	20	28
	500	11	8	60
	600	7.5	6	75
	700	5.5	4.1	100
2218-T61	-320	72	52	15
	-112	61	45	14
	- 18	59	44	13
	75	59	44	13
	212	56	42	15
	300	41	35	17
	400	22	16	30
	500	10	6	70
	600	5.5	3	85
	700	4	2.5	100
2219-T62	-320	73	49	16
	-112	63	44	13
	- 18	60	42	12
	75	58	40	12
	212	54	37	14
	300	45	33	17
	400	34	25	20
	500	27	20	21
	600	10	8	40
	700	4.4	3.7	75
2219-T81,T8551	-320	83	61	15
	-112	71	54	13
	- 18	69	52	12
	75	66	50	12
	212	60	47	15
	300	49	40	17
	400	36	29	20
	500	29	23	21
	600	7	6	55
	700	4.4	3.7	75
2618-T61	-320	78	61	12
	-112	67	55	11
	- 18	64	54	10
	75	64	54	10
	212	62	54	10
	300	50	44	14
	400	32	26	24
	500	13	9	50
	600	7.5	4.5	80
	700	5	3.5	120
3003-0	-320	33	8.5	46
	-112	20	7	42
	-18	17	6.5	41
	75	16	6	40
	212	13	5.5	43
	300	11	5	47
	400	8.5	4.3	60
	500	6	3.4	65
	600	4	2.4	70
	700	2.8	1.8	70
3003-H14	-320	35	25	30
	-112	24	22	18
	- 18	22	21	16
	75	22	21	16
	212	21	19	16
	300	18	16	16
	400	14	9	20
	500	7.5	4	60
	600	4	2.4	70
	700	2.8	1.8	70
3003-H18	-320	41	33	23
	-112	32	29	11
	- 18	30	28	10
	75	29	27	10
	212	26	21	10
	300	23	16	11
	400	14	9	18
	500	7.5	4	60
	600	4	2.4	70
	700	2.8	1.8	70
3004-0	-320	42	13	38
	-112	28	11	30
	- 18	26	10	26
	75	26	10	25
	212	26	10	25
	300	22	10	35
	400	14	9.5	55
	500	10	7.5	70
	600	7.5	5	80
	700	5	3	90
3004-H34	-320	52	34	26
	-112	38	30	16
	- 18	36	29	13
	75	35	29	12
	212	34	28	13
	300	28	25	22
	400	21	15	35

(*continues*)

Table 4—Continued

ALLOY AND TEMPER	TEMP. °F	TENSILE STRENGTH, ksi ULTIMATE	TENSILE STRENGTH, ksi YIELD②	ELONGATION IN 2 IN., PERCENT
	500	14	7.5	55
	600	7.5	5	80
	700	5	3	90
3004-H38	−320	58	43	20
	−112	44	38	10
	− 18	42	36	7
	75	41	36	6
	212	40	36	7
	300	31	27	15
	400	22	15	30
	500	12	7.5	50
	600	7.5	5	80
	700	5	3	90
4032-T6	−320	66	48	11
	−112	58	46	10
	− 18	56	46	9
	75	55	46	9
	212	50	44	9
	300	37	33	9
	400	13	9	30
	500	8	5.5	50
	600	5	3.2	70
	700	3.4	2	90
5050-0	−320	37	10	..
	−112	22	8.5	..
	− 18	21	8	..
	75	21	8	..
	212	21	8	..
	300	19	8	..
	400	14	7.5	..
	500	9	6	..
	600	6	4.2	..
	700	3.9	2.6	..
5050-H34	−320	44	30	..
	-112	30	25	..
	- 18	28	24	..
	75	28	24	..
	212	28	24	..
	300	25	22	..
	400	14	7.5	..
	500	9	6	..
	600	6	4.2	..
	700	3.9	2.6	..
5050-H38	−320	46	36	..
	−112	34	30	..
	− 18	32	29	..
	75	32	29	..
	212	31	29	..
	300	27	25	..
	400	14	7.5	..
	500	9	6	..
	600	6	4.2	..
	700	3.9	2.6	..
5052-0	−320	44	16	46
	−112	29	13	35
	− 18	28	13	32
	75	28	13	30
	212	28	13	36
	300	23	13	50
	400	17	11	60
	500	12	7.5	80
	600	7.5	5.5	110
	700	5	3.1	130
5052-H34	−320	55	36	28
	−112	40	32	21
	− 18	38	31	18
	75	38	31	18
	212	38	31	18
	300	30	27	27
	400	24	15	45
	500	12	7.5	80
	600	7.5	5.5	110
	700	5	3.1	130
5052-H38	−320	60	44	25
	−112	44	38	18
	− 18	42	37	15
	75	42	37	14
	212	40	36	16
	300	34	28	24
	400	25	15	45
	500	12	7.5	80
	600	7.5	5.5	110
	700	5	3.1	130
5083-0	−320	59	24	36
	−112	43	21	30
	− 18	42	21	27
	75	42	21	25
	212	40	21	36
	300	31	19	50
	400	22	17	60
	500	17	11	80
	600	11	7.5	110
	700	6	4.2	130
5086-0	−320	55	19	46
	−112	39	17	35
	− 18	38	17	32
	75	38	17	30
	212	38	17	36
	300	29	16	50
	400	22	15	60
	500	17	11	80
	600	11	7.5	110
	700	6	4.2	130
5154-0	−320	52	19	46
	−112	36	17	35
	− 18	35	17	32

(*continues*)

Table 4—Continued

ALLOY AND TEMPER	TEMP.	TENSILE STRENGTH, ksi		ELONGATION IN 2 IN., PERCENT
	°F	ULTIMATE	YIELD②	
	75	35	17	30
	212	35	17	36
	300	29	16	50
	400	22	15	60
	500	17	11	80
	600	11	7.5	110
	700	6	4.2	130
5254-0	−320	52	19	46
	−112	36	17	35
	− 18	35	17	32
	75	35	17	30
	212	35	17	36
	300	29	16	50
	400	22	15	60
	500	17	11	80
	600	11	7.5	110
	700	6	4.2	130
5454-0	−320	54	19	39
	−112	37	17	30
	− 18	36	17	27
	75	36	17	25
	212	36	17	31
	300	29	16	50
	400	22	15	60
	500	17	11	80
	600	11	7.5	10
	700	6	4.2	30
5454-H32	−320	59	36	32
	−112	42	31	23
	− 18	41	30	20
	75	40	30	18
	212	39	29	20
	300	32	26	37
	400	25	19	45
	500	17	11	80
	600	11	7.5	110
	700	6	4.2	130
5454-H34	-320	63	41	30
	-112	46	36	21
	- 18	44	35	18
	75	44	35	16
	212	43	34	18
	300	34	28	32
	400	26	19	45
	500	17	11	80
	600	11	7.5	110
	700	6	4.2	130
5456-0	−320	62	26	32
	−112	46	23	25
	− 18	45	23	22
	75	45	23	20
	212	42	22	31
	300	31	20	50
	400	22	17	60
	500	17	11	80
	600	11	7.5	110
	700	6	4.2	130
5652-0	−320	44	16	46
	−112	29	13	35
	− 18	28	13	32
	75	28	13	30
	212	28	13	30
	300	23	13	50
	400	17	11	60
	500	12	7.5	80
	600	7.5	5.5	110
	700	5	3.1	130
5652-H34	320	55	36	28
	-112	40	32	21
	- 18	38	31	18
	75	38	31	16
	212	38	31	18
	300	30	27	27
	400	24	15	45
	500	12	7.5	80
	600	7.5	5.5	110
	700	5	3.1	130
5652-H38	−320	60	44	25
	−112	44	38	18
	− 18	42	37	15
	75	42	37	14
	212	40	36	16
	300	34	28	24
	400	25	15	45
	500	12	7.5	80
	600	7.5	5.5	110
	700	5	3.1	130
6053-T6, T651	75	37	32	13
	212	32	28	13
	300	25	24	13
	400	13	12	25
	500	5.5	4	70
	600	4	2.7	80
	700	2.9	2	90
6061-T6, T651	−320	60	47	22
	−112	49	42	18
	− 18	47	41	17
	75	45	40	17
	212	42	38	18
	300	34	31	20
	400	19	15	28
	500	7.5	5	60
	600	4.6	2.7	85
	700	3	1.8	95
6063-T1	−320	34	16	44
	−112	26	15	36
	− 18	24	14	34
	75	22	13	33
	212	22	14	18
	300	21	15	20
	400	9	6.5	40
	500	4.5	3.5	75
	600	3.2	2.5	80
	700	2.3	2	105

(*continues*)

Table 4—Continued

ALLOY AND TEMPER	TEMP. °F	TENSILE STRENGTH, ksi ULTIMATE	TENSILE STRENGTH, ksi YIELD②	ELONGATION IN 2 IN., PERCENT
6063-T5	-320	37	24	28
	-112	29	22	24
	- 18	28	22	23
	75	27	21	22
	212	24	20	18
	300	20	18	20
	400	9	6.5	40
	500	4.5	3.5	75
	600	3.2	2.5	80
	700	2.3	2	105
6063-T6	-320	47	36	24
	-112	38	33	20
	- 18	36	32	19
	75	35	31	18
	212	31	28	15
	300	21	20	20
	400	9	6.5	40
	500	4.5	3.5	75
	600	3.3	2.5	80
	700	2.3	2	105
6101-T6	-320	43	33	24
	-112	36	30	20
	- 18	34	29	19
	75	32	28	19
	212	28	25	20
	300	21	19	20
	400	10	7	40
	500	4.8	3.3	80
	600	3	2.3	100
	700	2.5	1.8	105
6151-T6	-320	57	50	20
	-112	50	46	17
	- 18	49	45	17
	75	48	43	17
	212	43	40	17
	300	28	27	20
	400	14	12	30
	500	6.5	5	50
	600	5	3.9	43
	700	4	3.2	35
6262-T651	-320	60	47	22
	-112	49	42	18
	- 18	47	41	17
	75	45	40	17
	212	42	38	18
	300	34	31	20
6262-T9	-320	74	67	14
	-112	62	58	10
	- 18	60	56	10
	75	58	55	10
	212	53	52	10
	300	38	37	14
	400	15	13	34
	500	8.5	6	48
	600	4.6	2.7	85
	700	3	1.8	95
7075-T6, T651	-320	102	92	9
	-112	90	79	11
	- 18	86	75	11
	75	83	73	11
	212	70	65	14
	300	31	27	30
	400	16	13	55
	500	11	9	65
	600	8	6.5	70
	700	6	4.6	70
7075-T73, T7351	-320	92	72	14
	-112	79	67	14
	- 18	76	65	13
	75	73	63	13
	212	63	58	15
	300	31	27	30
	400	16	13	55
	500	11	9	65
	600	8	6.5	70
	700	6	4.6	70
7175-T74	-320	106	98	13
	-112	90	83	14
	- 18	87	80	16
	212	72	69	17
	300	35	31	30
	400	18	13	65
7178-T6, T651	-320	106	94	5
	-112	94	84	8
	- 18	91	81	9
	75	88	78	11
	212	73	68	14
	300	31	27	40
	400	15	12	70
	500	11	9	76
	600	8.5	7	80
	700	6.5	5.5	80
7178-T76, T7651	-320	106	89	10
	-112	91	78	10
	- 18	88	76	10
	75	83	73	11
	212	69	64	17
	300	31	27	40
	400	15	12	70
	500	11	9	76
	600	8.5	7	80
	700	6.5	5.5	80
7475-T61 Sheet	-320	99	87	10
	-112	88	79	12
	- 18	84	75	12
	75	80	72	12
	212	70	65	14
	300	30	26	28
	400	14	11	55
	500	9.5	7	70
	600	6.5	5.5	80
	700	5	3.8	85
7475-T761 Sheet	-320	95	82	11
	-112	84	73	12
	- 18	80	70	12
	75	76	67	12
	212	64	61	14
	300	30	26	38
	400	14	11	55
	500	9.5	7	70
	600	6.5	5.5	80
	700	5	3.8	85

Source: Aluminum Association, *Aluminum Standards and Data*, Tenth Edition, 900 19th St., N.W., Washington, D.C., 1990, p. 34, Table 2.2.

Note: The above values are for comparison purposes only. Actual characteristics vary. To convert ksi values in the table to Kgf/mm^2, multiply by 0.7045.

Table 5

Chemical Composition Limits of Wrought Aluminum Alloys

AA DESIGNATION	SILICON	IRON	COPPER	MANGANESE	MAGNESIUM	CHROMIUM	NICKEL	ZINC	TITANIUM	OTHERS[2]		ALUMINUM Min.[4]
										Each[3]	Total[3]	
1050	0.25	0.40	0.05	0.05	0.05	..	..	0.05	0.03	0.03[5]	..	99.50
1060	0.25	0.35	0.05	0.03	0.03	..	..	0.05	0.03	0.03[5]	..	99.60
1100	0.95 Si + Fe		0.05–0.20	0.05	..	..	..	0.10	..	0.05[6]	0.15	99.00
1145[8]	0.55 Si + Fe		0.05	0.05	0.05	..	..	0.05	0.03	0.03[5]	..	99.45
1175[7]	0.15 Si + Fe		0.10	0.02	0.02	..	..	0.04	0.02	0.02[5]	..	99.75
1200	1.00 Si + Fe		0.05	0.05	..	..	..	0.10	0.05	0.05	0.15	99.00
1230[7]	0.70 Si + Fe		0.10	0.05	0.05	..	..	0.10	0.03	0.03[5]	..	99.30
1235	0.65 Si + Fe		0.05	0.05	0.05	..	..	0.10	0.06	0.03[5]	..	99.35
1345	0.03	0.40	0.10	0.05	0.05	..	..	0.05	0.03	0.03[5]	..	99.45
1350[9]	0.10	0.40	0.05	0.01	..	0.01	..	0.05	..	0.03[5]	0.10	99.50
2011	0.40	0.7	5.0–6.0	..	..	..	..	0.30	..	0.05[10]	0.15	Remainder
2014	0.05–1.2	0.7	3.9–5.0	0.40–1.2	0.20–0.8	0.10	..	0.25	0.15	0.05	0.15	Remainder
2017	0.20–0.8	0.7	3.5–4.5	0.40–1.0	0.40–0.8	0.10	..	0.25	0.15	0.05	0.15	Remainder
2018	0.9	1.0	3.5–4.5	0.20	0.45–0.9	0.10	1.7–2.3	0.25	..	0.05	0.15	Remainder
2024	0.50	0.50	3.8–4.9	0.30–0.9	1.2–1.8	0.10	..	0.25	0.15	0.05	0.15	Remainder
2025	0.50–1.2	1.0	3.9–5.0	0.40–1.2	0.05	0.10	..	0.25	0.15	0.05	0.15	Remainder
2036	0.50	0.50	2.2–3.0	0.10–0.40	0.30–0.6	0.10	..	0.25	0.15	0.05	0.15	Remainder
2117	0.8	0.7	2.2–3.0	0.20	0.20–0.50	0.10	..	0.25	..	0.05	0.15	Remainder
2124	0.20	0.30	3.8–4.9	0.30–0.9	1.2–1.8	0.10	..	0.25	0.15	0.05	0.15	Remainder
2218	0.9	1.0	3.5–4.5	0.20	1.2–1.8	0.10	1.7–2.3	0.25	..	0.05	0.15	Remainder
2219	0.20	0.30	5.8–6.8	0.20–0.40	0.02	..	..	0.10	0.02–0.10	0.05[11]	0.15	Remainder
2319	0.20	0.30	5.8–6.8	0.20–0.40	0.02	..	..	0.10	0.10–0.20	0.05[11]	0.15	Remainder
2618	0.10–0.25	0.9–1.3	1.9–2.7	..	1.3–1.8	..	0.9–1.2	0.10	0.04–0.10	0.05	0.15	Remainder
3003	0.6	0.7	0.05–0.20	1.0–1.5	..	..	..	0.10	..	0.05	0.15	Remainder
3004	0.30	0.7	0.25	1.0–1.5	0.8–1.3	..	..	0.25	..	0.05	0.15	Remainder
3005	0.6	0.7	0.30	1.0–1.5	0.20–0.6	0.10	..	0.25	0.10	0.05	0.15	Remainder
3105	0.6	0.7	0.30	0.30–0.8	0.20–0.8	0.20	..	0.40	0.10	0.05	0.15	Remainder
4032	11.0–13.5	1.0	0.50–1.3	..	0.8–1.3	0.10	0.50–1.3	0.25	..	0.05	0.15	Remainder
4043	4.5–6.0	0.8	0.30	0.05	0.05	..	..	0.10	0.20	0.05[12]	0.15	Remainder
4045[13]	9.0–11.0	0.8	0.30	0.05	0.05	..	..	0.10	0.20	0.05	0.15	Remainder
4047[13]	11.0–13.0	0.8	0.30	0.15	0.10	..	..	0.20	..	0.05[12]	0.15	Remainder
4145[13]	9.3–10.7	0.8	3.3–4.7	0.15	0.15	0.15	..	0.20	..	0.05[12]	0.15	Remainder
4343[13]	6.8–8.2	0.8	0.25	0.10	..	..	..	0.20	..	0.05	0.15	Remainder
5005	0.30	0.7	0.20	0.20	0.50–1.1	0.10	..	0.25	..	0.05	0.15	Remainder
5050	0.40	0.7	0.20	0.10	1.1–1.8	0.10	..	0.25	..	0.05	0.15	Remainder
5052	0.25	0.40	0.10	0.10	2.2–2.8	0.15–0.35	..	0.10	..	0.05	0.15	Remainder
5056	0.30	0.40	0.10	0.05–0.20	4.5–5.6	0.05–0.20	..	0.10	..	0.05	0.15	Remainder
5083	0.40	0.40	0.10	0.40–1.0	4.0–4.9	0.05–0.25	..	0.25	0.15	0.05	0.15	Remainder
5086	0.40	0.50	0.10	0.20–0.7	3.5–4.5	0.05–0.25	..	0.25	0.15	0.05	0.15	Remainder
5154	0.25	0.40	0.10	0.10	3.1–3.9	0.15–0.35	..	0.20	0.20	0.05	0.15	Remainder
5183	0.40	0.40	0.10	0.50–1.0	4.3–5.2	0.05–0.25	..	0.25	0.15	0.05[14]	0.15	Remainder
5252	0.08	0.10	0.10	0.10	2.2–2.8	..	..	0.05	..	0.03[9]	0.10	Remainder
5254	0.45 Si + Fe		0.05	0.01	3.1–3.9	0.15–0.35	..	0.20	0.05	0.05	0.15	Remainder
5356	0.25	0.40	0.10	0.05–0.20	4.5–5.5	0.05–0.20	..	0.10	0.06–0.20	0.05[15]	0.15	Remainder
5454	0.25	0.40	0.10	0.50–1.0	2.4–3.0	0.05–0.20	..	0.25	0.20	0.05	0.15	Remainder
5456	0.25	0.40	0.10	0.50–1.0	4.7–5.5	0.05–0.20	..	0.25	0.20	0.05	0.15	Remainder
5457	0.08	0.10	0.20	0.15–0.45	0.8–1.2	..	..	0.05	..	0.03[9]	0.10	Remainder
5554	0.25	0.40	0.10	0.50–1.0	2.4–3.0	0.05–0.20	..	0.25	0.05–0.20	0.05[14]	0.15	Remainder
5556	0.25	0.40	0.10	0.50–1.0	4.7–5.5	0.05–0.20	..	0.25	0.05–0.20	0.05[15]	0.15	Remainder
5652	0.40 Si + Fe		0.04	0.01	2.2–2.8	0.15–0.35	..	0.10	..	0.05	0.15	Remainder
5654	0.45 Si + Fe		0.05	0.01	3.1–3.9	0.15–0.35	..	0.20	0.05–0.15	0.05[14]	0.15	Remainder
5657	0.08	0.10	0.10	0.03	0.6–1.0	..	..	0.05	..	0.02[11]	0.05	Remainder

For all numbered footnotes, see page 98.

(*continues*)

Table 5—Continued

AA DESIGNATION	SILICON	IRON	COPPER	MANGANESE	MAGNESIUM	CHROMIUM	NICKEL	ZINC	TITANIUM	OTHERS[22] Each[23]	OTHERS[22] Total[3]	ALUMINUM Min.[4]
6003[7]	0.35-1.0	0.6	0.10	0.8	0.8-1.5	0.35	..	0.20	0.10	0.05	0.15	Remainder
6005	0.6-0.9	0.35	0.10	0.10	0.40-0.6	0.10	..	0.10	0.10	0.05	0.15	Remainder
6053	[15]	0.35	0.10	..	1.1-1.4	0.15-0.35	..	0.10	..	0.05	0.15	Remainder
6061	0.40-0.8	0.7	0.15-0.40	0.15	0.8-1.2	0.04-0.35	..	0.25	0.15	0.05	0.15	Remainder
6063	0.20-0.6	0.35	0.10	0.10	0.45-0.9	0.10	..	0.10	0.10	0.05	0.15	Remainder
6066	0.9-1.8	0.50	0.7-1.2	0.6-1.1	0.8-1.4	0.40	..	0.25	0.20	0.05	0.15	Remainder
6070	1.0-1.7	0.50	0.15-0.40	0.40-1.0	0.50-1.2	0.10	..	0.25	0.15	0.05	0.15	Remainder
6101[12]	0.30-0.7	0.50	0.10	0.03	0.35-0.8	0.03	..	0.10	..	0.03[17]	0.10	Remainder
6105	0.6-1.0	0.35	0.10	0.10	0.45-0.8	0.10	..	0.10	0.10	0.05	0.15	Remainder
6151	0.6-1.2	1.0	0.35	0.20	0.45-0.8	0.15-0.35	..	0.25	0.15	0.05	0.15	Remainder
6162	0.40-0.8	0.50	0.20	0.10	0.7-1.1	0.10	..	0.25	0.10	0.05	0.15	Remainder
6201	0.50-0.9	0.50	0.10	0.03	0.6-0.9	0.03	..	0.10	..	0.03[17]	0.10	Remainder
6253[7]	[15]	0.50	0.10	..	1.0-1.5	0.04-0.35	..	1.6-2.4	..	0.05	0.15	Remainder
6262	0.40-0.8	0.7	0.15-0.40	0.15	0.8-1.2	0.04-0.14	..	0.25	0.15	0.05[5]	0.15	Remainder
6351	0.7-1.3	0.50	0.10	0.40-0.8	0.40-0.8	..	..	0.20	0.20	0.05	0.15	Remainder
6463	0.20-0.6	0.15	0.20	0.05	0.45-0.9	..	..	0.05	..	0.05	0.15	Remainder
6951	0.20-0.50	0.8	0.15-0.40	0.10	0.40-0.8	..	..	0.20	..	0.05	0.15	Remainder
7005	0.35	0.40	0.10	0.20-0.7	1.0-1.8	0.06-0.20	..	4.0-5.0	0.01-0.06	0.05[18]	0.15	Remainder
7008[7]	0.10	0.10	0.05	0.05	0.7-1.4	0.12-0.25	..	4.5-5.5	0.05	0.05	0.10	Remainder
7049	0.25	0.35	1.2-1.9	0.20	2.0-2.9	0.10-0.22	..	7.2-8.2	0.10	0.05	0.15	Remainder
7050	0.12	0.15	2.0-2.6	0.10	1.9-2.6	0.04	..	5.7-6.7	0.06	0.05[21]	0.15	Remainder
7072[7]	0.7 Si + Fe		0.10	0.10	0.10	..	..	0.8-1.3	..	0.05	0.15	Remainder
7075	0.40	0.50	1.2-2.0	0.30	2.1-2.9	0.18-0.28	..	5.1-6.1	0.20	0.05	0.15	Remainder
7175	0.15	0.20	1.2-2.0	0.10	2.1-2.9	0.18-0.28	..	5.1-6.1	0.10	0.05	0.15	Remainder
7178	0.40	0.50	1.6-2.4	0.30	2.4-3.1	0.18-0.28	..	6.3-7.3	0.20	0.05	0.15	Remainder
7475	0.10	0.12	1.2-1.9	0.06	1.9-2.6	0.18-0.25	..	5.2-6.2	0.06	0.05	0.15	Remainder
8017	0.10	0.55-0.8	0.10-0.20	..	0.01-0.05	..	..	0.05	..	0.03[13]	0.10	Remainder
8030	0.10	0.30-0.8	0.15-0.30	..	0.05	..	..	0.05	..	0.03[24]	0.10	Remainder
8176	0.03-0.15	0.40-1.0	..	..	..	..	..	0.10	..	0.05[25]	0.15	Remainder
8177	0.10	0.25-0.45	0.04	..	0.04-0.12	..	..	0.05	..	0.03[26]	0.10	Remainder

Note: Listed herein are designations and chemical composition limits for some wrought unalloyed aluminum and for wrought aluminum alloys registered with The Aluminum Association. This list does not include all alloys registered with The Aluminum Association. A complete list of registered designations is contained in the "Registration Record of International Alloy Designations and Chemical Composition Limits for Wrought Aluminum and Wrought Aluminum Alloys." These lists are maintained by the Technical Committee on Product Standards of the Aluminum Association.

① Composition in percent by weight maximum unless shown as a range or a minimum.

② Except for "aluminum" and "others," analysis normally is made for elements for which specific limits are shown. For purposes of determining conformance to these limits, an observed value or a calculated value obtained from analysis is rounded off to the nearest unit in the last right-hand place of figures used in expressing the specified limit, in accordance with ASTM Recommended Practice E 29.

③ The sum of those "other" metallic elements 0.010 percent or more each, expressed to the second decimal before determining the sum.

④ The aluminum content for unalloyed aluminum not made by a refining process is the difference between 100.00 percent and the sum of all other metallic elements present in amounts of 0.010 percent or more each, expressed to the second decimal before determining the sum.

⑤ Also contains 0.40-0.7 percent each of lead and bismuth.

⑥ Electric conductor. Formerly designated EC.

⑦ Cladding alloy. See Table 6.1.

⑧ Foil.

⑨ Vanadium 0.05 percent maximum.

⑩ Also contains 0.20-0.6 percent each of lead and bismuth.

⑪ Brazing alloy.

⑫ Bus conductor.

⑬ Vanadium plus titanium 0.02 percent maximum; boron 0.05 percent maximum; gallium 0.03 percent maximum.

⑭ Zirconium 0.08-0.20

⑮ Silicon 45 to 65 percent of actual magnesium content.

⑯ Beryllium 0.0008 maximum for welding electrode and welding rod only.

⑰ Boron 0.06 percent maximum.

⑱ Vanadium 0.05-0.15; zirconium 0.10-0.25.

⑲ Gallium 0.03 percent maximum; vanadium 0.05 percent maximum.

⑳ In addition to those alloys referencing footnote ⑯, a 0.0008 weight percent maximum beryllium is applicable to any alloy to be used as welding electrode or welding rod.

㉑ Zirconium 0.08-0.15.

㉒ Includes listed elements for which no specific limit is shown.

㉓ Boron 0.04 percent maximum; lithium 0.003 percent maximum.

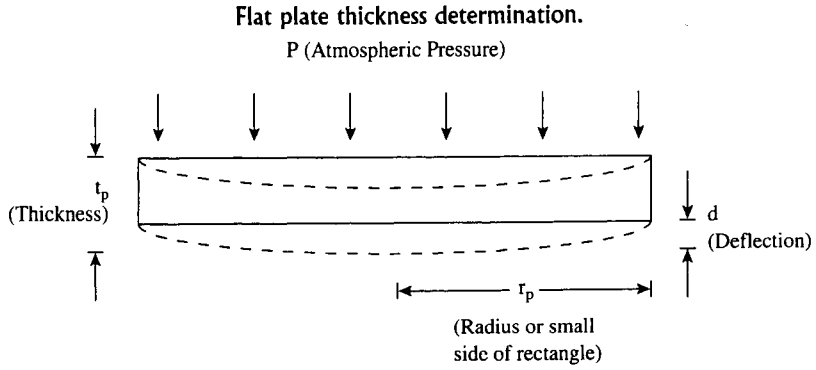
㉔ Boron 0.001-0.04.

㉕ Gallium 0.03 percent maximum.

㉖ Boron 0.04 percent maximum.

Source: Aluminum Association, *Aluminum Standards and Data*, Tenth Edition, 900 19th St., N.W., Washington, D.C., 1990, p. 97, Table 6.2.

Fig. 9.



Must limit deflection to an acceptable level.

FLAT PLATES

The thickness of flat plates is normally determined by the maximum deflection permissible (see Figure 9) and can be approximated as [48]:

$$t_p = \left[\frac{kPr_p^4}{Ed}\right]^{1/3} \tag{2}$$

where t_p = thickness of plate
k = 0.696 if unclamped, = 0.171 if clamped
P = atmospheric pressure
r_p = radius of circular plate or small side of rectangular plate
E = modulus of elasticity—Al, 7,470 Kgf/mm^2;
SS, 19,700 Kgf/mm^2
d = acceptable deflection

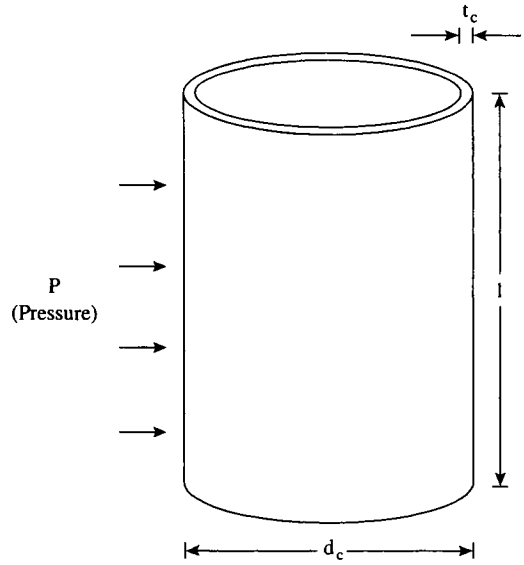
CYLINDERS

Cylinder walls can usually be much thinner than flat plates. The principal failure mode for cylinders pressurized on the outside is buckling due to instability of large unsupported surfaces. (Thickness of cylinders with pressure on the inside need only keep stresses below tensile strength—see peak stress formulas below.) The theoretical thickness necessary to prevent buckling in cylinders longer than a critical length, $l_c = 1.11\, d_c(d_c/t_c)^{0.5}$ (see Figure 10) is [49]

$$t_c = d_c\left[\frac{P(1-m^2)}{2E}\right]^{1/3} \tag{3}$$

Fig. 10.

Cylinder thickness determination.



Must be sufficient to prevent instability and buckling in large unsupported surfaces (when externally pressurized).

where t_c = thickness of cylinder wall
d_c = outside diameter of cylinder
P = pressure (typically 4 atm for safety)
m = Poisson's ratio (approximately 0.3 for both Al and SS)
E = stress/strain at actual operating point (elastic modulus)
l = length of cylinder or distance between supporting ribs

For cylinders with a length, l, shorter than l_c, thickness can be approximated using the following distillation of graphical and analytical relationships offered by Brownell (in any consistent unit system) [50]:

$$t_c = 0.697 \left[\frac{P}{E}\right]^{0.402} d_c^{0.585}\, l^{0.414} \tag{4}$$

RELATIVE THICKNESS OF ALUMINUM AND STAINLESS STEEL

Inspection of these formulas demonstrates that the minimum thickness (t_{AS}) of aluminum relative to stainless steel for the same application is related to their elastic moduli, E, as follows:

$$t_{AS}(\text{plate}) = \left[\frac{E_{SS}}{E_{Al}}\right]^{1/3} = \left[\frac{19{,}400}{7{,}470}\right]^{1/3} = 1.37 \tag{5}$$

$$t_{AS}(\text{long cylinder}) = \left[\frac{E_{SS}}{E_{Al}}\right]^{1/3} = 1.37 \tag{6}$$

$$t_{AS}(\text{short cylinder}) = \left[\frac{E_{SS}}{E_{Al}}\right]^{0.402} = 1.47 \tag{7}$$

Thus, aluminum chambers must be at most 50% thicker than stainless chambers. Often other considerations, such as structural stresses or manufacturing convenience, may require greater thickness than is necessary to merely withstand vacuum/atmospheric pressure differential. In such cases, aluminum and stainless steel may be of comparable thickness.

PEAK STRESSES AND STRAINS

Peak stresses should be calculated to ensure that material strength is not exceeded. These can be approximated as follows:

$$\text{Plate (clamped) stress} = 0.75\frac{Pr_p^{\,2}}{t_p^{\,2}} \tag{8}$$

$$\text{Plate (unclamped) stress [51]} = 1.24\frac{Pr_p^{\,2}}{t_p^{\,2}} \tag{9}$$

$$\text{Cylinder circumferential stress} = \frac{d_c P}{2t_c} \tag{10}$$

$$\text{Cylinder radial stress} = P \tag{11}$$

$$\text{Cylinder axial (force of end plates) stress [52]} = \frac{d_c P}{4t_c} \tag{12}$$

To ensure that deformation is within acceptable limits, strains should be estimated by dividing these stresses by elastic modulus, E.

WEIGHT AND THERMAL EXPANSION

Aluminum has about one-third the weight density of stainless steel: 2.70 versus 8.03 g/cm^3 respectively [53]. This makes aluminum appendages much lighter and easier to support than those made of stainless steel. However, allowances must be made for its 48% greater thermal expansion: 23.4 [54] versus 15.8 [55] microstrain per degrees Celsius respectively.

4.8.5 THERMAL CONDUCTIVITY AND EMISSIVITY

The thermal characteristics of aluminum are distinctly different from those of stainless steel and offer significant advantages where hot spots are a problem or where a chamber must be rapidly heated or cooled.

4.8.5.1 Thermal Gradients

Heat is often concentrated on a chamber wall by sources such as conduction from a hot appendage or by radiation from a nearby hot filament. The excessive outgassing and hazards caused by high temperatures at these points are greatly diminished by using aluminum, because it has a thermal conductivity 10 times higher than that of stainless steel. As an illustration, the temperature rise caused by a localized heat source on a large flat plate due only to heat conduction (Figure 11) is [56]

$$T_{\text{rise}} = q \frac{ln\left(\frac{d_2}{d_1}\right)}{2\pi kt} \tag{13}$$

where T_{rise} = maximum steady-state temperature rise
q = Incident heat
d_2 = diameter of outer perimeter at room temperature
d_1 = diameter of heat source (hot spot)
k = thermal conductivity: $k_{Al} = 206$, $k_{SS} = 16.2$ W/m·°C
t = thickness of surface

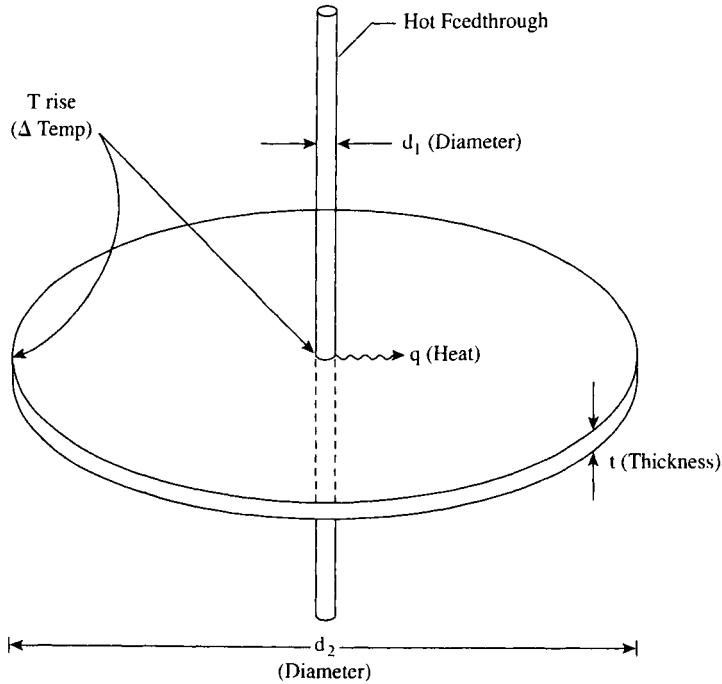
For purposes of comparison using the above formula, 10W of heat from a 6mm (0.25″) diameter feedthrough on a 1 m diameter by 6 mm thick plate will produce a 6.6°C rise in aluminum and an 83.8°C rise in stainless steel. The temperature rise is proportional to power and can become substantial for even moderate heat loads.

4.8.5.2 Absorption of Radiated Heat

Aluminum is highly reflective of incident radiated heat. It has an emissivity in the range of 0.04–0.07 compared to 0.3–0.7 for stainless steel [57]. In a closed system, all reflected heat is ultimately absorbed but in aluminum chambers tends to be uniformly distributed and not concentrated on the point of initial incidence.

Fig. 11.

Temperature rise caused by incident heat.



However, heat can concentrate on nonaluminum surfaces with high emissivity if these have substantial exposed cross sections.

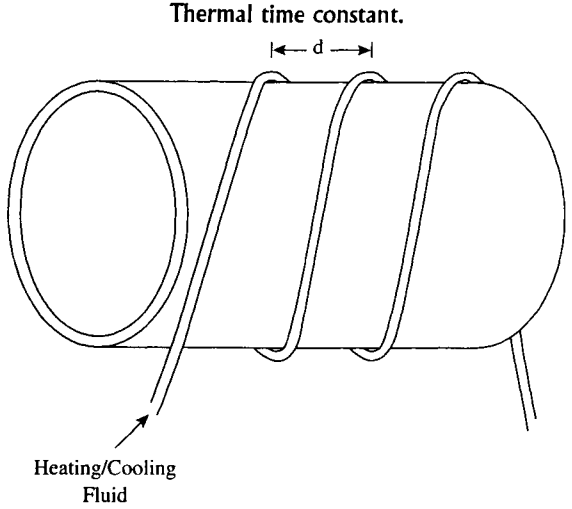
4.8.5.3 Thermal Time Constant

Because of its high thermal conductivity, an aluminum chamber can be uniformly heated and cooled in a fraction of the time required for a stainless steel chamber. The thermal time constant of a cylindrical system wrapped with tubing carrying heating/cooling fluid (Figure 12) can be estimated as follows:

$$TC = \frac{d^2 C_p D}{8k} \tag{14}$$

where TC = thermal time constant
d = distance between heating/cooling tubes
C_p = specific heat: $C_{p\,Al} = 870$, $C_{p\,SS} = 460$ W·sec/kg·°C
D = density: $D_{Al} = 2.70$, $D_{SS} = 8.03$ g/cm^3
k = thermal conductivity: $k_{Al} = 206$, $k_{SS} = 16.2$ W/m·°C

Fig. 12.



High thermal conductivity makes rapid heating and cooling of aluminum chambers feasible.

This assumes constant chamber thickness and approximates the distributed nature of its heat capacity and conductance by a lumped mass at the end of a thermal resistance corresponding to half the maximum distance from the heating fluid. It is intended to offer a feasibility perspective, not an exact prediction. For purposes of comparison, if heating/cooling tubes are spaced 20cm (8 inches) apart, the thermal time constants for aluminum and stainless steel are 1 minute and 19 minutes respectively.

This short thermal time constant of aluminum systems enables rapid partial baking of commercial systems, especially since, according to Figure 1, aluminum will start to effectively degas at temperatures corresponding to one time constant (102°C).

4.8.6

CORROSION

Both aluminum and stainless steel corrode in active gases and have limited life. Often the practical differences are minimal so that material selection is based on other considerations. For example, many semiconductor processes interact with chamber walls to produce particles or vapors. In such cases, aluminum is much less of a contamination problem than the heavy metals in stainless steel.

4.8.6.1 Suggested Approach to Predicting Corrosion Performance

Corrosion of metals occurs by many different chemical mechanisms, making the process so complex and difficult to characterize that accurate generalizations are nearly impossible. Nevertheless, a designer can use published corrosion data and a basic understanding of the mechanisms to make a reasonable guess. Performance must then be confirmed by testing under actual operating conditions. Published corrosion rates of aluminum and stainless steel for the most chemically active gases are given in Table 6 [58–79].

4.8.6.2 Principles for Extrapolating Published Data

The effects of other corrosive gas compounds is generally no worse than that of their halogen component alone. The principal mechanism of gas corrosion involves the scale (oxide layer) functioning as a semiconductor. Metal ions and electrons diffuse through it to the surface where they react with adsorbed corrosive gas. The scale grows at the outer surface, not at its interface with the base metal. This diffusion process appears to be the rate-limiting step in most cases. It is an activated process that increases with temperature but decreases with scale thickness [80].

The presence of moisture introduces electrochemical and acid/base mechanisms that usually increase the net corrosion rate. However, this is not always the case; for example, moisture greatly reduces chlorine corrosion of aluminum at high temperatures as shown in Table 6. In general, the transport barrier characteristics and high electrical resistivity of aluminum scales tend to inhibit the electrochemical effects of moisture. However, these scales are attacked if pH falls outside the range of about 4 to 8.5 [81]. It has been reported that corrosion begins when an HX/moisture mixture approaches its dewpoint. For example, HBr at 1 atm with 300 ppm moisture has a room-temperature dew point and causes corrosion; when moisture partial pressure falls below 100 ppm corrosion stops. Dew point temperature, however, is difficult to predict, because it appears to be positively correlated to moisture partial pressure and total pressure. Published data for making such estimates are very limited [82].

4.8.6.3 Heuristic Predictor of Corrosion Resistance

Scales that provide substantial corrosion protection usually have a volume per mole of metal in the range of one to two times that of the base metal; those out-

Table 6

Published Corrosion Rate Data Comparing Aluminum and Stainless Steel

Corrosive Gas	Aluminum Corrosion Rate	Stainless Steel Corrosion Rate
Atmosphere	For most alloys: <0.1 mm/yr decreasing to constant 0.003 mm/yr after 6 mo–2 yrs in the most severe sea coast locations [58] (2000 and 7000 series corrode at 2–3 times this rate [59])	Extremely low [60]
Chlorine	For 1100 alloy [61]: <0.5 mm/yr at 120°C, dry <0.5 mm/yr at 130°C, 0.06% moisture <0.5 mm/yr at 200°C, 1.5% moisture <0.5 mm/yr at 545°C, 30% moisture Al-Mg Alloys: lower rate [62] Al-Si Alloys: 2× higher rate [63]	For 304: 0.3 mm/yr at 100°C, dry 0.76 mm/yr at 290°C, dry 1.5 mm/yr at 315°C, dry 30.5 mm/yr at 40°C, 0.4% moisture [64] For 316: temperatures slightly higher [65]
Fluorine	For 1100 alloy: <0.05 mm/yr to 215°C, dry <0.5 mm/yr to 240°C dry [66] A durable protective scale forms [67] Moisture greatly accelerates corrosion [68] 2024 and 5224: 0.01 mm/yr at 547°C (200× higher for 1000 series) [69] Al-Si Alloys: lower rate [70]	For 304: <0.05 mm/yr at 210°C, dry <0.5 mm/yr at 249°C, dry [71] Moisture greatly accelerates corrosion [72] For 316: temperatures slightly higher [73]
Hydrogen chloride	Corrosive. However, aluminum has been used for handling *dry* HCl gas at 288°C [74]	<0.05 mm/yr to 293°C, dry (304, 316) [75]
Hydrogen fluoride	<0.5 mm/yr to 49°C, dry [76] <4.9 mm/yr to 500°C, dry <14.6 mm/yr to 600°C, dry [77]	<0.05 mm/yr to 49°C, dry <0.5 mm/yr to 204°C, dry [78] 13.4 mm/yr at 600°C, dry [79]

side this range usually do not. This can serve as a useful design tool even though the reasons for it are not well understood [83].

Thus, for corrosion protection,

$$1 < \frac{W_s G_m}{n W_m G_s} < 2 \tag{15}$$

where $W_{s,m}$ = molecular weight of the scale and metal respectively
$G_{s,m}$ = specific gravity of the scale and metal respectively
n = number of metal atoms in the scale molecule

This ratio for aluminum oxide (Al_2O_3) is

$$(101.94 \times 2.70)/(2 \times 26.97 \times 4.00) = 1.3$$

which is within the range where corrosion protection is expected, whereas the same ratio for iron oxide (Fe_2O_3) is 2.2 [84], which is outside this range. Iron oxides (rusts) are indeed less protective than aluminum oxides. However, it should be noted that these occur mainly on carbon steel; the nickel/chromium/iron scales on stainless steels are highly protective. In general, even though aluminum is very active chemically, it rapidly forms scales that provide effective passivation against further corrosion.

One notable exception to the preceding rule is the scale formed by chlorine ($AlCl_3$), which has a ratio of 1.8 but provides limited protection because its high vapor pressure and sublimation temperature of only 178°C causes it to evaporate away [85]. In contrast, AlF_3 formed by fluorine is known to be highly protective [86]. It melts at 1197°K and has a vapor pressure of only 0.075 torr at 1145°C [87].

4.8.6.4 Protective Coatings

Protective coatings of electroless nickel and hard-anodize (thick γ-Al_2O_3) provide very effective protection against corrosive gas at moderate temperatures, although the latter tends to increase outgassing by a factor of 40. In a plasma-activated chlorine environment, these are somewhat less effective (particularly the nickel). Proprietary, Al_2O_3 coatings with a high percentage of α-phase have been reported that virtually eliminate corrosion as a failure mode in turbomolecular pumps on reactive plasma etch systems [88,89]. A special process is required to create these, because α-Al_2O_3 normally forms above 800°C, which is above the melting point of aluminum.

4.8.7 WELDING ALUMINUM FOR VACUUM APPLICATIONS

Techniques for welding aluminum are well established and documented [90,91, 92]. These are somewhat different from those for stainless steel but not much more difficult. The main objective in fabricating vacuum chambers is to produce welds free of porosity and cracks that can cause leak paths or virtual leaks. Considerations for structural applications such as strength, color, and corrosion resistance are of secondary importance. The welding of aluminum need not be considered an art, because mechanisms affecting weld quality follow well-understood principles.

4.8.7.1 Avoiding Porosity and Particulate Defects

Hydrogen is the main cause of porosity in welds. Much more of it can dissolve in molten metal than can stay dissolved when the metal cools and solidifies. As hydrogen comes out of solution, it forms voids that interconnect to form leak paths. The molten/solid hydrogen solubility ratio for aluminum is 36 times higher than for iron. This makes aluminum welds much more sensitive to this source of porosity than those of stainless steel [93]. The principal sources of hydrogen are moisture, hydrocarbons, and hydroxides on the surface and in the surrounding atmosphere. Fortunately, relatively simple and practical measures can reduce these enough to produce good UHV welds.

Gloves should be worn at all times, because skin oils cause considerable contamination. Oils and finger prints can be removed from the work piece and filler rod with ethanol or acetone.

Oxides entrap moisture (a prime source of hydrogen) and have high melting temperatures that can cause incomplete fusion. They grow rapidly on bare aluminum and must be removed within four hours of welding. The surfaces within 1 cm of a joint should be scraped with a hard-edged tool [94]. Wire brushes may drive contamination into the soft surface and should be avoided. Since filler represents most of the metal in a joint, it is a prime source of potential contamination and worthy of special care in sourcing, packaging, storing, and handling—especially in wire form. Filler rod can be cleaned off with several passes of Scotch-Brite® immediately before welding.

Guard gas should be premium ultra-high-purity type (99.999%) and flushed through the tip for 20 seconds before initiating an arc to assure a moisture content below 10 ppm (preferably <3 ppm) [95]. Relative humidity in the welding area should be less than 75% and preferably less than 50%, which may necessitate working in an air-conditioned room. Fast weld speed has been shown to reduce porosity [96]. Efforts should be made to assure that there are no leaks in the inert gas lines, because these can be an inadvertent source of moisture build-up.

AC-TIG welding with argon gas should be used whenever possible, since during its positive half-cycles, shielding gas is ionized and accelerated toward the work piece, providing a sputter-etch cleaning action that removes residual contamination.

4.8.7.2 Avoiding Crack Defects

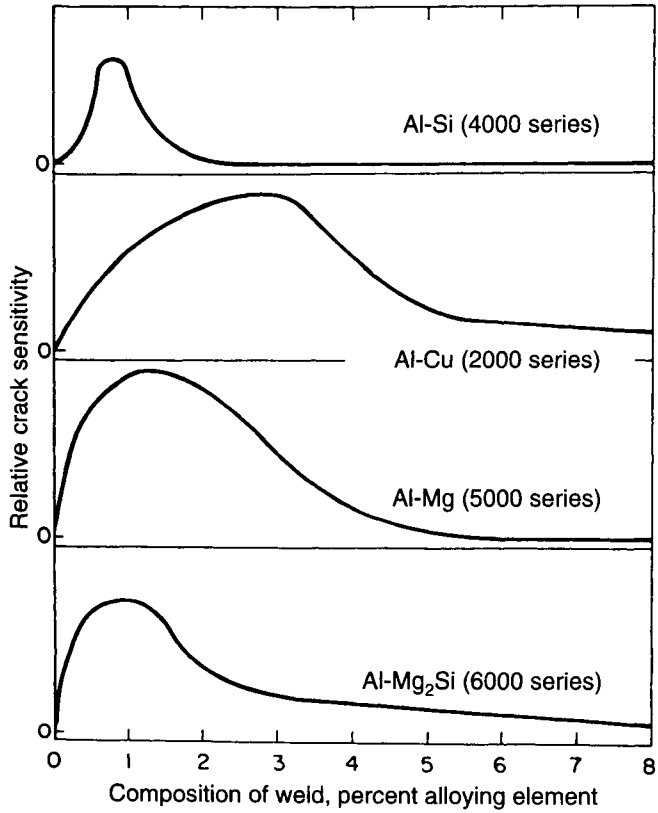
Cracks usually occur when there is a wide temperature range over which an alloy solidifies (freezing range). While the joint cools in this range, the bead is mushy and easily fractured by small thermal contractions. This process is partially mitigated by the fact that alloys contain some eutectics that continue to flow at the

low-temperature end of the freezing range and tend to heal cracks. These then solidify without cracks over a narrow temperature range [97].

Unfortunately, alloy compositions that produce good structural properties often also have wide freezing ranges. Weld crack sensitivity when joining these can be greatly reduced by selecting filler alloys that are less susceptible to cracking (i.e., narrower freezing range and/or high eutectic content), even when partially diluted by the base metal alloy. Most cracks occur near the center of a bead, which cools last. Fortunately, this area is least diluted by the base metal. Figure 13 shows the effect of various alloy compositions on crack sensitivity. It should be noted that for $AlMg_2Si$ (6000 series) alloys, crack resistance is increased by increasing either Mg or Si content. Fillers that solidify at much higher temperatures than the

Fig. 13.

Relationship of aluminum crack sensitivity to the presence of various alloying elements.



[Courtesy of J. H. Dudas and F. R. Collins, "Preventing Weld Cracks in High-Strength Aluminum Alloys," *Welding Journal* (June 1966, p. 4).]

Table 7

Guide to the Choice of Filler Metal for General Purpose Welding

Base Metal	319.0,333.0 354.0,355.0 C355.0,380.0	356.0,A356.0 A357.0,359.0 413.0,A444.0 443.0	511.0,512.0 513.0,514.0	7005[k],7039 710.0,711.0 712.0	6070	6061 6063,6101 6201,6151 6351,6951	5456	5454	5154 5254[a]	5086	5083	5052 5652[a]	5005 5050	3004 Alc.3004	2219 2519	2014 2036	1100 3003 Alc.3003	1060 1070,1080 1350
1060,1070, 1080,1350	4145[c,i]	4043[i,f]	4043[e,i]	4043[j]	4043[j]	4043[j]	5356[c]	4043[j]	4043[e,i]	5356[c]	5356[c]	4043[j]	1100[c]	4043	4145	4145	1100[c]	1188[j]
1100,3003 Alclad 3003	4145[c,i]	4043[i,f]	4043[e,i]	4043[j]	4043[j]	4043[j]	5356[c]	4043[e,i]	4043[e,i]	5356[c]	5356[c]	4043[e,i]	4043[e]	4043[e]	4145	4145	1100[c]	
2014,2036	4145[g]	4145			4145	4145									4145[g]	4145[g]		
2219,2519	4145[g,c,i]	4145[c,i]	4043[i]	4043[j]	4043[i,j]	4043[i,j]	4043	4043[j]	4043[i]	4043	4043	4043i	4043	4043	2319[c,f,i]			
3004 Alclad 3004	4043[j]	4043[i]	5654[b]	5356[e]	4043[e]	4043[b]	5356[e]	5654[b]	5654[b]	5356[e]	5356[e]	4043[e,i]	4043[e]	4043[e]				
5005,5050	4043[j]	4043[j]	5654[b]	5356e	4043[e]	4043[b]	5356[e]	5654[b]	5654[b]	5356[e]	5356[e]	4043[e,i]	4043[d,e]					
5052,5652[a]	4043[j]	4043[b,i]	5654[b]	5356[e]	5356[b,c]	5356[b,c]	5356[b]	5654[b]	5654b	5356[e]	5356[e]	5654[a,b,c]						
5083		5356[c,e,i]	5356[e]	5183[e]	5356[e]	5356[e]	5183[e]	5356[e]	5356[e]	5356[e]	5183[e]							
5086		5356[c,e,i]	5356[e]	5356[e]	5356[e]	5356[e]	5356[e]	5356[b]	5356[b]	5356[e]								
5154,5254[a]		4043[b,i]	5654[b]	5356[b]	5356[b,c]	5356[b,c]	5356[b]	5654[b]	5654[a,b]									
5454	4043[j]	4043[b,i]	5654[b]	5356[b]	5356[b,c]	5356[b,c]	5356[b]	5554[c,e]										
5456		4043[b,i]	5356[e]	5556[e]	5356[e]	5356[e]	5556[e]											
6061,6063, 6351,6101 6201,6151, 6951	4145[c,i]	5356[c,e,i]	5356[b c]	5356[b,c,i]	4043[b,i]	4043[b,i]												
6070	4145[c,i]	4043[e,i]	5356[c,e]	5356[c,e,i]	4043[e,i]													
7005[k],7039, 710.0,711.0 712.0	4043[j]	4043[b,i]	5356[b]	5356[e]														
511.0,512.0 513.0,514.0		4043[b,i]	5654[b,d]															
356.0,A356.0 A357.0,359.0 413.0 A444.0,443.0	4145[c,i]	4043[d,i]																
319.0,333.0 354.0,355.0, C355.0,380.0	4145[d,c,i]																	

NOTES:

1- Service conditions such as immersion in fresh or salt water, exposure to specific chemicals or a sustained high temperature (over 150 °F) may limit the choice of filler metals. Filler alloys 5356, 5183, 5556, and 5654 are not recommended for sustained elevated temperature service.

2- Recommendations in this table apply to gas shielded arc welding porcesses. For gas welding, only 1100, 1188 and 4043 filler metals are ordinarily used.

3- All filler metals are listed in AWS specification A5.10.

a Base metal alloys 5652 and 5254 are used for hydrogen peroxide service. 5654 filler metal is used for welding both alloys for low temperature service (150°F and below).

b 5183, 5356, 5554, 5556, and 5654 may be used. In some cases they provide: (1) improved color match after anodizing treatment; (2) highest weld ductility; and (3) higher weld strength. 5554 is suitable for elevated temperature service.

c 4043 may be used for some applications.

d Filler metal with the same analysis as the base metal is sometimes used.

e 5183, 5356, or 5556 may be used.

f 4145 may be used for some applications.

g 2319 may be used for some applications.

i 4047 may be used for some applications.

j 1100 may be used for some applications.

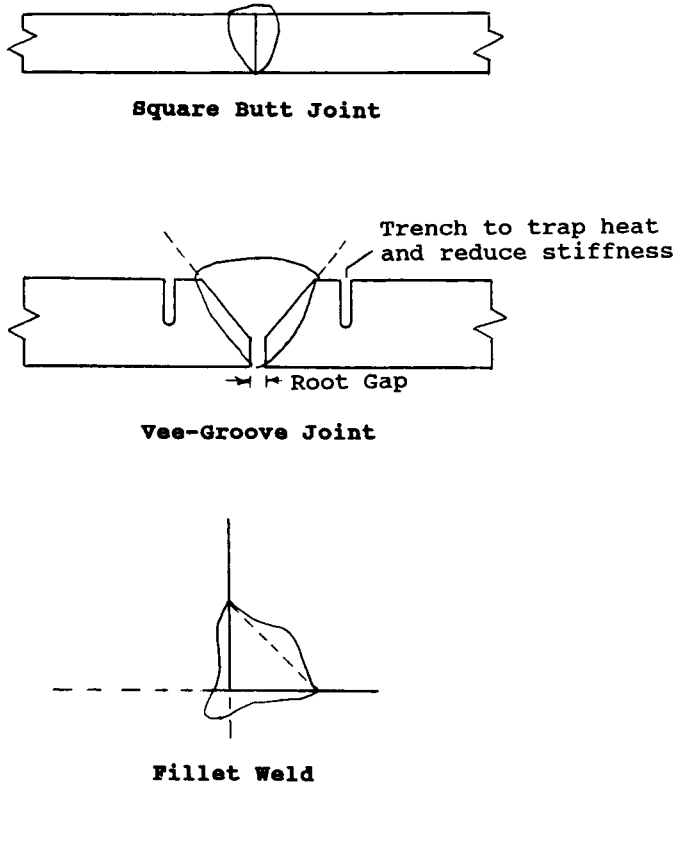
k This refers to 7005 extrusions only (X-prefix still applies to sheet and plate).

4- Where no filler metal is listed, the base metal combination is not recommended for welding.

base metal should be avoided, because thermal contraction forces will cause cracking at the molten edge of the bead [98]. Recommended fillers for various aluminum alloys are given in Table 7. Combinations, for which no filler is specified, should be avoided.

Several other techniques help minimize cracks. Fillet and Vee-groove joints are preferred, because they maximize crack-insensitive filler in the joint and minimize its dilution by crack-sensitive base metal. The weld bead should be slightly convex to promote flow of molten metal in healing incipient cracks. If a bead is allowed to become flat or concave, surface tension can inhibit flow. This can be accomplished by tilting the electrode 10 to 20 degrees and using the force of the arc to "dam up" the molten bead [99]. A crater must be avoided when terminating a seam using any of a number of established techniques. In especially problematic seams, trenches can be cut on either side of the joint (Figure 14). These trap

Fig. 14. Weld joint nomenclature.



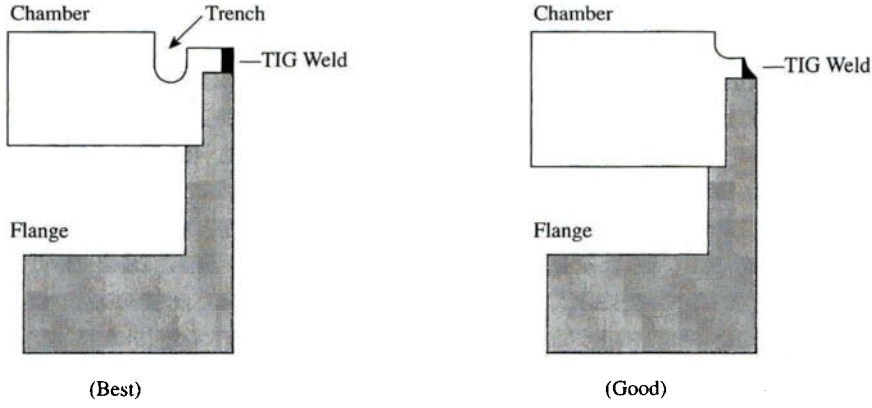
heat around the bead and minimize thermal expansion of the rest of the structure. They also relieve stress by reducing stiffness and permitting some mechanical deflection.

4.8.7.3 Joint Design

Except for a few special considerations, joint design for aluminum is basically the same as for stainless steel. See Figure 14 for nomenclature. Simple, square butt joints (touching) are suitable for thicknesses up to about 6 mm (0.25 in) and are especially desirable because they are easy to scrape clean. Thicker joints should employ Vee-grooves (60–90 degrees) or fillets to assure arc penetration and to minimize dilution of filler by the base metal. A 3 mm ($\frac{1}{8}$ in) root opening, 1.5 to 2 mm from the bottom, helps ensure full penetration. It is desirable for both sides of a joint to be machined to the same thickness for balanced heating that minimizes stress and distortion. Examples of joints between a thin flange neck and a thick chamber wall are shown in Figure 15. Tack welds can be made after fit-up to hold a joint in place. These are essential to prevent distortions that build up during welding from closing root gaps (when used) before a seam is completed. Tack size may be according to operator preference, since during the final weld they must be fully remelted and incorporated into the seam [100]. Seams may be crossed since this does not affect metallurgical soundness in aluminum as it does in other metals [101].

There is a "heat-affected zone" of about 25 mm on either side of a joint where the temper is affected and strength permanently reduced. This is important when welding highly tempered flanges but can be mitigated by the use of a heat sink on

Fig. 15. Recommended joint design for aluminum flanges.



the flange face (such as a mating flange) to keep the temperature down. If a weld seam is required on both sides of a joint for structural reasons, one side should be only skip-welded to prevent trapped gas from causing virtual leaks. After welding, shrinkage will be about twice that of stainless steel (about 6% of the width of a seam). This can cause distortion and should be accounted for by welding pieces together in a symmetrical sequence and machining after welding if strict dimensional tolerances are required.

4.8.7.4 Welding Techniques

Excellent vacuum-tight seams can be achieved with manual TIG welding. An experienced stainless steel welder can acquire the necessary skills in a matter of hours through demonstration and practice. The most important difference is that temperature cannot be gauged by color. Aluminum simply becomes shiny when it melts, and only this quality can be used to determine the extent and temperature of the bead. Welding style is also affected by the fact that molten aluminum tends to hang together like jelly instead of flowing like a liquid, as does stainless steel [102].

Because of its high thermal conductivity, aluminum requires about three times more heat to weld than does stainless steel. AC-TIG (alternating current tungsten inert gas) welding with argon gas is recommended for thin joints because it provides a cleaning action and is relatively insensitive to variations in arc length. A ball tip should be used, since the high arc power tends to melt pointed tips, causing contaminating particles of tungsten to be thrown off. However, for thicknesses of 1 cm and more, the high power required makes AC-TIG welding difficult to control. Instead, DC-TIG with helium should be considered, since it produces much more concentrated heating for better control and smaller beads for less distortion. The presence of helium increases the heat concentration and appears to have a cleaning action that partially makes up for the absence of the plasma cleaning provided by AC-TIG. Preheating is not recommended, because it can destroy the tempered properties of the metal. Trial welds on test pieces are highly recommended.

A dip and transfer motion should be used with AC-TIG welding. After a spot is melted, filler rod is added which, in turn, cools the spot. The tip is held in place while the rod is withdrawn to allow the spot to thoroughly reheat. The tip is then moved forward enough to get good penetration but not so much that the arc burns through the work piece. In DC-TIG welding, filler is supplied continuously, because the intense, concentrated heat is sufficient to maintain bead temperature. In both cases, the tip should be normal to the seam or slightly leading the bead. Filler rod should be 25 to 50% of the width of the bead, depending on the operator's personal preference. Thicker rod provides a little extra stiffness for pushing

filler into square joints. Thin rod requires a high feed rate that can be difficult to control. In general, the amount of heat used in welding should be minimized since excessive heat causes loss of material strength, weld defects, and porosity.

4.8.7.5 Quality Assurance

Compliance with MIL-STD-2219 Class A weld requirements has been found to produce welds that are suitable for UHV applications [103]. These are similar to several other commercial specifications [104,105,106]. The main requirement is for subsurface pores and inclusions (as seen in X-ray photographs) relative to weld thickness (T) to be as follows [107]:

1. Individual size must be less than 0.33T or 0.060 in, whichever is less
2. Spacing must be at least four times the size of the larger adjacent imperfection
3. Accumulated length in any 3 in of weld must be less than 1.33T or 0.24 in, whichever is less.

4.8.7.6 Repairing Defects

Defects that cause leaks can nearly always be repaired by removal with a point grinder, such as a Dremel® tool, and rewelding. Material must be removed 25 mm (1 in) on either side of the defect but need not fully penetrate the seam. The tool used should have a metal cutting tip rather than an abrasive tip, because the latter tends to force contamination into the surface. The cleaning action of AC-TIG welding is particularly useful for removing entrenched contamination. It can also be used to "boil out" hydrogen from porous areas. Care should be taken not to overheat the repair, because this can cause cracking further along the seam that sometimes nccessitates a series of repairs that "chase" a leak down the seam [108].

REFERENCES

1. H. F. Dylla, D. M. Manos, and P. H. La Marche, Correlation of outgassing of stainless steel and aluminum with various surface treatments, *J. Vac. Sci. Technol.*, **A11**(5) (September–October 1993) 2629.
2. John F. O'Hanlon, *A User's Guide to Vacuum Technology*, 2nd ed. (Wiley, New York, 1989), Appendices C.2 and C.3, pp. 445–446.
3. M. Suemitsu, H. Shimoyamada, N. Miyamoto, T. Tokai, Y. Moriya, H. Ikeda, and H. Yokoyama, Ultrahigh-vacuum compatible mirror-polished aluminum-alloy surface: Observation of surface-roughness-correlated outgassing rates, *J. Vac. Sci. Technol.*, **A10**(3) (May–June 1992) 570.

4. M. Mohri, S. Maeda, H. Odagiri, M. Hashiba, T. Yamashina, and H. Ishimaru, Surface study of Type 6063 aluminum alloys for vacuum chamber materials, *Vacuum*, **V34**(6) (1984) 643.
5. F. J. Norton, Gas permiation through the vacuum envelope, *Trans. 8th AVS Vac. Symposium* (Pergamon Press, New York, 1962), p. 14.
6. H. Ishimaru, Developments and Applications for All-Aluminum Alloys Vacuum Systems, *MRS Bulletin*, **15**(7) (July 1990) 24.
7. H. J. Halama and J. C. Herrera, Thermal desorption of gases from aluminum alloy Al 6061, their rates and activation energies, *J. Vac. Sci. Technol.*, **13**(1) (January–February 1976) 466.
8. June-Rong Chen, Chih-Hao Lee, Jyh-Cheng Chen, Huan-Lurn Hsieh, and Yuen-Chung Liu, Outgassing behavior of A6063-EX aluminum alloy and SUS 304 stainless steel, *J. Vac. Sci. Technol.*, **A5**(6) (November–December 1987) 3422.
9. June-Rong Chen and Yuen-Chung Liu, A comparison of outgassing rate of 304 stainless steel and A6063-EX aluminum alloy vacuum chamber after filling with water, *J. Vac. Sci. Technol.*, **A5**(2) (March–April 1987) 263.
10. G. Moraw, The influence of ionization gauges on gas flow measurements, *Vacuum*, **24** (1974) 128.
11. J. R. Young, Outgassing characteristics of stainless steel and aluminum with different surface treatments, *J. Vac. Sci. and Technol.*, **6**(3) (1969) 400.
12. H. Iguchi, T. Momose, and H. Ishimaru, Cl2 gas corrosion and outgassing of various surface treated aluminum alloys, *J. Vac. Sci. and Technol.*, **A11**(4) (July–August 1993) 1708.
13. Chen et al. 1987, Iguchi 1993.
14. Yuko Hirohata, Satoshi Fujimoto, Tomoaki Hino, and Toshiro Yamashina, Hydrogen desorption behavior of aluminum materials used for extremely high vacuum chamber, *J. Vac. Sci. Technol.*, **A11**(5) (September–October 1993) 2637.
15. Ishimaru 1990.
16. Francis J. Norton, Gas permiation through the vacuum envelope, *Trans. 8th AVS Vac. Symposium* (Pergamon Press, New York, 1962) 13.
17. Ishimaru 1990, p. 27.
18. R. N. Peacock, Practical selection of elastomer materials for vacuum seals, *J. Vac. Sci. Technol.*, **17**(1) (January–February 1980) 330.
19. A. Roth, *Vacuum Technology*, 2nd ed. (Elsevier Science, New York, 1982) 398, Equation 7.12.
20. Aluminum Association, *Aluminum Standards and Data*, 1990, Aluminum Association, 900 19th St., N.W., Washington, DC 20006 (Tel: 301/645-0756), p. 30, Table 2.1.
21. J. Holden, L. Holland, and L. Laurenson, Bakeable vacuum seals using aluminum wire gaskets, *Journal of Scientific Instruments*, **36** (June 1959) 282.
22. J. D. Redman, et al. *Vacuum systems techniques and material studies*, Oak Ridge National Laboratory Report 3472. As referenced in Alex Roth, *Vacuum Sealing Techniques* (1966), Reprinted by American Institute of Physics, 1993, Ref. 1036a, p. 309.
23. Cefilac Etancheite, *Seals for Ultra-Vacuum and Cryogenics*, catalog p. 5.
24. I. Sakai, H. Ishimaru, and G. Horikoshi, Sealing concept of elastic metal gasket "Helicoflex," *Vacuum*, **32**(1) (1982) 34.
25. *Aluminum Standards and Data*, p. 30, Table 2.1.
26. J. D. Gonczy, R. C. Ferry, R. C. Niemann, and B. Roop, UHV seal studies for the advanced photon source storage ring vacuum system, Conference Record of the 1991 IEEE Particle Accelerator Conference, May 6–9, 1991, San Francisco, California, p. 2304.
27. W. Unterlerchner, Non-standard vacuum hardware for an accelerator vacuum system, CERN/LEP-VA/89–50.
28. H. Ishimaru, Bakeable aluminum vacuum chamber and bellows with an aluminum flange and metal seal for ultrahigh vacuum, *J Vac. Sci. Technol.*, **15**(6) (November–December 1978).
29. T. Ishigaki, K. Shibuya, H. Sakaue, and S. H. Be, Sealing technique for aluminum to stainless steel CF flange pairs, *J. Vac. Sci. and Technol.*, **A12**(4) (July–August 1994) 1705.

30. Greg Andronaco, Lawrence Berkeley Laboratory, private communication, 1994.
31. Ishimaru 1990, p. 23.
32. M. Suemitsu, H. Shimoyamada, N. Matsuzaki, and N. Miyamoto, Photoemission studies and outgassing-rate measurements on aluminum alloy surfaces lathed with various alcohols, *J. Vac. Sci. Technol.*, **A10**(1) (January–February 1992) 188.
33. A. G. Mathewson, J. P. Bacher, K. Booth, R. S. Calder, G. Dominichini, A. Grillot, N. Hilleret, D. Latorre, F. Le Normand, and W. Unterlerchner, Comparison of chemical cleaning methods of aluminum alloy vacuum chambers for electron storage rings, *J. Vac. Sci. Technol.*, **A7**(1) (January–February 1989) 77.
34. N. Kaufherr, A. Krauss, D. M. Gruen, and R. Nielsen, Chemical cleaning of aluminum alloy surfaces for use as vacuum material in synchrotron light sources, *J. Vac. Sci. and Technol.*, **A8**(3) (May–June 1990) 2849.
35. R. A. Rosenberg, M. W. McDowell, and J. R. Noonan, X-ray photoelectron spectroscopy analysis of aluminum and copper cleaning procedures for the Advanced Photon Source, *J. Vac. Sci. and Technol.*, **A12**(4) (July–August 1994) 1755.
36. John Philip Durrant and Beryl Durran, *Introduction to Advanced Inorganic Chemistry* (Longmans, Green, London, 1962).
37. Dylla 1993, p. 2636 (Appendix).
38. Y. Tito Sasaki, A survey of vacuum material cleaning procedures: A subcommittee report of the American Vacuum Society Recommended Practices Committee, *J. Vac. Sci. Technol.*, **A9**(3) (May–June 1991) 2031, describing SLAC cleaning procedure.
39. Dylla 1993, p. 2636 (Appendix).
40. Tito 1991, p. 2031.
41. Mathewson 1989; Kaufherr 1990; and Rosenberg 1994.
42. Tito 1991, p. 2031.
43. Mathewson 1989; Kaufherr 1989; Rosenberg 1994.
44. Tito 1991, p. 2031.
45. *Aluminum Standards and Data.*
46. *Metals Handbook*, 8th ed., Vol. 3: *Machining*, American Society of Metals, Metals Park, OH, p. 439.
47. George A. Goeppner, APS storage ring vacuum chamber fabrication, AIP Conference Proceedings 236 of the American Vacuum Society Series 12: *Vacuum Design of Synchrotron Light Sources*, p. 127 supported by Reference 3, which was unpublished.
48. E. A. Avallone and T. Baumeister, *Mark's Standard Handbook for Mechanical Engineers*, 9th ed. (McGraw-Hill, New York, 1978) pp. 5–52.
49. Lloyd E. Brownell and Edwin H. Young, *Process Equipment Design: Vessel Design* (Wiley, New York, 1959), p. 143.
50. Ibid., Figure 8.3, p. 144.
51. *Mark's Standard Handbook*, pp. 5–52.
52. Joseph E. Shigley and Charles R. Mischke, *Mechanical Engineering Design*, 5th ed. (McGraw-Hill, New York, 1989), p. 59.
53. *Mark's Standard Handbook*, pp. 6–11 and 6–44.
54. *Aluminum Standards and Data*, p. 40.
55. A. Roth, *Vacuum Sealing Techniques* (AIP Press, Amsterdam, 1966), p. 752.
56. Donald R. Pitts and Leighton E. Sissom, *Heat Transfer* (McGraw-Hill, New York, 1977), p. 17.
57. *Mark's Standard Handbook*, pp. 4–68.
58. Bruce Craig, *Handbook of Corrosion Data* (ASM International, Metals Park, OH, 1989), p. 146.
59. Ibid., pp. 162–163.
60. Ibid., p. 147.
61. Ibid., p. 227.
62. *DECHEMA-Werkstoff-Tabelle/Chemische Beständigkeit* (DECHEMA Deutsche Gesellschaft

für Chemisches Apparatewesen, Frankfurt am Main, Germany, December 1982), pp. Chlor B1.4.
63. Ibid.
64. *Handbook of Corrosion Data*, p. 227.
65. Ibid., p. 229.
66. Philip A. Schweitzer, *Corrosion Resistance Tables*, 2nd ed. (Marcel Dekker, New York, 1986), p. 466.
67. D. J. De Renzo, *Corrosion Resistant Materials Handbook*, 4th ed. (Noyes Data Corporation, Park Ridge, New Jersey, 1985), p. 618.
68. *Handbook of Corrosion Data*, p. 273.
69. DECHEMA, p. Fluor B1.3.
70. DECHEMA, p. Fluor B1.3.
71. *Corrosion Resistance Tables*, p. 466.
72. *Handbook of Corrosion Data*, p. 273.
73. *Corrosion Resistance Tables*, p. 466.
74. *Corrosion Resistant Materials Handbook*, p. 620.
75. *Corrosion Resistance Tables*, p. 568.
76. Ibid., p. 572.
77. *Handbook of Corrosion Data*, p. 326.
78. *Corrosion Resistance Tables*, p. 572.
79. *Handbook of Corrosion Data*, p. 326.
80. Stanslaw Mrowec and Teodor Werber, *Gas Corrosion of Metals*, Translated from Polish and Published for the National Bureau of Standards and the National Science Foundation, Washington, D.C., by the Foreign Scientific Publications Department of the National Center for Scientific Technical and Economic Information, Warsaw, Poland, 1978, Chapter 4.
81. *Metals Handbook*, p. 583.
82. S. M. Fine, Design and operation of UHP low vapor pressure and reactive gas delivery systems, *Semiconductor International*, October 1995, p. 138.
83. *Metals Handbook*, 9th ed., Vol. 13: *Corrosion* (ASM International, Metals Park, OH, 1987), p. 65.
84. Robert H. Perry and Cecil H. Chilton, *Chemical Engineers' Handbook*, 5th ed. (McGraw-Hill, New York, 1973), Section 3.
85. Ibid.
86. Monika Kuhn and P. Bachmann, Demands for turbo-molecular pumps in the aluminum etching process, *Vacuum*, **41**(7–9) (1990) 2028.
87. DECHEMA, Fluor B1.2.
88. Kuhn, p. 2031.
89. Rosemarie Becker and Monika Kuhn, Heated turbomolecular pumps in metal etch applications, *Solid State Technology*, November 1994, p. S13.
90. Aluminum Association, *Welding Aluminum Theory and Practice*, 2nd ed. Aluminum Association, 900 19th St., N.W., Washington, DC 20006 (1991) (Tel: 301/645-0756).
91. Aluminum Association, *Welding Aluminum*, Aluminum Association, 900 19th St., N.W., Washington, DC 20006 (1984) (Tel: 301/645-0756).
92. *Let's Talk About Aluminum Welding* (videotape). Published by ALCOA, available through distributors of AlcoTec, 2750 Aero Park Drive, Traverse City, MI 49684 (Tel: 616/941-4111).
93. Weiping Liu and L. Dorn, Improved filler wires for aluminum alloy welding—a review, Part I, *Aluminum*, **70**(9–10) (1994), 583.
94. Joseph Gagliano, Argonne National Laboratory, private communication regarding advanced photon source welding procedures.
95. Procedure used for advanced photon source storage ring welding at Argonne National Laboratory.
96. M. Kutsuna, J. Suzuki, S. Kimura, S. Sugiyama, M. Yuhki, and H. Yamaoka, CO2 laser welding

of A2219, A5083, and A6063 aluminum alloys, *Welding in the World* (Pergamon Press) **31**(2) (1993) 128.
97. J. H. Dudas and F. R. Collins, Preventing weld cracks in high-strength aluminum alloys, *Welding Journal*, June 1966, p. 3.
98. Ibid.
99. Ibid., p. 10.
100. *Welding Aluminum Theory and Practice*, p. 5.3.
101. Ibid., p. 10.14.
102. Daniel Van Lanen, Senior Welder, Argonne National Laboratory, Verbal interview November 4, 1994. Provided most of the specific suggestions and insights regarding welding techniques based on his many years of experience.
103. Joseph Gagliano, manager of APS welding, Argonne Laboratory, private communication, 1994.
104. ANSI/AWS D1.2, Structural welding code—aluminum, American Welding Society, 550 N.W. LeJeune Rd., Miami, FL 33126 (Tel: 800/334-9353).
105. ANSI/AWS A5.10-92, Specification for bare aluminum and aluminum alloy welding electrodes and rods, American Welding Society, 550 N.W. LeJeune Rd., Miami, FL 33126 (Tel: 800/334-9353).
106. Qualification Standard for Structural Welding of Aluminum, Publication 25, the Aluminum Association, 900-19th St., N.W., Washington, DC 20006 (Tel: 301/645-0756).
107. MIL-STD-2219, Fusion welding for aerospace applications, U.S. Department of Defense, Available at no charge through Naval Publications (Tel: 215/697-2179).
108. Van Lanen.

CHAPTER 4.9

Preparation and Cleaning of Vacuum Surfaces

Donald M. Mattox
Management Plus, Inc.

Surfaces that are in contact with the vacuum environment are called vacuum surfaces. These surfaces may be associated with the vacuum system, fixturing associated with a vacuum-related processing or the parts related to vacuum processing. Thus they may be a metal, ceramic, glass, semiconductor, or polymer. Surfaces used for vacuum applications should be prepared such that they do not contribute to contamination in the system [1–5]. Some surfaces, such as the fixtures, can be easily or routinely removed for cleaning, while others, such as the vacuum chamber and associated feedthroughs, are nonremovable and must be cleaned in place. Often the surface areas of the fixtures and parts are greater than that of the vacuum system, and these surfaces may determine the pumping performance and the residual gas and vapor contamination in the vacuum system. This type of contamination can be called "introduced contamination" or "process-related" contamination as converse to "system-related" contamination associated with nonremovable surfaces.

Surface preparation includes cleaning and surface modification. Cleaning means reducing the contamination level on the surface to an acceptable level for application of the vacuum environment. Surface modification can involve changing the surface morphology to be more rough or smooth, changing the chemical composition of the surface, changing the outgassing or outdiffusion properties of the material, or changing the mechanical properties of the surface.

ISBN 0-12-325065-7
$25.00

Surface roughness is important to the vacuum performance of a vacuum material. Surface roughness not only increases the surface area available for adsorption of contaminants, it also traps small particulates, making them more difficult to remove. A rough or porous surface also allows capillary condensation of water, making it more difficult to remove water vapor from the system. During fabrication of surfaces for use in a vacuum environment, the surface should be made as smooth as possible. For metals this means light machine cuts, sharp machining tools, cutting lubricants, etc. Fully dense ceramics and glasses can be formed with smooth surfaces by using fine-grit grinding and polishing materials in the final stages of forming. After fabrication the surface can be further smoothed by mechanical [6], chemical, electrochemical [7,8] or chemical/mechanical [9] polishing.

Mechanical polishing can be done by machine or by hand using a wet or dry polishing procedure. The polishing material may be in the form of a powder, contained in a slurry, loosely held in a polishing cloth, or rigidly held on a solid material such as a paper. Table 1 lists a number of common abrasive and polishing compounds and their size and nomenclature.

The cleaning and treatment of the surfaces should be spelled out in the operation and maintenance instructions for the system. This is particularly true for aluminum.

4.9.1 SURFACE MODIFICATION

4.9.1.1 Nonremovable Surfaces

Nonremovable surfaces are vacuum surfaces that, once assembled, cannot be removed or even reached without great difficulty.

STAINLESS STEEL

Stainless steel is the preferred material for nonremovable metal surfaces since it forms a natural passive oxide and is resistant to corrosion. There are many stainless steel alloys, such as

- 304. Common machinable alloy, nonmagnetic—carbides can precipitate in weld areas, which can result in galvanic corrosion if electrolytes are present.
- 304L. Low carbon—used for better intergranular corrosion resistance than obtained with 304. Use for fluid lines and gas lines containing moisture.
- 316. For general corrosion resistance—do not mix 304 and 316 when used in fluid transport because of galvanic corrosion at joints.

Table 1

Average Particle Size (microns)

Grit/Mesh Size	Grit/Mesh Emery	Al_2O_3, SiC, Garnet	B_4C	Diamond
Coarse Abrasive Polish				
24		1035	710	1036
30		930	590	
36		810	500	712
46		508	350	508
54		430	297	
60		406	250	407
70		328	210	
80		266	177	267
90		216	149	
100		173	125	213
120		142	105	142
150		122	74	122
Fine Abrasive Polish				
180		70–86	62	86
220		66	53	66
240		54–63	42.5	60
280	1	44	35	
320	1/0	29–32	27.7	52
400	2/0	20–23	16.3	40
600	3/0	12.5–17	8.3	14
800	4/0	9–12	5.5	8
1000		5–9	3.7	
1200		1.5–5	2.5	6
1500		1–1.5		4
2000		1.0		2
				1
				0.5
				0.25

- 316L. Better intergranular corrosion resistance than 316.
- 303. Has a high sulfur content and a higher tendency for porosity—do not use, it cannot be welded very well.
- 440. Hardenable but magnetic and more prone to corrosion than the 300 series.

Stainless steel is available as mill plate with several finishes.

- Unpolished #1—very dull finish produced by hot-rolling the steel followed by annealing and descaling. The surface is very rough and porous. Used where surface finish is not important.

- Unpolished #2D — Dull finish produced by a final cold roll after the hot rolling but before annealing and descaling. Used for deep drawing where the surface roughness retains the drawing lubricant.
- Unpolished #2B — Bright finish obtained by a light cold roll after annealing and descaling. Grain boundary etching due to descaling still present. General-purpose finish.
- Polished #3 — Intermediate polish using 50 or 80 grit abrasive compound. R_{max} of 140 microinches. Heavy polishing grooves.
- Polished #4 — General-purpose surface obtained with 100–150 grit abrasives. R_{max} of 45 microinches. Lighter polishing grooves.
- Buffed #6 — Polished with 200 grit abrasive.
- Buffed #7 — Polished with 200 grit abrasive with a top dressing using chrome oxide rouge. R_a of 8–20 microinches.
- Buffed #8 — Polished with 320 grit abrasive (or smaller) with an extensive top dressing using chrome oxide rouge. R_a of 4–14 microinches.* To the eye, the surface appears to be free of grinding lines.

Stainless steel plate is often ground to form a flat surface and then plates are welded to form a chamber. Stainless steel plates may also be shaped by deformation, which will create a wrinkled surface morphology and may trap lubricants in the subsurface region. Deformation or machining of stainless steel will work-harden the surface. This is desirable where the surface must provide a cutting edge such as when used on CF flanges or for providing a deformation and wear-resistant surface. Heating to above 450°C will anneal the work-hardened surface, and the hardness will be lost.

The surface of the steel can be mechanically polished to improve the smoothness. It can then be chemically polished or electropolished to make it more smooth. Electropolishing [10] decreases the R_a by about a factor of 2 and eliminates many of the microcracks, asperities, and crevices in the polished surface. Typically electropolishing is done in an electrolyte containing phosphoric acid and the smooth areas are protected by a thin phosphate layer, causing the peaks to be removed. This phosphate layer should be removed using an HCl rinse and then the surface rinsed to an acid-free condition prior to use. Directed streams of electrolyte ("jets") can be used to selectively electropolish local areas of a surface [11]. Commercial suppliers provide electropolishing services to the vacuum industry either at their plants or on site at the customer's plant.

*The center line average (CLA) or R_a is the average peak height above the center line of a profile over the surface. The center line is defined as the line that divides the profile in such a way as the net deviation is zero. The R_a tells nothing about maximum peak heights (R_{max}), slopes or the lateral dimensions of the peaks (waviness).

Typical electropolishing solutions for stainless steel are

1. Phosphoric acid	5 parts
Sulfuric acid	4 parts
Glycerin (USP)	1 part
Current density, amps/ft^2	450

or

2. H_2SO_4 (1.84 s.g.)	1000 ml
H_2O	370 ml
Glycerin (USP)	1370 ml

Note: Add acid slowly to water (avoid overheating), then add glycerin. Use carbon or lead cathode. Polish at 7.5 volts for about 30 sec.

Stainless steel can also be chemically polished. A chemical polish for stainless steel is [12]:

- Clean in a hot alkaline solution.
- Rinse.
- Activate in a hot 5% sulfuric acid solution for 5 minutes before polishing.
- Chemically polish at 75°C in a solution of

Nitric acid	4 parts
Hydrochloric acid	1 part
Phosphoric acid	1 part
Acetic acid	5 parts

Electropolishing decreases the actual surface area available for adsorption and reduces the contamination retention of the surface. The electropolished surface generally exhibits a lower coefficient of friction than a polished surfaces. The various surface treatments can alter the outgassing properties of the stainless steel surface [13–17]. The surface condition can be determined by scanning electron microscopy (SEM, 1,000–50,000×) for surface morphology and auger electron spectroscopy (AES) or X-ray photoelectron spectroscopy (XPS or ESCA) for

surface composition. The chemical composition of electropolished surfaces can be specified for critical applications [18,19].

Electropolishing, as well as acid treatments, "charge" the steel surface with hydrogen, and for UHV applications the stainless steel should be vacuum-baked at 1000°C for several hours to outgas hydrogen taken up by the surface [13]. The surface of stainless steel will form a natural passive oxide layer when dried and exposed to the ambient. The passive layer can be improved by heating in air. However, control of the temperature and dewpoint is very important. A smooth oxide film is formed on 316L stainless steel at 450°C and a dewpoint of −100°C but small nodules and surface coarsening result when the oxidation is done above 550°C in air with this dew point [20]. Type 304 and 316 stainless steels are more easily passivated than are the 400 series (hardenable) stainless steels [21].

The natural oxide on stainless steel can be removed thus:

- Vapor clean in trichloroethane (5 minutes).
- Rinse in cold water.
- Hot alkaline cleaner 5 minutes.
- Rinse in hot water.
- Potassium permanganate (100 ml DI water + 50 g NaOH + 5 g $KMnO_4$ at 95°C)—soak to condition oxide scale.
- Hydrochloric acid dip to sensitize surface (remove natural oxide passivation).
- Pickling solution (30 vol% HNO_3 + 3 vol % HF) at room temperature for 30 minutes.
- Rinse in hot, deionized water.

LOW-CARBON STEEL

Low-carbon or mild steel (such as 1018 steel) is low cost with respect to stainless steel and is often used in applications involving large vacuum vessels. Sometimes the steel is porous, usually due to "stringers" in the metal plate. This problem is generally masked by painting the exterior of the chamber to reduce the real leak rate. Removal of this paint can open up the porosity and increase the leak rate. Generally, when using such steels, the surface is used in the as-machined condition. Low-carbon steel can be cleaned by wiping with neutral pH perchloroethane or trichloroethane, followed by an acetone wipe followed by drying with anhydrous methanol. Water-based cleaning should be avoided because it can contribute to rusting of the steel.

Carbon steel and low alloy steels can be cleaned by pickling in a hydrochloric acid bath (8–12 wt %) at 40°C for 5–15 min to strip the oxide from the surface [22]. A simple technique for removing localized iron rust is (1) solvent clean, (2) soak (or wet) in acetic acid, (3) brush away residue, and (4) repeat as necessary. Low-carbon steel can also be electropolished, but this is not typically done.

ALUMINUM ALLOYS

Aluminum alloys are not normally used for vacuum processing systems because they are soft and easily corroded, especially by chlorine-containing chemicals. With proper fabrication, aluminum has proven to be a good high- and ultra-high-vacuum material [23]. However, mill-rolled aluminum has an outgassing rate ~100 times that of mill-rolled stainless steel [24]. A dense, thin oxide with good outgassing properties can be formed on aluminum surfaces by (1) machining under a dry, chlorine-free argon/oxygen gas, (2) machining under pure anhydrous ethanol, or (3) extrusion under a dry, chlorine-free argon/oxygen gas [25–27].

The natural oxide on aluminum can be removed (stripped) before polishing. A chemical strip for the oxide on aluminum is as follows:

- Soak in solution of 5% NaOH by weight at 70–75°C.
- Soak in a solution of 1 part concentrated HNO_3 to 1 part deionized water at 20°C, followed by a dip in a solution of 1 part concentrated HNO_3 with 64 g/liter NH_4HF_2 at 20°C (desmutting procedure—Cu and Si).
- Rinse well.

The aluminum can be chemically polished thus:

1. • Soak in 10% HCl.
   • Rinse in deionized water.

or

2. • Make a solution of

H_3PO_4	80%
CH_3COOH	15%
HNO_3	5%

   • Raise to temperature 90 to 110°C.
   • Dip for 2–4 min.

Heavily corroded aluminum surfaces can be electrocleaned thus:

- Pickle in 5% NaOH solution at 75°C
- Wash in 30% HNO_3.
- Dip in 12% H_2SO_4.
- Follow with an anodic electroetch at 90°C in a solution of 100 g H_3BO_3 and 0.5 g borax in 1 liter deionized water starting at 50 volts and increasing to 600 volts.

Anodized Aluminum

In special cases where the surface hardness must be increased or chemical corrosion resistance is necessary (e.g., plasma etching with chlorine), anodized aluminum surfaces can be useful [28,29]. Alloying elements, impurities and heat treatment can influence the nature and quality of the anodized coating—typically, the purer the aluminum alloy, the better the anodized layer. To build up a thick anodized layer on aluminum, it is necessary for the electrolyte to continuously corrode the oxide, giving a porous oxide layer. ASTM Specification B-580-73 designates seven thicknesses (up to 50 microns) for anodization with letter designations of *A* through *G*. Anodization baths for the various thicknesses are

- Oxalic anodize—very thick films (50 microns)
- Sulfuric acid—thick films (80% aluminum oxide, 18% aluminum sulfate, 2% water—15% porosity)
- Chromic acid—thin films (1–2 microns)
- Phosphoric acid—very porous films (base for organic coatings)

After formation, the porous aluminum oxide can be "sealed" by hydration, which swells the amorphous oxide. Sealing is done in hot (95–100°C) deionized water (sulfuric acid anodize only) or by the use of sodium dichromate solution (improve corrosion resistance) or with nickel or cobalt acetate solutions. Sealing reduces the hardness of the anodized film. Steam sealing can be used to avoid the use of nickel-containing hot water to prevent the possibility of nickel contamination in semiconductor manufacturing. For vacuum use, the anodized surface should be vacuum-baked before use. To increase the corrosion protection or lubricity of the anodized surface, other materials can be deposited in the porous surface. Examples are the Magnaplate® coating to improve corrosion protection and Tufram® coating using to improve the frictional properties of anodized aluminum surfaces.

Anodized aluminum does not provide a good surface for sealing with elastomer seals. In anodized systems, the sealing surfaces are often machined to reveal the underlying aluminum. These surfaces can be protected from corrosion with a thin layer of a chemically resistant grease such as Krytox®.

Aluminum can be anodized with a dense oxide (barrier anodization) [30] but this technique has not been evaluated for vacuum application, because it is rather thin.

COPPER

Copper is often used in vacuum systems as an electrical conductor or as a shear-sealing material. The polishing of copper often means the removal of the oxide layer. Copper can be polished (smoothed) thus:

- Immerse in a solution of
  60 ml phosphoric acid (s.g. 1.75)

10 ml nitric acid (s.g. 1.42)
10 ml acetic anhydride and
8 ml water
for 4 minutes at room temperature

The oxide can be removed from copper thus:

1. • Clean in perchloroethylene.
   • Ultrasonic clean in alkaline detergent (pH = 9.7) at 60°C for 5 to 10 minutes.
   • Rinse.
   • Deoxidize in 50 vol % HCl at room temperature for 5 to 10 minutes.
   • Rinse.

or

2. • Solvent clean.
   • Immerse in solution of

     60 ml phosphoric acid (s.g. 1.75)

     10 ml nitric acid (s.g. 1.42)

     10 ml acetic anhydride

     8 ml water

     for 4 minutes at room temperature
   • Rinse.

CERAMICS AND GLASSES

Ceramics and glasses develop surface microcracks when ground or polished. These microcracks reduce the strength of the material as well as contribute to surface retention of contamination. Oxide ceramics and glasses can be etched in a solution of hydrofluoric acid (HF) or ammonium bifluoride, which will mildly etch the surface and blunt the microcracks.

POLYMERS

The use of polymers in the vacuum system should be minimized. If used, the polymers should be outgassed by vacuum baking before being installed in the sys-

tem. Vacuum baking of polymeric materials should be done at temperatures below the degradation temperature of the material. The time necessary for outgassing will vary with the history of the material. In order to determine the necessary time and temperature for outgassing studies should be made using weight loss as a function of time and temperature [31,32]. Water vapor and solvents are the most common vapors outgassed from polymers but in some cases low-molecular-weight polymers can be evolved. A material can be said to be "outgassed" when it has less than 1% weight loss after being held at 25°C above the expected operating temperature for 24 hours at 5×10^{-5} torr (ASTM E595-90), though some applications may require better outgassing. Another technique is to monitor the outgassing products as a function of time and temperature.

4.9.1.2 Removable Surfaces

Removable surfaces are those that are routinely removed, are being processed, or that can be removed easily for cleaning such as fixtures, tooling, liners, and shields. In some cases the removable surface can be discarded, such as when aluminum foil is used in the vacuum system. Note that typical aluminum foil is coated with a vegetable oil and must be cleaned before being used in the vacuum system.

Common materials for removable surfaces are stainless steel and aluminum with some use of titanium. Occasionally a polymer may be used for fixturing and often a polymer is being processed in a vacuum system. If possible, the removable surfaces should be smoothed, as has been discussed. If there has been no contamination of the vacuum surface during processing, the surface that has been removed from the chamber can be cleaned in the manner discussed in Section 4.9.2.3. When using the vacuum system for film or coating deposition, plasma, or low-pressure chemical vapor deposition (CVD), the surface will have to be "stripped" of the deposit buildup.

LINERS AND SHIELDS

The nonremovable surface should be protected from corrosion or abrasion. This may necessitate the use of liners and shields in the system to protect the surface from the processing environment or minimize the need for subsequent cleaning.

STRIPPING

Stripping is the term given to the removal of large amounts of materials from a surface, usually by chemical or mechanical means. The most simple stripping technique is to apply an adhesive tape and pull the deposit buildup from the sur-

face. This can be assisted by having a release agent on the surface. Common release agents are carbon [33] and boron nitride (e.g., Combat®) applied to the vacuum surface in a water slurry. Release agents can also be applied to nonremovable surfaces and can be applied by glow discharge decomposition of a hydrocarbon vapor [34]. The oxide on the surface of stainless steel acts as a natural release agent for films of deposited materials such as copper or gold that are not very oxygen active.

Deposit buildups can also be removed by abrasion with grit blasting and dry or wet glass bead blasting [35–37] being common techniques. Dry glass bead blasting is a commonly used cleaning technique, but, as with other grit abrasive techniques, it can leave shards of glass embedded in soft surfaces. The bead blasting can also flow the surface and trap oil contamination if the surface is not clean before bead blasting. The amount of grit embedded depends on how long the glass beads have been used, i.e., how much they have been fractured. Water-soluble particles can be used for abrasive cleaning and allow easy removal of the water-soluble embedded particles. For example, 5 micron sodium bicarbonate (baking soda) particles entrained in a high-velocity water stream can be used for mild abrasive cleaning. Polymer beads can be used in some cases [38].

Grit blasting uses grit such as fractured cast iron, alumina, silica, plastic, etc. of varying sizes and shapes accelerated in a gas stream to deform and gouge the surface [39]. Particles can be entrained in a high-velocity gas stream by using a siphon system or a pressure system such as used in sand blasting equipment. In addition to removing gross contamination, grit blasting roughens the surface. The Society of Automotive Engineers (SAE) has specifications on grit size and type, e.g., SAE Specification J444 for cast iron grit in the range G10 (2820 microns maximum size) to G325 (120 microns maximum size), which gives the percentage by weight allowed on standard screens. Bombardment of a surface by grit is like "shot peening" and places the surface in compressive stress, which may produce unacceptable distortion of thin materials.

In some cases the surfaces of fixtures are deliberately roughened so as to prevent the easy removal of deposit buildup, since flaking of deposited material can be a source of particulates in the vacuum system. Roughening is typically done using grit blasting.

Chemical etching can often be used to remove the deposit buildup [40] without attacking the underlying material. Table 2 lists a number of etchant solutions that can be used to remove the material indicated. Also listed are some plasmas that can be used to remove the material indicated. Chemical etching is also used to remove coatings from coated parts to "rework" the parts.

In extreme cases the surface must be cleaned by aggressive techniques such as mechanical abrasive cleaning. Wet or dry mechanical abrasion using an abrasive pad such as Scotchbrite®, or emery paper can be used, although these techniques abrade the surface and increase the surface area.

Table 2

Chemical Etchants for System/Fixture/Substrate Stripping

Material to Be Removed	Etchant	Ratio (vol.)	Useful on These Surfaces	Can Damage
Al	$H_3PO_4/HNO_3/H_2O$	20:2:5	stainless steel (SS), glass (G), ceramic (C)	Cu
	NaOH	Molar	SS,G,C	Ti, Ag
	BCl_3 (plasma)			
C	H_2O_2	10–30%	SS,G,C	Cu, Fe
	KOH/H_2O	saturated/hot	G,C	
	O_2 (plasma)		SS,G,C	Ag, Cu
	H_2 (plasma)		SS,G,C	
Cr	HCl/glycerine	1:1	SS,G,C,Cu	Fe
	$KMnO_4/NaOH/H_2O$	5gm: 7.5gm: 30mL	SS,G,C	Al
Cu	HNO_3/H_2O	1:1	SS,G,C	Fe
Au	HCl/HNO_3 (aqua regia)	3:1	G,C	SS, Cu
Fe	HCl/H_2O	1:1	SS,G,C	
Mo	$HNO_3/H_2SO_4/H_2O$	1:1:3	SS,G,C	Cu, Fe
	H_2O_2	10–30%	SS,G,C	Cu, Fe
Ni	$HNO_3/C_2H_4O_2/C_3H_6O$	1:1:1	SS,G,C	Cu
Pd	HCl/HNO_3	3:1	G,C	SS, Cu
Ag	NH_4OH/H_2O_2–30%	1:1	SS,G,C,Cu	
Ta	HF/HNO_3	1:1	SS	G, C, Cu
Ti	NH_4OH/H_2O_2–30%	1:2	SS,G,C,Cu	
W	HF/HNO_3	1:1	SS	G, C, Cu
	H_2O_2	30%	SS,G,C	Cu, Fe
	$CF_4 + O_2$ (plasma)		SS	Cu, G, C
Si	HF/HNO_3	1:1	SS	G, C, Cu
	$CF_4 + O_2$ (plasma)		SS	Cu, G, C
Ti-W	H_2O_2	30%	SS,G,C,Al	
TiN	H_2O_2/NH_4OH (30%)/H_2O	1:1:1	SS,G,C,Al	
	HF/H_2O	1:1	SS,Cu	G, C
	$CF_4 + O_2$ (plasma)		SS	Cu, G, C
NiCr	$HNO_3/HCl/H_2O$	1:1:3	SS,G,C,Cu	
SiO_2	HF/H_2O	1:1	SS,Cu	G, C
	CF_4 (plasma)		SS,Cu	G, C
ZrO_2	H_2SO_4/H_2O	1:1	G,C	
	HF/H_2O	1:1	SS,Cu	G, C
MgF_2	HNO_3/H_2O	1:1	G,C	
SiO	HF/H_2O	1:1	SS,Cu	G, C
Cd plating	NH_4NO_3/H_2O	120g/liter	steel, brass, Cu brass,	
Zn plating	HCl/H_2O	120ml/liter	Cu alloys	

SURFACE MECHANICAL PROPERTIES

The mechanical properties of a surface can affect the generation of particulates in the vacuum system. Friable (easily fractured) surfaces will generate particulates due to abrasion and wear in the system. System vibration can contribute to this source of particulates.

Surface wear is a function of the porosity, surface hardness, friction coefficient, and fracture toughness of the surfaces in contact. To reduce wear, the surfaces should be fully dense, be as hard as practicable, and have a compressive stress. Surfaces of metals can be hardened by shot peening, which also introduces a compressive stress, by coating with a hard material, and by chemical conversion such as nitriding and boriding [41].

Surfaces can be hardened and dispersion strengthened by forming nitride, carbide, or boride dispersed phases in the near-surface region by thermal diffusion of a reactive species into the surface. Steels that contain aluminum, chromium, molybdenum, vanadium, or tungsten can be hardened by thermal diffusion of nitrogen into the surface. Typically, nitriding is carried out at 500–550°C for 48 hours in a gaseous atmosphere, giving a "case" depth of several hundred microns. In carburizing, the carbon content of a low-carbon steel (0.1 to 0.2%) is increased to 0.65 to 0.8% by diffusion from a carbon-containing vapor at about 900°C. Carbonitriding can be performed on ferrous materials by diffusing both carbon and nitrogen into the surface. Nitrogen diffuses faster than the carbon, so a nitrogen-rich layer is formed below the carbonitrided layer and, if quenched, increases the fatigue strength of the carbonitrided layer. Hardening by boronizing can be done on any material having a constituent that forms a stable boride such as Fe_2B, CrB_2, MoB, or NiB_2.

Using a plasma for ion bombardment enhances the chemical reactions and diffusion rates and also allows heating of the substrate and *in situ* cleaning by sputtering. Typically a plasma containing NH_3, N_2, or N_2–H_2 (9:1, i.e., forming gas) is used along with substrate heating to 500–600°C to nitride steel. The term "ion-itriding" has been given to the plasma-nitriding process. Bombardment from a nitrogen plasma can be used to plasma-nitride a steel surface before the deposition of a PVD-TiN coating. Plasma carburizing is done in a carbon environment. Table 3 shows some properties of materials that have been nitrided and borided.

The coefficient of friction can be lowered by coating with a low-friction, low-vapor-pressure solid material such as molybdenum disulfide [42]. In high-torque applications, the surfaces can be coated with a low-shear metal, such as gold, silver, or lead, which will act as a lubricant. It should be noted that graphite is not a good lubricant in vacuum because of the lack of moisture. Oils will creep away from the point of contact.

Table 3

Hardening of Surfaces by Diffusion

Treatment	Substrate	Microhardness (kg/mm^2)	Case Thickness (microns)
Carburizing	Steel: Low C, Med C, C-Mn, Cr-Mo, Ni-Mo, Ni-Cr-Mo	650–950	50–3000
Nitriding (ion)	Steel: Al, Cr, Mo, V or W, austinic stainless	900–1300	25–750
Carbonitriding	Steel: Low C, Med C, Cr, Cr-Mo, Ni-Cr-Mo	550–950	25–750
Boriding	Steel: Mo, Cr, Ti, cast Fe; Cobalt-based alloys; Nickel-based alloys	1600–2000	25–500

4.9.1.3 Parts to Be Processed

When a vacuum system is to be used for processing, the parts introduced will contribute to the gas load and contamination in the system. This contribution should be consistent from run to run if the processing is to be reproducible. The gas load and contamination from the processing with a specific number of parts and associated fixturing should be determined and factored into the initial design of the vacuum system. The surface of parts to be processed in the vacuum system can be modified to change their morphology, chemistry, or mechanical properties.

SURFACE MORPHOLOGY

The surface morphology of the parts being introduced into the vacuum system can be important to the "introduced contamination." Rough surfaces will have a high surface area, which will affect the removal of water vapor during pumping and will tend to retain particulates, which may be released in the vacuum system.

SURFACE CHEMISTRY

Outgassing and outdiffusion to the surface from the bulk can be a major concern to the performance of the vacuum system, particularly if there is heating from the operation of the system or from vacuum processing. Metals typically outgas hydrogen, which is often of no importance to the vacuum application.

Polymers outgas water vapor, solvent vapors, and absorbed gases. These can often be removed to a tolerable level by vacuum baking before inserting the polymer into the vacuum system. In some cases, the vacuum pumping system will have to tolerate the gas/vapor load from this source during processing. Polymers can also diffuse (outdiffuse) low-molecular-weight fractions to the surface without

volatilization. These surface contaminants can then affect the processing by volatilizing when heated. In some cases, these contaminants can be sealed in by using a "basecoat" on the polymer surface.

4.9.2 EXTERNAL CLEANING

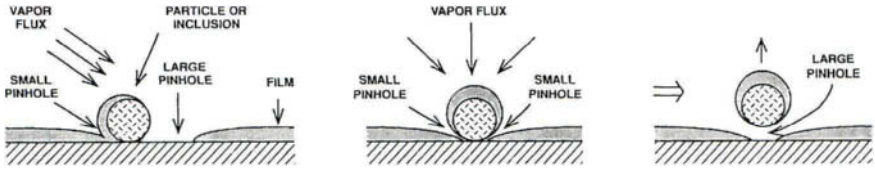
Cleaning is the reduction of contamination to a tolerable level for the application for which the vacuum is needed. External or *ex situ* cleaning is defined as the use of cleaning processes that are performed in the ambient environment or are associated with a separate piece of cleaning equipment. *In situ* cleaning processes are done in the vacuum system used for processing. Generally *in situ* cleaning is used to remove recontamination that occurs after *ex situ* cleaning. However in some cases, such as when the system has been contaminated by oil backstreaming, plasma etching, or plasma-enhanced chemical vapor deposition (PECVD) processing, *in situ* cleaning can be used for cleaning very contaminated vacuum surfaces without disassembly.

External cleaning can vary from very benign cleaning, such as wiping with ethanol, to the very aggressive cleaning where the surface is etched [43]. It is not feasible nor economical to try to achieve ultra-high-vacuum type cleanliness if the requirements do not require it. In some cases a vacuum component, such as a fixture used in vacuum deposition, may be cleaned many times. Over time, this can result in changes in the surface, such as increasing the surface roughness. Often, when using pieces, parts, and components purchased from a supplier, the cleaning process can be simplified by specifying cleaning, handling, and packaging to be provided by the supplier. This may be an increased cost but can reduce the problems of the system builder. Commercial cleaning and packaging services for vacuum applications are available.

4.9.2.1 Contaminants and Contamination

A vacuum system is often used for some type of processing such as vacuum heat treating, freeze drying, or physical vapor deposition (PVD). A contaminant is a material on the surface or in the gas phase that interferes with processing or results in an unacceptable product. As a practical matter, a "clean" surface is one that contains no significant amounts of contamination; thus what constitutes a clean surface depends on the requirements. For PVD processing, general contamination, such as a surface layer, can cause a low nucleation density of the depositing film on a surface, causing poor overall adhesion of a film to a surface and pre-

Fig. 1.



Pinhole formation and pinhole flaking of material deposited on a particle or inclusion on a surface.

venting good electrical contact in the case of deposited electrical contacts. Local contamination (film or particle) can result in locally poor adhesion of a film to a surface giving pinholes in the film. For example, Figure 1 shows the effect of particulate contamination on the formation of pinholes in a vacuum-deposited thin film such as a mirror coating or semiconductor metallization. If the particle flakes off in the vacuum system ("pinhole flaking"), it produces particulate contamination in the system. Cleaning should address local surface conditions such as porosity, embedded particles, steps, roughness, etc., that affect film properties, produce pinholes, and local loss of adhesion. Contaminants in the vacuum system can be categorized as

- Gaseous contamination (e.g., residual gases)
- Vapor contamination (e.g., water vapor, oil vapors)
- Fluid contamination (e.g., oils on surfaces)
- Solid contamination (e.g., organic and inorganic particulates)

System-related contamination is contamination whose origin is the vacuum system. Process-related contamination is contamination associated with the material being processed and materials or fixtures used in the processing and is "brought into" the vacuum system.

4.9.2.2 Cleaning Nonremovable Surfaces

One hopes that nonremovable surfaces will not require drastic cleaning techniques even with repeated use. Particulates can be removed from the vacuum chamber using a dedicated vacuum cleaner with an efficient filter material such as HYSURF® (DuPont).

STAINLESS STEEL

Stainless steel vacuum surfaces that have been contaminated by casual contact with the environment can be cleaned without disturbing the natural oxide by wash-

ing with a neutral detergent solution followed by a thorough rinse in deionized water and drying using anhydrous ethanol or by wiping with neutral pH perchloroethane or trichloroethane, followed by an acetone wipe followed by drying with anhydrous ethanol or methanol [4].

OTHER MATERIALS

Aluminum [23], copper, glass, and ceramics can be cleaned by using neutral pH perchloroethane or trichloroethane, followed by acetone followed by a rinse with anhydrous ethanol. Polymers such as neoprene, Viton®, and Teflon® should be cleaned with methanol or isopropanol and, if low outgassing is a requirement, they should be vacuum baked before use. For elastomer sealing, it may be desirable to very lightly grease the seal with a low-vapor-pressure lubricant before assembly.

4.9.2.3 Cleaning of Removable Surfaces

Surfaces should be cleaned as thoroughly as possible external to the vacuum system. External cleaning includes "gross cleaning" to remove large amounts of contaminants, often by removing some of the surface material, and "specific cleaning," which is directed toward removing specific contaminants such as particulates or hydrocarbons. A clean processing environment and proper handling and storage after the external cleaning are important to minimizing recontamination of the cleaned surface before it is placed in the vacuum system.

REMOVAL OF PARTICULATES

The ability to remove particles from a surface depends on the size, shape, and composition of the particle as well as the surface to which it adheres [44–46]. Removal of particulate contaminants from a surface is best done by mechanical disturbance in a flowing fluid environment [47]. The mechanical disturbance should be done in a fluid environment containing detergents and wetting agents, and the fluid should be continually filtered. Dry or wet brushing is often used for particle removal. Camel hair and mohair are used for dry brushing. Polypropylene, Teflon®, and Nylon® are used for wet brushing. Mechanical scrubbing is often combined with high-pressure fluid jets (2000–3000 psi) as a cleaning procedure.

Blowoff techniques for particle removal have the advantage that they can be done after the parts have been placed in fixtures and even in the deposition system. The best means of blowoff is to use filtered gas from a liquid nitrogen tank. The gas is filtered with a 0.2-micron or smaller filter in the nozzle, and the nozzle should allow ionizing of the gas with a radioactive or electronic source. Ionized

gas should be used when blowing off insulator/organic surfaces to prevent electrostatic charge buildup on the insulator surface.

Generally fluid sprays are effective for removing large particles but are not effective on submicron-sized particles [47]. In ultrasonic cleaning (see later section), the jetting action by the collapsing cavitation bubble acts as a high-pressure fluid spray that displaces the particle. Ultrasonic jetting is good at removing large particles but as the particle size decreases to submicron, the cleaning effectiveness decreases.

Particles can be removed from surfaces by covering the surface with a liquid polymer, allowing it to solidify, then stripping the polymer from the surface. This technique is used by the optics industry to remove particles from mirror surfaces and protect surfaces from abrasion during assembly [48]. There are many types of strip coats, each coating leaving different residues after stripping and having differing corrosion compatibility with surfaces.

REMOVING CONTAMINANT FILMS

Abrasives

Abrasive powders in a paste or fluid carrier are useful in removing contaminant films. Precipitated calcium carbonate ($CaCO_3$), CeO [49], and Snow-Floss® (diatomaceous earth with the calcium carbonate removed, leaving a friable silica network) are used to clean glass surfaces with a mild abrasive action.

The use of CO_2 "snow" as a mild abrasive is useful for reducing particulates and films (50–52). Carbon dioxide "snow" can be formed by adiabatic cooling of compressed CO_2 gas. The size of the snow particles depends on the geometry of the expansion nozzle and the gas flow and can be varied over a wide range. The CO_2-snow is soft and has been shown to be a good abrasive cleaner that does not damage sensitive surfaces such as mirrors and other optics. The CO_2-snow exhibits some solvating action, which is probably due to a liquid phase between the impinging particle and the surface. If the snow is made from high-purity CO_2, it will leave no residue on the surface. This technique removes particulates, fingerprints, and silicone oil from surfaces and has been found effective as a solvent cleaner for the removal of hydrocarbons in many cases. It does not remove heavier oils very well. A major processing variable is the purity of the compressed CO_2 gas and a purifier should be included in the cleaning system. Commercial CO_2-snow cleaning units are available.

Chemical Etching

Chemical etching can be used to remove surface material along with the contaminants. This is a very useful technique for achieving a "known" surface condition

[53–55]. Chemical etchants can be highly selective in their action. This can result in preferential etching of grain boundaries or certain crystallographic orientations. Etching removes surface layers such as oxides, eliminates or blunts surface cracks in brittle materials, and removes strongly adherent contaminates. Common wet chemical etchants for glass include sodium or ammonium bifluoride (e.g., 100 grams of ammonium bifluoride salt in 800 cc deionized water), trisodium phosphate (a mild etchant), and hydrofluoric acid (a very strong etchant). These materials are dangerous and should be handled with proper precautions to prevent injury. When using etchants for cleaning, care must be taken to prevent selective removal of surface constituents that are important to further processing [56]. Sometimes chemical etching does not remove some constituents from a surface and leaves a "smut" that must be removed by further etching. For example, etching aluminum alloys with sodium hydroxide (NaOH) leaves a copper smut and/or a silicon smut on the surface. The copper smut can be removed by a nitric acid (HNO_3) etch, and a copper/silicon smut can be removed with a HNO_3/HF etch.

Reactive plasma cleaning [57] can be called dry chemical cleaning in that a volatile reaction product is formed and leaves the surface. Typically a chlorine or fluorine-containing gas provides the reactive species. This technique is used for *in situ* cleaning of some vacuum processing systems such as those used for PECVD and plasma etching (Section 4.9.4.2).

Solvent Cleaning

Some contaminants can be removed from surfaces by solvents that dissolve (take into solution) specific contaminants. Polar solvents such as water and water/alcohol mixtures are used to dissolve ionic materials (salts) that are polar contaminates. Nonpolar solvents such as the chlorinated hydrocarbon solvents, are used to remove nonpolar contaminates such as grease. Often there is a mixture of solvents to dissolve both polar and nonpolar contaminates. Solvents can vary greatly as to their ability to dissolve contaminants, and their effectiveness needs to be determined by determining the "solubility parameter" for specific contaminants. The solubility parameter is the maximum (saturation) amount of a specific contaminant that can be dissolved in a specific amount of the solvent. Many nonchlorinated hydrocarbon-based or petroleum-based materials are used as solvents. One of the most common is acetone (CH_3COCH_3). Acetone removes heavy greases quite effectively but tends to leave a residue and it is quite flammable. Acetone cleaning or "wipedown" should be followed by a methanol rinse or wipedown to remove the residue. For instance, a solvent wipedown cleaning sequence is trichloroethylene–acetone–methanol–isopropanol.

Chlorinated hydrocarbon nonpolar solvents such as trichloroethylene are often preferred to hydrocarbon-based or petroleum-based solvents because of their lower flammability (i.e., higher flashpoint, as determined by ASTM D1310-63).

However, because of the toxicity and carcinogenic properties of some of these materials, they should be used in well-ventilated areas or the vapors should be contained in a closed recycling system. Volatile organic compounds (VOCs) are those that have boiling points below 138°C. The discharge of VOCs into the environment is regulated by local, state, and federal laws. To comply with these regulations it may be necessary to recycle the material by condensation of the vapors or to thermally destroy the vapors by burning [58], that is, instead of releasing them into the atmosphere.

Chlorofluorocarbon (CFC) solvents are more stable and less toxic than the chlorinated solvents; however, the well-substantiated atmospheric ozone depletion and the controversial increase in the greenhouse effect have caused their use to be banned. Since January 1, 1996, CFC solvents have not been produced in the United States. Other possible solvents for removal of greases are N-methyl-2-pyrrolidone-based solvents [59,60], terpene-based solvents, as well as the Fluoroinerts®, which do not contain chlorocarbon bonds. Reports indicate that the terpenes may be as effective as the CFCs in many instances, though they have a greater tendency to leave residues. Terpenes suffer from the fact that they have low flashpoints (about 120°F) and reduced lower explosive limits (LELs) than the CFCs. Other approaches use nonlinear alcohols and purely aqueous cleaning [61]. This area of solvent development is rapidly changing.

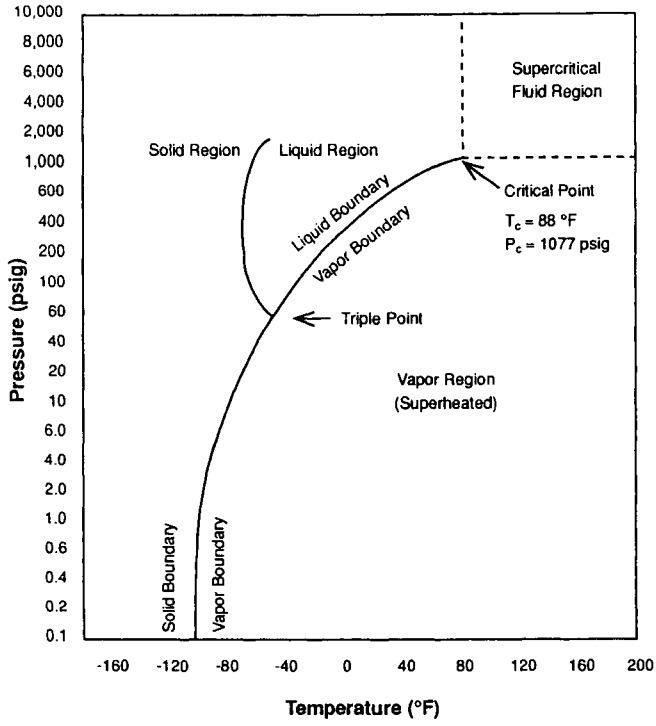
Another type of solvents are the supercritical fluids. If a gas such as CO_2 is compressed to its "critical pressure" (CO_2 = 1077 psi), it liquefies to become a "critical fluid." If it is also heated above its "critical temperature" (CO_2 = 31.1°C), it becomes a supercritical fluid (SCF), as shown in Figure 2. Critical fluids and supercritical fluids are good solvents for many medium-molecular-weight, nonpolar or slightly polar organics. The more dense they are, the better their solvency power.

Solvents can be densified most easily when they are in the supercritical state. Carbon dioxide has been shown to have a Hildebrand solubility parameter that can vary from 0 in the gas to 10 under high-pressure supercritical conditions [62]. Values of 6 to 8 are typical, which is about the same as hexane and carbon tetrachloride. Supercritical CO_2 fluid (SCF-CO_2, critical point 31°C, 74 bar pressure) has the advantage that it is stable, has low toxicity, minimal cost, and is a solvent for many organic materials and has shown promise as a solvent cleaning technique [63]. Table 4 shows typical operating parameters for SCF cleaning.

Saponifiers

Alkaline cleaners (generally silicate- and phosphate-based) are saponifiers that convert organic fats to water-soluble soaps. Alkaline cleaners have a pH of about 11 and are generally used hot [64]. After using alkaline cleaners, the surface should receive an acid dip prior to the water rinse since alkali salts adhere strongly to sur-

Fig. 2.



PHASE DIAGRAM FOR PURE CO_2

Phase diagram for pure CO_2.

faces. Clean oxide surfaces strongly adsorb hydrocarbons, and detergents or solvents normally do not completely remove the hydrocarbons; alkaline or oxidative cleaners must be used to remove the remaining hydrocarbons. Strong alkaline

Table 4

Operating Conditions for CO_2–SCF Cleaning

Parameter	Range
Pressure	1450–4350 psi
Temperature	100–185°F
SC CO_2 density	30–50 lb/ft^3
SC CO_2 flow rate	2–11 lb/hr
Cleaning time	0.5–3 hours

cleaners can etch aluminum and oxide surfaces, particularly glasses, so solution strength (pH), temperature, and exposure times should be carefully controlled.

Detergent Cleaning

Detergent (soap) cleaning is a comparatively mild cleaning technique [65]. In detergent cleaning, the detergent surrounds contaminants, taking them into suspension without actually dissolving the material. This action is assisted by wetting agents and surfactants, which loosen the contaminants from the surface. Liquid dishwasher soap is an excellent detergent for many applications such as cleaning polymer surfaces. A major problem with soaps is that metal ions, such as the calcium and magnesium, which are found in hard water, make the soaps insoluble, thus leaving a residue. Deionized (DI) water should always be used for residue-free detergent cleaning. Many detergents contain phosphates, which can be environmentally harmful and subject to pollution regulations.

Surface Tension

The ability to wet a surface depends on the relative surface tensions of the surface and the fluid. The lower the surface tension, the better the fluid can wet a surface and penetrate into small spaces. The surface tension of clean oxide surfaces is close to 1000 dyne/cm. The surface tension of water can be lowered by the addition of chemicals. Table 5 shows the surface tension of some water solutions:

Table 5

Surface Tension of Fluids

Material		Surface Tension (in air) mJ/m^2 (dyne/cm)
Pure H_2O	at 18°C	= 73.05
	at 50°C	= 67.91
	at 100°C	= 58.9
n-propanol	at 25°C	= 23.32
H_2O + 30 vol% *n*-propanol	at 18°C	= 26.9
Ethyl alcohol	at 30°C	= 21.5
H_2O + 50 vol% ethanol	at 30°C	= 27.5
1000 g H_2O + 34 g NH_4OH	at 18°C	= 57.05
1000 g H_2O + 17.7 g HCl	at 20°C	= 65.75
1000 g H_2O + 14 g NaOH	at 18°C	= 101.05
1000 g H_2O + 6 g NaCl	at 20°C	= 82.55

Surfactants

Surfactants are the generic name for surface-active agents that reduce the interfacial energy of materials in contact. Surfactants used with water have both hydrophobic ("water hating") and hydrophilic ("water loving") groups. They dissolve in water by virtue of their hydrophilic groups and lower the surface energy of water to about 30 mJ/m^2. The surfactant collects at the interface between immiscible substances, such as oil and water, and lowers the interfacial energy. Surfactants should best be used in deionized water.

pH Adjusters and Chelating Agents

In solutions, pH adjusters are used to aid in the cleaning action. Generally it is found that basic solutions clean better than acidic solutions if chemical etching is not involved. The pH of the cleaning solution is often adjusted to basic using ammonia or ammonium hydroxide. Chelating agents keep in solution the normally insoluble phosphates that are formed in hard-water detergent cleaning. Glass-cleaning solutions often use chelating agents such as ethylene diamine tetraacetic acid (EDTA) and citric acid with salts containing hydroxyl and amine substitutes.

Reactive Cleaning

Reactive cleaning uses liquids, gases, vapors, or plasmas to react with a contaminant to form a volatile or soluble reaction product.

Fluids

Reactive cleaning liquids are often oxidizing solutions. Many acid-based systems can be used as oxidants. One system commonly used in the semiconductor industry is the "piranha solution." The piranha solution is hot (>50°C) concentrated sulfuric acid plus ammonium persulfate [66]. The addition of the solid ammonium persulfate to the hot sulfuric acid produces peroxydisulfuric acid, which reacts with water to form H_2SO_5 (Caro's acid) which further decomposes to form free atomic oxygen. The ammonium persulfate should be added just prior to the immersion of the part into the solution. The effectiveness of this oxidation technique can be shown by first placing a piece of paper in the hot sulfuric acid where it is carbonized, then adding the ammonium persulfate and watching the carbon disappear. This treatment is sometimes followed by a brief dip in a 10:1 solution of water and HF or immersion for 20 minutes in a hot solution of hydrogen peroxide and ammonium hydroxide in the ratio H_2O : H_2O_2 (30%) : NH_4OH (29%) at 80°C. Another similar oxidizing solution uses stabilized sulfuric acid and hydrogen peroxide.

Hydrogen peroxide (H_2O_2) is a good oxidizing solution for cleaning glass. Often boiling 30% unstabilized H_2O_2 is used. The most common hydrogen peroxide has been stabilized, which reduces the release of free oxygen. Unstabilized H_2O_2 must be stored in a refrigerator to slow decomposition. Hydrogen peroxide is sometimes used with ammonium hydroxide, to increase the complexing of surface contaminants, and is used at a ratio of

$$8\ (30\%\ H_2O_2) : 1\ (NH_4OH) : 1\ (H_2O)$$

However, the decomposition rate of the H_2O_2 is greatly increased by combination with ammonium hydroxide.

Oxidative cleaning can be performed using chlorine-containing chemicals. For example, a water slurry of sodium dichloroisocyanurate (i.e., swimming pool chlorine), which has 63% available chlorine, can be used to scrub an oxide surface to remove hydrocarbon contamination. This combines mechanical scrubbing with oxidation and improves the cleaning action.

Gases

Reactive gas cleaning uses a reactive gas at high temperature to form a volatile material. For example, air firing of an oxide surface oxidizes all the hydrocarbons, and they are volatilized. High-temperature air fire is an excellent way to clean surfaces that are not degraded by high temperature [67,68]. For example, alumina can be cleaned of hydrocarbons by heating to 1000°C in air. Some care must be taken in furnace firing, because particulate generation from the furnace liner can be a source of undesirable particulates and sodium from the insulating material may be an undesirable contaminant for semiconductor device fabrication. Self-cleaning kitchen ovens clean by oxidation at about 425°C.

The oxidation by ozone (O_3) created by ultraviolet radiation (UV) at atmospheric pressure and low temperature is useful for removing hydrocarbons. The ultraviolet radiation also causes bond scission of the hydrocarbon contaminants, which aids in cleaning. The UV/O_3 cleaning has greatly simplified the production, storage, and maintenance of hydrocarbon-free surfaces [69,70].

The UV is produced by a mercury vapor lamp in a quartz envelope so that both the 1849 Å and the 2537 Å radiation is transmitted. The short-wavelength radiation generates ozone and causes bond scission in the hydrocarbon contaminants. The mercury lamps can be custom made to a variety of shapes for specific applications. Ozone adsorbs UV so the surfaces should be as close as possible to the UV source. UV radiation intensity should be maintained to about 1–10 milliwatts/cm^2 at the surface. In the UV/O_3 chamber the air may be stagnant or flowing. If flowing air is used, the air should be filtered. In a correctly operating system, ozone can be detected by smell when the chamber is opened. The smell is similar to that of the air after a lightning storm and indicates that the ozone con-

centration is less than 10 parts per million by volume (ppmbv). Higher concentrations of ozone deaden the olfactory nerves and are harmful.

Typical exposure times for UV/O_3 cleaning are from a few minutes to remove a few monolayers of hydrocarbon contamination to hours or days or weeks for storage of cleaned surfaces. The UV/O_3 cleaning technique has the advantage that it can be used as a dry in-line cleaning technique [71]. The UV/O_3 cleaning technique is also useful for cleaning holes (vias) in surfaces [72].

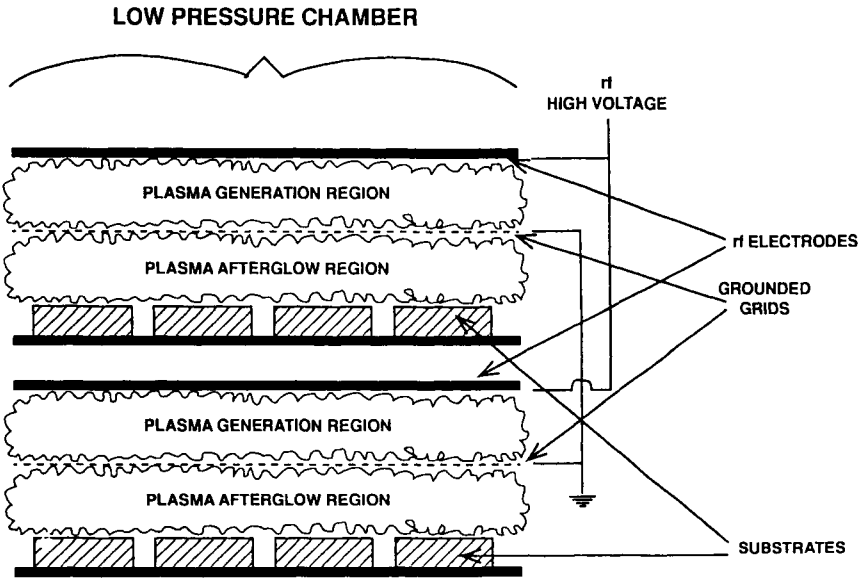
Plasmas

Plasma cleaning can be done in a plasma system separate from the deposition system. Reactive plasma cleaning uses a reactive species in the plasma to react with the contaminants to form a volatile species, which leave the surface at much lower temperatures than those necessary for reactive gas cleaning. The additional requirement on reactive plasma cleaning is that it does not leave a residue. Oxygen (from pure ["medical"] air) and hydrogen (pure or as "forming gas") [73] are the most common materials used for plasma-cleaning vacuum surfaces both as an external cleaning process and as an *in situ* cleaning technique (Section 4.9.4.2) [74].

Oxygen (or air) plasmas are very effective in removing hydrocarbons and absorbed water vapor from surfaces. The reaction of the oxygen with carbon on the surface can be monitored using a mass spectrometer to monitor the CO and CO_2 that is produced [75]. Hydrogen plasmas can be used to remove hydrocarbon contamination when oxygen plasmas are unacceptable. This technique has been used to clean vacuum surfaces (stainless steel) in nuclear fusion reactors [76]. Hydrogen plasmas have been used to clean metals [77] and semiconductor materials [78].

Figure 3 shows a typical plasma cleaner where the plasma is generated by an RF discharge and the surfaces to be cleaned are in a "remote" or "afterglow" location and not in the plasma generation region. Figure 4 shows the processes that occur on a surface exposed to a plasma. The surface attains a potential (sheath potential) that is negative with respect to the plasma, and ions are accelerated from the plasma to the surface. For the case of a "cold plasma," which has low energy particles, this sheath potential will only be a few volts. When the plasma particles are more energetic or the electrons are accelerated to the surface, the sheath potential can be tens of volts. In addition to being bombarded by ions, the surface in contact with the plasma will be bombarded by "activated species," excited species, and other high-energy species. Ions and excited species release their energies of ionization or excitation when they impinge on the surface. For example, when a singly charged argon ion impinges on a surface, it will give up the kinetic energy it attained by acceleration through the sheath potential and the ionization energy, which is 15.7 eV. For a high ion flux, this release of heat can result in an appreciable temperature rise at the surface.

Fig. 3.



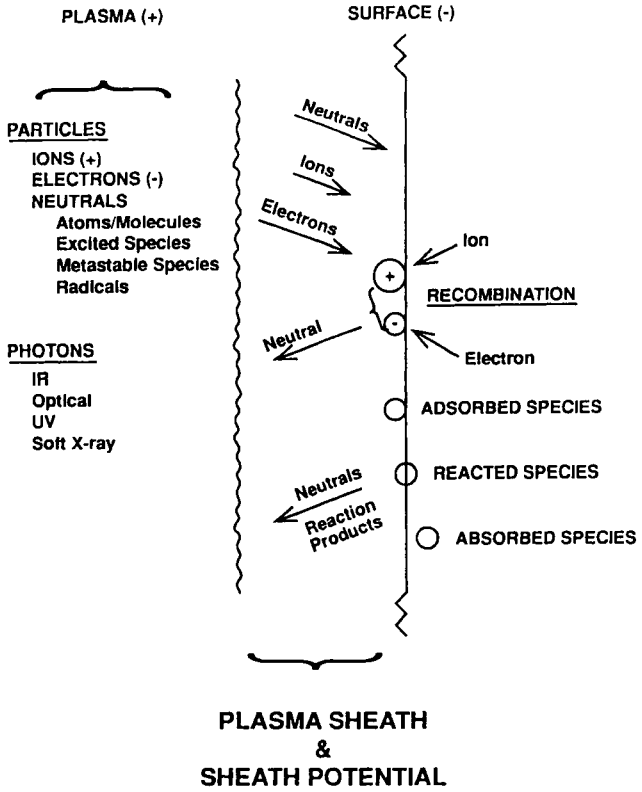
Plasma cleaner using parallel plate electrodes. The surfaces to be cleaned are placed in the "afterglow region" outside of the plasma generation region.

RINSING

After any wet cleaning process, the surface should be thoroughly rinsed in an ultra-pure liquid, usually water, before being allowed to dry. This avoids leaving residues on the surface. The most common rinsing technique is to use successive rinses (cascading rinsing) in ultrapure water until the rinse water retains a high resistivity (> 12 megohm/cm). This is called "rinse to resistivity." After rinsing, the surface should be dried as quickly as possible, because the residual water film on the surface will cause particles to stick to the surface and, on drying, the particles will adhere very tenaciously.

Ultra-pure or deionized (DI) water is used for rinsing since it is easily prepared in large quantities. Water purity is measured using a conductivity cell that measures the ionic concentration in the water. Spontaneous dissociation of the water molecule limits the resistance of water to 18.2 megohms between electrodes spaced 1 centimeter apart (resistivity of 18.2 megohm/cm). Conductivity measurements do not measure the organic or biological contamination and some type

Fig. 4.



Plasma–surface interactions.

of residue analysis must be used to measure these impurities. There are a number of techniques for determining the nature of the contaminants in water [79]. A simple technique is to let some of the water evaporate on a clean glass surface and then look for a "haze" or for loss of contact angle (discussed later).

Typical semiconductor specifications for ultra-pure water for end use are

- Resistivity—18 megohm/cm continuous at 25°C
- Particle count—less than 500 particles (0.5 microns or larger) per liter
- Bacteria count—less than one colony (cultured) per cc
- Organics—less than one part per million

- Total electrolytes — less than 5 parts per billion NaCl equivalent
- Quantity requirements
- Peak-level usage

High volumes of ultrapure water are made by

- Pretreatment — pH adjustment, coagulation, filtration
- Reverse osmosis — semipermeable membrane (pore size of 10^{-3} to 10^{-4} microns), which rejects salts, dissolved solids (90–98%) and organics (99%), and requires 400 to 600 psi feedwater
- Degasification — removes dissolved CO_2
- Ion exchange (anion and cation) — ion exchange resins remove ions by exchanging H^+ for cations and OH^- for anions
- Absorption materials (activated carbon) — remove organics
- Filtration — removes particulates and biological matter (0.2 microns filter pore size, for bacterial, 1.0 microns pore size for prefilter)
- Ultraviolet radiation or ozone bubbling — kills bacteria on filters
- Endpoint filtration — 0.2 micron filter pore size

Smaller amounts of ultra-pure water can be prepared by the same process steps beginning with the ion exchange process. Triply distilled water can also be used but it is relatively expensive. Note that in the production of "soft water" anions and cations are exchanged for Na^+ and Cl^-, and the water will leave a residue.

APPLICATION OF FLUIDS

Fluids are often used in cleaning processes. Fluid baths should be continuously filtered and monitored so as to replace or replenish the active ingredients as they are used or become contaminated. In cases of contaminants that are less dense than the fluid, the surface of the solution can be "skimmed" as contaminants rise to the surface, or the container can be of an "overflow" type where the fluid flows over the lip of the container. There are a number of ways to apply the fluids to the surface to be cleaned.

Soaking

Soaking involves immersing the part in a solution for an extended time, and therefore soaking has not been a desirable technique for production. This may change in the future when less aggressive cleaning chemicals must be used because of environmental concerns. Immersion of a surface in a stagnant solution is generally a poor technique, because the contaminants that are taken into solution are concentrated near the surface and must diffuse away.

Mechanical Disturbance

Mechanical disturbance uses agitation, wiping, brushing, or scrubbing in a fluid environment to break up the stagnant fluid layer near the surface, loosen particles, and aid in carrying contamination away from the surface. Care must be taken to ensure that any material that is used in a fluid does not produce particulates and is compatible with the fluid and surfaces it contacts. When using any mechanical rubbing, care should be taken to prevent contamination by abrasive transfer from the rubbing media—gentle pressure should be used. There are a variety of brush materials used in fluids including polypropylene, Teflon®, and Nylon®. If wiping or scrubbing with a cloth is used, care should be taken that the cloth is lint-free and desized by multiple washing before use. Special particulate-free sponge, and cloth materials are available for wiping. In semiconductor technology, mechanical scrubbing combined with high-pressure fluid jets (2000–3000 psi) and spinning are standard cleaning procedures.

Spraying

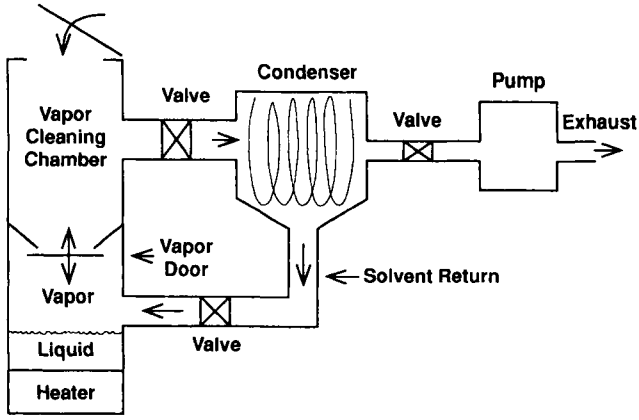
Liquid sprays can be used to remove contaminants from surfaces. Spray pressures can be low, at less than a hundred psi, or high, at several thousand psi. Spraying parameters include the type of fluid, pressure, angle of incidence, and volume of fluid. Sprays should be directed at an oblique angle to the surface. Spray systems often use copious amounts of material so the fluid should be recycled. The spraying fluid can also contain suspended particles that aid in cleaning by adding a mild abrasion to the cleaning. For example, CeO polishing compound is often added to fluids in cleaning glass for mirrors [49].

The fluid should be monitored by residue analysis, and when it is contaminated above a given level it should be replaced. With increasing concern about solvent vapors, many of the newer solvent spray systems are self-contained with condensers to trap the solvent vapors as used with vapor cleaning. Some systems allow the purification of the solvents by distillation. It should be noted that spraying can induce resonant vibrations that can cause component failure or deterioration.

Vapor Cleaning

In vapor cleaning, a cold surface is held in a hot vapor of the solvent. The solvent condenses on the surface, dissolves the contaminants and flows off into a hot "sump" [80,81]. When the part reaches the temperature of the vapor, condensation ceases and the cleaning action stops. The vaporization of the solvent in the sump provides continuous distillation of the cleaning fluid. Parts should never be immersed in the sump fluid. Fluid in the sump should be changed when it

Fig. 5.



ENCLOSED VAPOR CLEANING SYSTEM

Enclosed vapor cleaner and dryer.

becomes contaminated. The old vapor cleaners ("vapor degreasers"), which were open to the air, were very effective but with the advent of the concern over ozone depletion and VOCs (volatile organic compounds) emission, they are not feasible. Instead, enclosed vapor cleaners that trap and recycle the vapors are being used [82]. Figure 5 shows a schematic of a typical enclosed vapor cleaner. This type of equipment is also amenable to vapor drying and spray cleaners. In many cases, older vapor degreasers are being retrofitted into immersion cleaners.

Mixture of cleaning fluids can be used for vapor cleaning. A mixture of fluids that has the same composition in the vapor as in the fluid is called an *azeotrope*.

Ultrasonics

Ultrasonic cleaning is widely used to aid in cleaning surfaces [83–85]. Low-frequency ultrasonic cleaning relies on the jetting action of collapsing cavitation bubble in contact with a surface to provide a high-pressure jet of fluid against the surface. Ultrasonic cleaning is often a good way to remove loosely adhering particles after a grinding or abrasive procedure and can be used with solvents to remove adsorbed contaminants.

The cavitation bubbles are formed by the tension wave portion of an ultrasonic wave in a fluid media. The ultrasonic wave is produced by magnetostrictive or electrostrictive transducers(s), typically operating at 18–120 kHz, and at an energy density of about 100 watts/gal of fluid. The ultrasonic cleaner size can be from 5 gallons for a small cleaner, up to very large systems using many transduc-

ers. The size of cavitation bubbles in the fluid depends on the vapor pressure, surface energy, and temperature of the fluid. For example, water at 60°C and 40 kHz has a cavitation bubble size of about 3 microns. The jet pressure from the collapsing bubble can be as high as 300 psi. The cavitation jetting is more energetic for cooler media and when there are no gases in the bubble to hinder its collapse. The ultrasonic energy density decreases with distance from the transducer; therefore the cavitation energy is greatest near the transducer surface. Acoustic streaming results in an overall movement of fluid away from the transducer surface (bottom of the tank). This brings contaminants that have settled to the bottom of the tank up into the cleaning region. Therefore the cavitating fluid should be continuously filtered.

When using a fixed-frequency transducer, there are nodes and antinodes formed (standing waves) in the fluid, which produce variations of cavitation energy with position. These standing wave patterns can be modified by reflection of the pressure waves from surfaces in the tank. This variation in cavitation with position can be overcome somewhat using swept-frequency generation. A typical system uses 40 kHz ± 2 kHz. If frequency sweeping is not used or there are large variations of cavitation energy with position, the parts should be moved from one region to another in the tank during cleaning. The ultrasonic frequencies are above the hearing range of the human ear, and the audible noise that is heard from an ultrasonic cleaner is due to vibration of surfaces in the cleaner.

Variables in ultrasonic cleaning include [86,87]:

- Amplitude and frequency of pressure wave (energy density, standing wave pattern)
- Nature of the transducer fluid (density, viscosity, surface tension, vapor pressure)
- Nature of the cleaning fluid if different from the transducer media
- Surfaces in the transducer media that must transmit the pressure waves
- Flow and filtering of the cleaner fluid
- Temperature of fluid
- Fluid contaminants such as water
- Gas content of the fluid
- Energy of cavitation implosion (temperature, pulse height of ultrasonic wave)
- Cavitation density changes with position in tank
- Cavitation density changes with time
- Shape of the pressure pulse
- Nature of ultrasonic cycle train ("quiet time," "degas time," cycles per train)
- Geometry of the system and associated fixtures

Quite often ultrasonic cleaning is not as effective as one would like. Ultrasonic erosion or deformation of aluminum foil or removal of an aluminum metallization from a glass surface can be used to determine the cavitation power that a sur-

face is exposed to in the ultrasonic cleaner. The cavitation intensity can be studied by observing the cavitation damage on a series of aluminum foils with increasing thickness. The damage changes from hole generation, to dimpling, to pitting, to no damage, with foil thickness. The cavitation intensity of an ultrasonic cleaner should be plotted as a function of position with fixtures and parts in position, since reflections from surfaces can change the cavitation energy distribution. The cavitation pattern should be checked periodically.

Fixturing is very important in ultrasonic cleaning to ensure that all surfaces are cleaned. Generally the total area of parts, in cm^2, should not exceed the volume of the tank, in cm^3. Parts should be separated and suspended with the surface to be cleaned parallel to the stress wave propagation direction. Surfaces that are oriented parallel to the stress wave will not be cleaned except by action of the solvent. The parts must not trap gases that prevent wetting of the surface by the cavitating fluid. Metal or glass holding fixtures of small mass and an open structure should be used. Energy adsorbing materials such as polyethylene or fluoropolymers should not be used in fixturing, since they adsorb the ultrasonic energy.

A common mistake in ultrasonic cleaning is to place the parts to be cleaned in the bottom of a container containing the cleaning fluid and then to immerse the container in the ultrasonic cleaner. The parts will just bounce up and down and no cavitation bubbles will form on the surface of the parts.

Ultrasonic cleaning must be used with care since the jetting action can produce high pressures that cause erosion and introduce fractures in the surface of brittle materials. For example, in high-power laser applications it has been shown that extended ultrasonic cleaning of glass surfaces increases the light scattering from the surfaces, indicating surface damage. Ultrasonic agitation has been shown to create particles by erosion of the container surface. The erosion of stainless steel creates 500 times as many particles as the erosion of Pyrex® glass containers. In all cases studied, particles of the container material were produced on prolonged use. Resonance effects may also mechanically damage devices in an ultrasonic cleaner. Ultrasonic cavitation can also be a source of pitting and the loss of adhesion of thin films [88]. Surface damage can be controlled by adjusting the energy density of the cavitation and/or controlling the time of application.

DRYING

Drying is the vaporization or displacement of water or other fluid adsorbed on the surface or absorbed in the bulk. Porous and rough surfaces retain fluids more readily than do smooth surfaces and are more difficult to dry since the fluids are trapped in capillaries. Oxide layers on metals are often porous and retain fluid molecules readily. Drying, by removal or displacing of the water, has the advantage that when the water is removed it takes the bulk of the potential residues with

it, whereas vaporization of large amounts of fluid concentrates the residues, giving a "water spot" of residue. After fluid cleaning and rinsing, it is important to dry the surface quickly to prevent the water film from collecting particles.

Anhydrous alcohol, such as isopropyl alcohol (IPA), anhydrous ethyl alcohol denatured with acetone or methanol, or a commercial drying agent such as a high-vapor-pressure Freon® are good drying agents. They displace the water, and when the surface is removed from the fluid the surface dries rapidly. Drying fluids should be residue free and should be discarded or recycled as they take up water, either from the drying process or from the ambient. The water content of the drying fluid can be monitored by its specific gravity or by monitoring the infrared (IR) adsorption peak for water.

One of the best drying techniques is a "vapor dry" where the cold surface is immersed in the vapor above a heated anhydrous alcohol sump [89–91]. The cold surface condenses the alcohol vapor, which flows off into the sump, taking water and particulates with it. When the surface becomes hot, condensation ceases, and the hot surface, when withdrawn, dries rapidly. Such a system is shown in Figure 5.

Fluids can be blown from the surface using a low- or high-velocity gas stream. When blowing, a nozzle with a 0.2 micron or smaller particulate filter should be used in the nozzle. In addition, when drying insulator surfaces, the gas should be ionized to prevent charge buildup on the surface.

A high-velocity jet of gas can be shaped to blow off a moving surface. The jet is often shaped into a long, thin configuration, and this "air knife" is used to remove fluid from a moving surface such as a large glass plate. Exiting the air knife, the gas velocity can be as high as 35,000 fpm (feet per minute). The jet should hit the oncoming wet surface at about a 30-degree angle. At the trailing edge, a droplet will form and spread back over the surface when the jet is past. Thus, the water on the surface should be ultra-pure so as not to leave a residue when it drys. The size of the droplets can be reduced by decreasing the surface tension of the water by the addition of alcohol. This technique leaves a thin layer of water on the surface so subsequent evaporative drying may be necessary.

Water can be removed from a surface by thermal evaporation. The most common means is to use vacuum baking or gas drying in a hot oven. The vacuum environment aids in evaporation, but a circulating gas dryer is most often used. This technique suffers from the fact that the residue is concentrated as the water dries, leaving a "water spot" unless the water is very pure.

MONITORING CLEANING

The best monitoring technique for cleaning is the ability of the process to provide surfaces that can be processed or used in an acceptable manner. The testing of a surface invariably results in contamination of the surface, so tested surfaces gen-

erally cannot be used for subsequent processing. In some cases, witness sample surfaces can be tested for cleaning to determine surface conditions.

The cleanliness of smooth surfaces can be determined during the rinse operation by observing the wetting and "sheeting" of water on the surface. Sheeting is the flow of the water over the surface as it drains, giving a smooth water surface. If there is hydrophobic contamination on the surface, the water will avoid that area and the sheet of water will "break up." This test is often called the "water break" test.

A common check on the cleanliness of a glass surface uses the contact angle of a liquid drop on the surface of the cleaned glass [92]. If the surface is clean, it has a high surface energy, and the liquid wets and spreads over the surface. In the case of water on a clean glass surface, the contact angle is less than 5 degrees as measured with a contact angle goniometer. This technique must be used with some care since if a hydrophilic contaminant, such as a soap residue, is present, the contact angle may be low even though the surface is contaminated.

A smooth, clean surface will give uniform nucleation of a vapor on the surface. A common test is to breathe on the surface and look at the nucleation pattern of the moisture from the breath. This is called the "black breath test." For example, nucleation of water on the mirror in a shower room will show up the "swipes" where the mirror surface has not been cleaned very well.

Ionic salts are soluble in water and water–50% alcohol mixtures. The removal of ionic salts can be monitored by measuring the resistivity of the rinsing solution. A surface that is relatively free of ionic contaminants should exhibit a rinse resistivity of greater than 10 megohm-cm.

SAFETY

Appropriate laboratory safety methods and procedures should be used at all times in the cleaning process [93,94]. Various industrial organizations have formulated guidelines for the safe use of industrial chemicals. For instance the Institute for Interconnecting and Packaging Electronic Circuits has issued a guideline entitled "Guidelines for Chemical Handling Safety in Printed Board Manufacture" (IPC-CS-70). In the United States the federal government (OSHA) establishes exposure limits for various toxic and carcinogenic chemicals. These limits should be strictly adhered to in the workplace. Reference should be made to current OSHA guidelines since they change frequently. Chemical manufacturers and distributors are required to provide Material Safety Data Sheets (MSDSs) for the materials with shipment of all materials. OSHA has mandated that employees must be provided with this information and trained with respect to the hazards of the materials that they are using (Hazard Communication Standard 29 CFR 1910.1200) [95].

4.9.2.4 Parts to Be Processed

Parts to be processed are cleaned in much the same way as removable surfaces, although the process is often more highly automated, using "cleaning lines" to advantage. Problems with recontamination are lessened if the cleaning procedure is integrated into the processing line, the prepared surfaces are used immediately, and/or if there is a final *in situ* cleaning process in the vacuum system.

4.9.3 ASSEMBLY, HANDLING, AND STORAGE

An important aspect of vacuum engineering is the preparation of the surfaces that go into fabricating and assembling the vacuum system. If properly prepared and handled during fabrication and assembly, the system will have better vacuum performance than if they are improperly prepared, cleaned, or handled during assembly. Nonremovable surfaces should be scrupulously cleaned and handled during assembly. Ionic salts such as sodium chloride (NaCl) from sweat should be particularly avoided. These salts are deliquescent and will take up moisture when exposed to the environment. Cleaning should be performed as described in Section 4.9.2.3.

4.9.3.1 Assembly

Parts and subassemblies should be cleaned as early in the assembly stage as possible. They should then be handled and stored in a noncontaminating manner [96]. Cleaning can be simplified by having the supplier of parts and components clean and package them in a noncontaminating manner. Commercial cleaning and packaging services are available.

4.9.3.2 Processing Environment

Clean surfaces are very reactive and easily recontaminated. Recontamination can occur from the adsorption of vapors, collection of particles, contact with other surfaces, or reaction with reactive gases. An important aspect of cleaning are the conditions existing in the processing area, the handling of the surfaces and storage of the cleaned part. Dust is a particular concern in physical vapor deposition (PVD) processing, because particulates on the part surface will result in pinholes in a de-

posited film. Cleaning should be done in an environment and with procedures that are compatible with the level of cleanliness desired. Reduction of contamination in the environment can range from very simple to very elaborate and costly.

Several types of contaminants should be controlled in the cleaning environment, including

- Particulates — organic and inorganic
- Vapors — organic and inorganic
- Reactive gases — unavoidable and avoidable

Several steps should be taken to reduce the contamination in the processing environment. Avoidable contaminants in the processing area should be excluded by good "housekeeping," separation of contaminant-producing processing from the cleaning area; elimination of particulate-producing materials such as excessive amounts of common paper, short-fiber materials such as rugs, cotton fiber clothes and furniture coverings; and the elimination of vapor-producing materials such as many molded plastics and vinyl coverings [24].

A further action that can be taken to reduce contamination is to contain contaminant-producing sources as much as feasible. Humans and their clothing shed large amounts of particulates that are "pumped out" through the loose weave of the clothing as the person moves about. The use of head coverings, mouth coverings, facial hair coverings and coats or coveralls of tightly woven long-fiber cloth or a coated paper such as Tyvek®, will contain the particulates. Body covering comes with several degrees of contamination control capabilities [97]. Button-up "lab coats" provide the least contamination control. Zip-up smocks are better. Supplies can be found in the Buyers Guide of *Micro* (formally *Microcontamination*) *Magazine* [98]. Necessary vapor-producing and aerosol-producing processes should be performed in ventilated work areas such as "chemical hoods."

Vapors, such as water vapor, solvents, and many low-molecular-weight organics and polymers, are easily condensed and adsorbed. Water vapor, in particular, can often be an important and variable contaminant in a vacuum system. The relative humidity of the processing environment should be controlled. Usually a value around 45% is comfortable and not excessive. Organic and polymer vapors often originate from construction materials and furniture covering such as vinyl.

Reactive gases react with clean surfaces to form surface layers. Reactive gases can be unavoidable, such as oxygen, or avoidable, such as chlorine. Chlorine, for instance, reacts with aluminum to corrode the surface. Undesirable reactive gases should be excluded from the work area when possible.

PARTICLE FILTRATION

Airborne particulates are the contaminants most often monitored and controlled in the cleaning environment. Particulates range in size from submicron to many

microns and are measured by light scattering techniques and specified as to the number of particles per cubic foot of air above a certain size (e.g., 0.5 microns). The smaller and lighter the particle, the less its settling rate and the more easily it can be stirred up to become airborne. The smallest particles are generally formed by evaporation of aerosols such as sea spray and bodily fluids from talking and sneezing. Such particles are often inorganic salts. Common organic particles come from broken fabric fibers (wool or cotton) and skin sloughing.

Particulates are removed from the cleaning environment by mechanical filtering [99–101]. Small processing areas with filtered air supplies can be used to provide local clean areas [100]. These can be "clean benches," which use a horizontal laminar flow of filtered air to remove particulates or "soft-wall" clean areas that are separated from the processing area by plastic curtains and use downward laminar flow of filtered air that enters at the ceiling and exits underneath the curtain. The ultimate clean area is the "cleanroom." Cleanroom standards are measured by the number of particles per cubic foot of air with particle size greater than 0.5 microns. Class 1 cleanrooms are attainable with very stringent design, fabrication, and operating conditions. Class 100 cleanrooms are becoming common in fabrication and assembly of particulate-sensitive components and devices. Class 10,000 rooms are usually attainable with modification of existing facilities and moderate control of conditions in the room and are used as general cleaning and assembly areas. Uncontrolled areas can be Class 100,000 or greater. The term "cleanroom" is somewhat of a misnomer in that, in a cleanroom, no attempt is made to filter vapors from the air. Efficient cleanrooms are very expensive to build, operate, and maintain. Careful consideration should be given to other techniques for contamination control before deciding that a high-class (really a low Class) cleanroom is needed.

4.9.3.3 Handling

An integral, and often neglected, aspect of cleaning is that of handling and storage before the next processing step or usage. Handling and storage after cleaning is a major source of recontamination. If possible, parts should be held in fixtures made of materials that can be easily cleaned and do not contain potentially contaminating materials such as the plasticizers on molded polymers. Glass, ceramics, hard metals, or unplasticized polymers are used as handling fixturing materials. It is best if the fixtures do not touch areas of concern, since "abrasive transfer" of clean materials in contact can result in contamination. The use of cleaning fixtures reduces the amount of touching of the critical surfaces by the operator during the cleaning process.

It is preferable to handle surfaces using fixtures or tools; however, in many cases the surfaces must be handled directly, so gloves must be used. Gloves may

be of a woven fabric or of a polymer film that is either molded to shape or heat welded from flat sheet. Polymer gloves for general use are often powdered to make donning the gloves easier, but for cleaning applications unpowdered gloves must be specified, to avoid particulate contamination. Lengths can vary from finger cots, which only cover the finger tips, to wrist length, to elbow length. Woven polyester glove liners that absorb moisture are available, and make the wearing of polymer film gloves more comfortable.

There are a number of choices for polymer glove material, including latex rubber, nitrile rubber, vinyl, polyethylene, and fluorocarbon materials such as Teflon® as well as polymer blends such as latex-nitrile-neoprene-natural rubber blends for use with acids. All glove material should have low extractables for the chemicals that they might contact [102]. Vinyl gloves are comfortable and are often used in handling surfaces. A problem with the vinyl is that when it contacts alcohol (a common wipedown material and drying agent), the alcohol extracts phthalate plasticizers from the vinyl. These extracted materials on the glove surface can then contaminate surfaces touched by the gloves. Tools with plastic handles should be avoided for the same reason. Generally it is best not to have vinyl gloves or plastic-handled tools in the cleaning area.

Unplasticized polyethylene gloves are compatible with alcohol and most cleaning chemicals, and are good gloves for clean handling. An advantage of the polyethylene gloves is that they are rather awkward and uncomfortable and are readily discarded when they are not required. Latex rubber gloves are often used in "suiting up" for the clean room. A problem is that they are then used all day long, thereby transferring contamination from one place to another. When handling clean surfaces, an unplasticized polyethylene glove should be put on over the latex glove and then discarded when the handling is over. A disadvantage of the polymer gloves is that the soft polymer can be easily transferred to a clean surface by "abrasive transfer."

Polymer gloves are slippery, and it may be desirable to use fabric gloves such as desized and lint-free Nylon® or Dacron® woven fabric gloves when friction in handling is desirable or abrasive transfer from softer polymer gloves is a problem. Woven fabrics will wick oils from the skin to the glove surface, so polyethylene or latex gloves or finger cots should be used under the fabric gloves.

4.9.3.4 Storage and Packaging

Cleaned surfaces should be stored in a nonrecontaminating environment. Often surfaces to be stored are held in clean fixtures to reduce the necessity for handling the surfaces directly.

PASSIVE STORAGE

Passive storage environments are those which have been carefully cleaned. A commonly used passive environment is a clean glass container such as a petri dish. Clean surfaces can be stored by wrapping them in a clean material. Wrapping the surfaces in clean nylon fabric or clean aluminum foil works well. Usually it is best to avoid wrapping in paper or polymer wrap because of contaminants in those materials and abrasive transfer of materials.

ACTIVE STORAGE

Active storage environments are those where the contaminants are continually removed from the storage environment by adsorption or reaction. Hydrocarbon adsorption can be on freshly oxidized aluminum [103] or activated carbon. Hydrocarbon contaminants can be continually removed by having an oxidizing atmosphere. A UV/O_3 cleaning cabinet provides such an environment (section 4.9.2.3). The UV/O_3 cleaning chamber is by far the best technique for storing surfaces where surface oxidation is not a problem.

For some applications, moisture is the main contaminant to be considered and an actively desiccated environment is desired. Common desiccants include silica gel, phosphorous pentoxide (P_2O_5), and magnesium perchlorate ($Mg[ClO_4]_2$). Phosphorous pentoxide is probably the most effective desiccant material. It should be fused to reduce particle formation. Desiccants must be used with care, since they tend to be friable and produce particulates. It is best to isolate the desiccants from the storage chamber by means of a particle filter. After prolonged use, desiccants must be regenerated by heating.

4.9.4 *IN SITU* CLEANING

External cleaning should be performed to the best possible extent. However, recontamination can and will occur before the surfaces can be placed into a clean vacuum environment and contamination can occur in the vacuum system due to system-related contamination and/or process-related contamination. *In situ* cleaning is performed in the vacuum system as part of the processing cycle. Vacuum processes requiring long times, such as outgassing, should be done in a separate system to minimized "tieing-up" the vacuum processing system.

4.9.4.1 Origin of In-Chamber Contamination

SYSTEM-RELATED CONTAMINATION

System-related contamination is contamination whose origin is the vacuum system itself. System-related contamination can originate from

- Residual atmospheric gases and vapors, such as O_2, N_2, water vapor
- Desorption of vapors from vacuum surfaces, such as water vapor, vapors from previous processing
- Backstreaming from pumping system, such as pump oils, lubricating oils
- Wall creep from the pumping system
- Outgassing of gases and vapors from construction materials, such as water vapor from organics, hydrogen from metals
- Vaporization of construction materials or solid contaminants
- Leakage of gases and vapors from real leaks
- Leakage of gases and vapors from virtual leaks
- Permeation of gases and vapors through construction materials such as rubber O-rings
- Particles generated by wear between moving surfaces
- Particles from previous processing

Many of these sources have been addressed in other sections of this book.

Minimization of system-related contamination can also be attained by having the vacuum surfaces heated when the system is open to the ambient in order to minimize water vapor adsorption. This can be done by having heater coils, or flood panels on the exterior surfaces of the vacuum system and flowing hot water through them when the system is open. Operational procedures such as having the system opened to the ambient for as short a period of time as possible and venting using dry gas also minimize contamination in the system. In load-lock system design, the processing chamber is kept under vacuum and materials are introduced through a separate chamber where they can be preprocessed to minimize contaminants brought into the chamber. This procedure aids in reducing contamination and in allowing a process to be reproducible.

Oil-sealed and oil-lubricated mechanical pumps are commonly used to reduce the gas pressure in a vacuum chamber to the range of 100 mtorr. An important factor in using these pumps is to minimize the "backstreaming" and "wall creep" of the mechanical pump oils into the vacuum chamber and high-vacuum pump. If oil gets into the processing chamber, it can contaminate the surfaces before processing begins or be decomposed in a plasma to deposits contaminants such as carbon. If the oil gets into a cryopump, it will fill the pores of the adsorbing media and decrease the pumping speed and capacity. If the mechanical pump oil

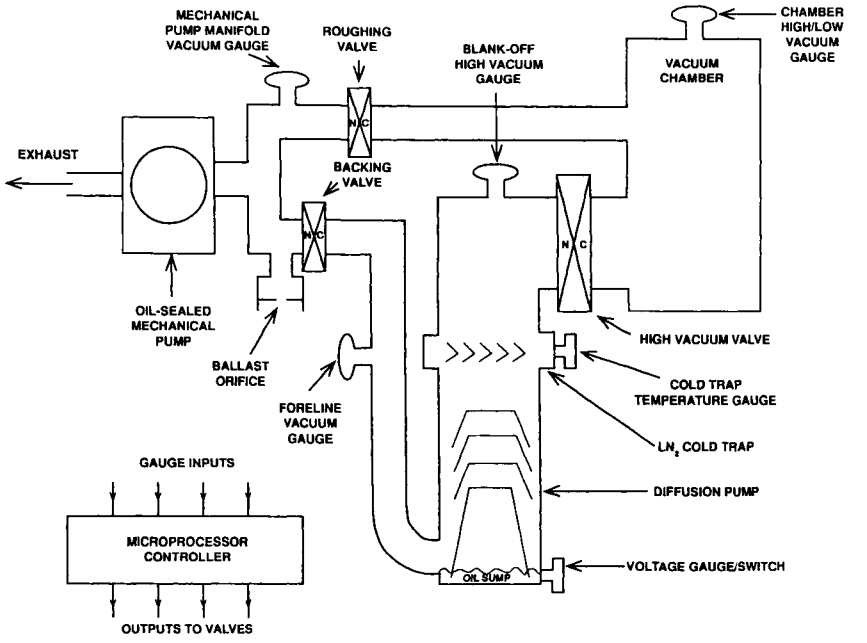
gets into an oil diffusion pump, the high-vapor-pressure mechanical pump oil will quickly make its way into the deposition chamber.

To minimize backstreaming and wall creep, keep the manifold to the mechanical pump in the turbulent (viscous) gas flow pressure regime as much as possible. If at any time the flow should fall below this value, oil will make its way up the pumping manifold, depositing on the walls and valves, and if the valves are open, it will enter the vacuum chamber and high-vacuum pump. There is viscous flow in a pipe of diameter D, when P (torr) $\times$ D (inches) > 0.18. For example, for a 2-inch diameter pipe, the gas flow is viscous above 90 mtorr pressure. The "crossover pressure" from roughing to high-vacuum pumping should be close to this pressure to minimize backstreaming through the roughing line.

Another source of backstreaming is when there is a power failure and the mechanical pump stops. The oil seal in the mechanical pump is not effective in holding a large pressure differential and air will "suck" back through the pump, carrying oil with it into the pumping manifold. This backstreaming can be eliminated by having a ballast orifice or ballast valve on the roughing pump manifold to give a continuous gas flow through the mechanical pump even when the roughing and foreline valves are closed so as to keep the manifold pressure in the viscous flow range. In the event of a power failure, this leak brings the pumping manifold up to ambient pressure, thereby preventing air (and oil) from being sucked back through the mechanical pump. This permanent leak in the roughing manifold adds a pumping load to the mechanical pump, which must be allowed for in the system design. If such a permanent leak is not used, a normally open (NO) (when power is off) "leak valve" or "ballast valve," which opens when there is a power failure, can be used in the manifold between the mechanical pump and the roughing valve. However, this arrangement lets the pumping manifold have a low pressure during normal operation, which allows oil to migrate through the pumping manifold.

The roughing, backing, and high-vacuum valves should be pneumatic or solenoid-operated normally closed (NC) (when power is off), which will close on power failure and not reopen until the proper signal is sent from the microprocessor. The roughing valve and backing valve are activated from a preset vacuum signal to prevent lowering the manifold pressure below the viscous flow range. It is also advisable to have the microprocessor programmed so that the roughing valve will not open if the pumping manifold is at a much higher pressure than the high-vacuum side of the valve. For example, if there is a short power outage, the manifold will be brought to ambient pressure through the permanent leak or the actuated leak valve, but the diffusion pump and/or the vacuum chamber can still be under a good vacuum. If power returns and the roughing valve or backing valve opens, then the gas flow will be reversed and gas will flow from the mechanical pump manifold toward the high-vacuum pump or chamber.

Fig. 6.



FAIL-SAFE DESIGN FOR DIFFUSION PUMPED SYSTEM

Vacuum manifolding for a fail-safe design of a diffusion-pumpepd system in case of a power outage.

Figure 6 shows ways that the vacuum manifolding can be designed to minimize oil contamination from the mechanical pumping system in a diffusion-pumped system. In the diffusion-pumped system, the diffusion pump can be interlocked so as to not heat up until the liquid nitrogen (LN2) cold trap has been cooled. Also shown in the figure is a high-vacuum gauge between the high-vacuum pump and the high-vacuum valve. This gauge allows monitoring the status of the pumping system in a "blanked off" mode. A major change in the pump performance in the blanked-off mode indicates a problem in the pumping system, such as oil contamination of a cryopump, a low oil level in the oil-sealed mechanical pump, a low oil level in the diffusion pump, an incorrect oil sump temperature in the diffusion pump, and so on.

Oil contamination from the pumping system can be completely eliminated by using "dry pumps" that contain no oil. Such a system is a turbomolecular pump with a molecular drag stage backed by a diaphragm pump. This type of system is used in the load locks in semiconductor manufacturing.

If a vacuum system has been contaminated with oil, it is is often possible to

clean the system with an oxygen plasma. The oxygen plasma will volatilize hydrocarbon oils and will oxidize silicone oils to silica, which will then present a particulate cleaning problem.

When a volume of gas is expanded rapidly, the gas cools. If the gas has a high relative humidity, the gas can be cooled past its dewpoint and become supersaturated. The supersaturated water vapor will condense on particles and ions in the gas to form water droplets. This droplet formation is similar to the formation of contrails behind aircraft and has been used to detect ionization tracks in cloud track chambers. The droplets can coalesce to form drops that "rain" down on surfaces in a vacuum chamber [104,105]. When the drops evaporate, they leave a residue that can affect film adhesion and cause pinholes in the deposited film. Laser particle counters are available for detecting particles in the vacuum system during pumpdown. To prevent this type of particle formation mechanism, the vacuum system should be filled with low-humidity air, flushed with dry air before pumping, or evacuated slowly.

In processes where the parts have appreciable outgassing of vapors such as water vapor, solvents, or low-molecular-weight organics, it may be desirable to have large-area cold-panels in the processing chamber. These traps should be cooled to the appropriate temperature to reduce the vapor pressure of the contaminant to the desired level. For example, to trap water vapor, the surfaces should be cooled to $-150°C$. These panels can be "defrosted" or removed periodically in order to minimize contaminant carryover from previous processing.

Structures, such as fixtures, that are put into and taken out of the deposition chamber often represent the largest surface areas in the vacuum system. In many applications, particularly in sputter deposition and in transfer chambers, the fixture may "crowd" the volume to such an extent that it interferes with the pumping action, particularly of vapors. Often this problem can be overcome by using cold panels in the chamber.

Fixtures and tooling often have moving surfaces in contact as the fixtures are moved to randomize positions or move from one point to another. This moving contact can generate particulates in the system. Wear can be minimized by using smooth, hard surfaces in contact and vacuum-compatible lubricants such as liquid lubricants, solid film lubricants (e.g., sputter-deposited or burnished MoS_2) [42] or low-shear metals (e.g., Pb, Sn, or Ag).

In PVD processing, buildup of excess film material on the vacuum surfaces can cause an increased surface area that exacerbates the water vapor adsorption/desorption problem. This film buildup also can generate particulate contamination in the system by flaking from the surface. This problem is compounded when depositing highly stressed films. The system should be designed so that flakes do not fall on part surfaces before deposition and fall to areas where they can easily be removed. This means that systems subject to such problems should not be base-pumped, because particulates will fall into the pumping system.

Turbulence has been shown to "stir up" particles in the vacuum system and "soft pump" and "soft vent" valves have been developed to limit the flow of high-velocity gas into the processing chamber. Rapid venting has been shown to be a greater problem than rapid roughing. If particulates are a problem, the system should be "vacuumed" with a vacuum cleaner on a regular schedule.

PROCESS-RELATED CONTAMINATION

Process-related contamination such as gases, vapors, and particulates in the vacuum processing chamber can originate from

- Particulates brought in with the parts, fixture, tooling, gases, or deposition sources
- Particulates generated during processing (e.g., from sputtering targets, spits from evaporation sources, wear particles from moving fixturing)
- Surface films "brought in" on surfaces
- Desorption, outgassing, and vaporization from introduced fixtures, tooling, parts, deposition source materials, etc.

Some types of vacuum processing produce very fine particles. Sputter deposition can produce "ultra-fine particles" [106] and plasma etching [107] can produce copious amounts of particulate matter in the system. These particles should be swept away through the pumping system.

Contamination from desorption should be minimized by cleaning and outgassing of the materials before insertion into the vacuum system as described in Section 4.9.1.1. Where there is still appreciable contamination brought into the vacuum system, such as high outgassing materials, the system and processing procedures must be designed to accommodate the contamination.

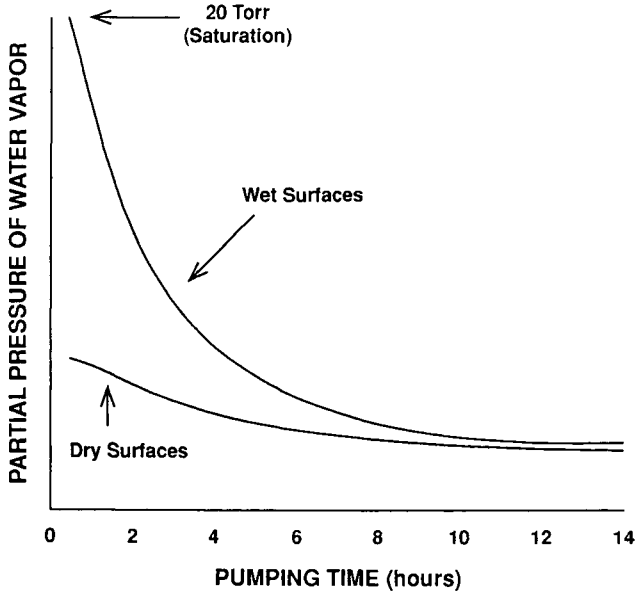
4.9.4.2 Conditioning of Vacuum Surfaces

One type of *in situ* cleaning is sometimes termed "conditioning" and is used to remove the recontamination on surfaces that has occurred after *ex situ* cleaning or after the system has been opened to the ambient. These contaminant species are predominantly water vapor and hydrocarbon vapors. Vacuum system surface conditioning can consist of one or a combination of procedures as follows:

DESORPTION

Desorption is the nonchemical removal of a species from the surface. In thermal desorption, the process is due to heating. When a system is evacuated, the water

Fig. 7.



WATER VAPOR PUMPING CURVES

Starting with "wet" surfaces and "dry" surfaces

Removal of water vapor from a vacuum system that initially had wet surfaces and from one that had dry surfaces.

vapor begins to desorb from the vacuum surfaces. The rate of desorption depends on the surface area, surface condition, temperature, and amount of water on the surfaces. Water vapor is the most common vapor contamination in a vacuum system. Figure 7 shows the removal of water from a vacuum system that initially had a wet surface and one that initially had a dry surface. Note that it takes a very long time to remove most of the water vapor from a system, particularly if it starts in a wet condition.

One technique for increasing the thermal desorption rate is to flush the system with a hot, dry gas [108]. Another is to heat the system ("bakeout"), preferably to >400°C [109–111]. Bakeout for ultra-high-vacuum applications can take days or weeks. When using the bakeout procedure, there should be no "cold spots" to condense volatilized contaminants.

An effective way to desorb species from a surface is to use inert gas "ion scrubbing," which uses low-energy ions accelerated across the plasma–wall sheath to enhance desorption as shown in Figure 4. The plasma is established by having a

high-voltage negative electrode ("glow bar") in the vacuum system and establishing a gas pressure of 10–100 mtorr of inert or reactive gas. A DC gas discharge is established, and all surfaces in contact with the plasma attain a negative potential with respect to the plasma across the "wall sheath." Bombardment by ions accelerated across the wall sheath enhances desorption of adsorbed contaminants from the wall. To obtain the most uniform bombardment, the glow bar should extend throughout the vacuum vessel. Because the gas throughput of the system is low during many types of plasma cleaning, often a conditioning cycle of "pump–plasma clean–pump" is used to flush desorbed contamination from the system.

Desorption can be stimulated by the adsorption of short-wavelength radiation. This process is called *photodesorption* and is stimulated by the short-wavelength radiation from a quartz-ultraviolet lamp [112,113] or plasma in the vacuum system.

REACTION AND VOLATILIZATION

Ion scrubbing with a reactive gas such as oxygen or hydrogen results in oxidation or hydrogenation of contaminants [76,114,115]. If the reaction products are volatile, they are pumped away. The cleaning can be monitored by residual gas analysis such as CO and CO_2 in an oxygen plasma and C_xH_y in a hydrogen plasma. If the reaction products are nonvolatile, they will be left as a residue on the surface. *Caution*: A vacuum system using a compression-type pump with hydrocarbon oil should not be used to pump pure oxygen, because an explosion can occur — instead, use air (21% O_2) as the plasma gas.

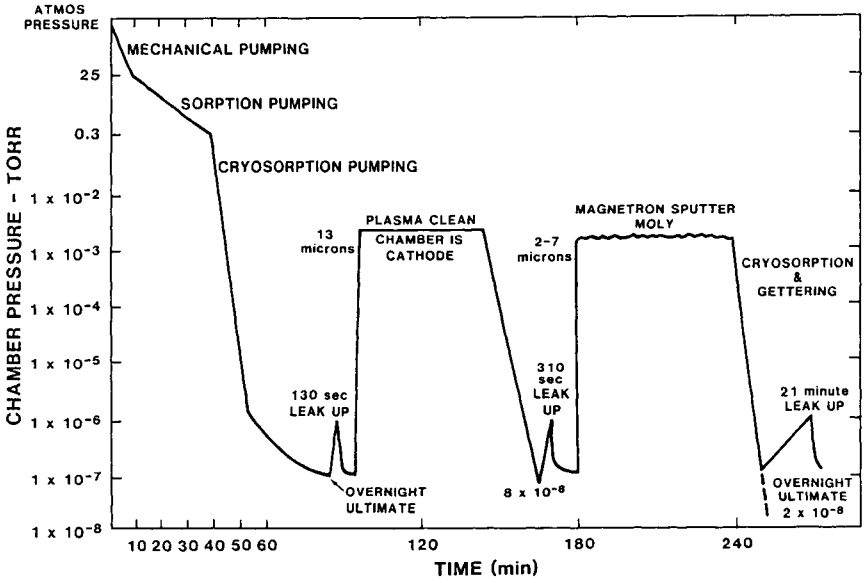
The UV/O_3 cleaning process (Section 4.9.2.3) can be used for *in situ* cleaning. This procedure requires a partial pressure of oxygen and a UV source in the vacuum chamber.

Instead of oxygen or hydrogen plasmas, more chemically reactive materials such as chlorine and fluorine can be used for reactive plasma cleaning [57,107]. The chlorine or fluorine reacts with the contaminants or surface material to form a volatile fluoride or chloride compound. This process is similar to plasma etching and usually generates a lot of particulate matter in the system. For example, chlorine from BCl_3 is used to etch aluminum and chlorine from CCl_4 and fluorine from CF_4 is used to etch silicon. Fluorine from NF_3 is used to etch glasses and ceramics such as that deposited by PECVD processing [116].

SPUTTER CLEANING

The chamber surfaces of a vacuum system can be sputtered using the "glow bar" as a high-voltage positive electrode and the vacuum surfaces as a negative ground [117]. Figure 8 shows a conditioning cycle for a stainless steel chamber used to sputter molybdenum.

Fig. 8.



Conditioning of stainless steel chamber for the sputter deposition of molybdenum. Conditioning is by sputter cleaning the vacuum surfaces. Deposition of molybdenum provides getter pumping in the system. Effectiveness of the cleaning is shown by the leak-up rate. [117]

4.9.5 DOCUMENTATION

The key to reproducible cleaning and surface preparation is documentation. Documentation is also important in the transfer of a process or product from research and development to manufacturing, in improving the process over time, and in qualifying for the ISO 9000 certifications. In many instances, lack of proper documentation has resulted in loss of product yield and even in loss of the process itself, if critical information has not been documented.

Documentation starts with the laboratory/engineering (L/E) notebook, where the experiments, trials, and results of development work are documented. Where the data are not amenable to direct entry, a summary of the findings can be entered into the L/E notebook and reference made to particular charts, graphs, memos, etc. To ensure unquestionable entries, the L/E notebook should be hardbound and have numbered pages, and entries should be handwritten, dated, and initialed. If an entry is made about a patentable process, product, or idea, the entry should be read by another person, then initialed and dated with the statement

"Read and understood" beside the entry. Patents are developed from L/E notebooks, and dated entries are important if questions are ever raised about when and where an idea was conceived or a finding made.

Process specifications ("specs") are essentially the "recipe" for the process and are the goal of a focused R&D process development effort. Specifications define what is done, the critical process parameters, and the process parameter limits that will produce the desired product. The specification can also define the substrate material, materials to be used in the processing, handling and storage condition, packaging, process monitoring and control techniques, safety considerations, and any other important aspect of the processing. Specifications should be dated, and there should be a procedure available that allows changes to the specifications. Reference should be made to the particular "issue" (date) of specifications. Specifications should be based on accurate measurements, so it is important that calibrated instrumentation be used to establish the parameter windows for the process. Specifications usually do not specify specific equipment and noncritical process parameters. Specifications can also be used to define the functional and stability properties of the product, and associated test methods.

Manufacturing processing instructions (MPIs) are derived from the specifications as they are applied to specific equipment and manufacturing procedures. Often the MPIs contain information that is not found in the specifications but is important to the manufacturing flow. This may be something such as the type of gloves to be used with specific chemicals (e.g., no vinyl gloves around alcohol, rubber gloves for acids). The MPIs should be dated and updated in a controlled manner. The MPIs should also include the appropriate manufacturing safety data sheets (MSDSs) for the materials being used. In many cases the MPIs should be reviewed with the R&D staff who have been involved in writing the specifications, to ensure that mistakes are not made. The R&D staff should be included in process review meetings for the same reason. In some cases, MPIs and specifications must be written from an existing process. Care must be taken that the operators reveal all the important steps and parameters.

In manufacturing, it is important to keep equipment logs for the equipment and instrumentation being used. These logs contain information as to when and how long the equipment was used, its performance, any modifications that are made, and any maintenance and service that has been performed. For example, for a vacuum system used for thin-film deposition, the log should include entries on performance such as

- Time to crossover pressure (roughing to high-vacuum pumping)
- Time to the basepressure specified
- Leak-up rate between specified pressure levels
- Process being performed
- Chamber pressure during processing
- Fixturing used

- Nnumber and type of substrates being processed
- Mass spectrometer trace at basepressure and during processing
- Total run time

In critical applications, the system performance can be evaluated by statistical analysis [118].

The equipment logs can be used to establish routine maintenance schedules and determine the cost of ownership (COO) of that particular equipment. When the equipment is being repaired or serviced, it is important to log the date, action, and person doing the work. The equipment log should also contain the calibration log(s) for associated instrumentation, if applicable.

4.9.6 CONCLUSION

Surface preparation and cleaning are necessary to provide the best possible vacuum in the least amount of time. This is important to determining the "cycle time" in vacuum-based processing. Maintaining the surface condition is important to obtaining a reproducible processing environment with repeated use.

TRADE NAMES

Snow Floss® — Manville Corporation
Scotch Brite® — 3M Company
Soft-scour® — 3M Company
GAF NMP®, GAF M-PYROL® — GAF Corporation
Bioact EC-7® — Petroform Company
Genesolv 2000® — Allied Signal Company
Fluoroinert® — 3M Company
Krytox® — E. I. DuPont Company
Teflon® — E. I. DuPont Company
Viton® — E. I. DuPont Company
Combat®, boron nitride — Carborundum Corporation

REFERENCES

1. N.V. Cherepnin, *Treatment of Materials for Use in High Vacuum* (Ordentlich, Frankfurt, 1976).
2. Fred Rosebury, *Handbook of Electron Tubes and Vacuum Techniques* (Addison-Wesley, New York, 1965), (AVS reprint), p. 1.

3. J.F. O'Hanlon, Contamination reduction in vacuum processing systems, *J. Vac. Sci. Technol.*, **A7**(3) (1989) 2500.
4. Y.T. Sasakim A survey of vacuum material cleaning procedures: A subcommittee report on the American Vacuum Society Recommended Practices Committee, *J. Vac. Sci. Technol.*, **A9**(3) (1991) 2025.
5. B. S. Halliday, Cleaning materials and components for vacuum use, in *Modern Vacuum Practice: Design, Operation, Performance and Application of Vacuum Equipment*, edited by G. F. Weston, special issue of *Vacuum*, **37**(8/9), 1987.
6. L. E. Samuels, Mechanical grinding, abrasion and polishing, in *Metals Handbook*, 9th ed., Vol. 9: *Metallography and Microstructure*, (American Society for Metals, 1985), p. 33.
7. John F. Jumer, Electropolishing, in *Metal Finishing Guidebook Directory 1994* (Elsevier, New York, 1994), p. 383.
8. P.V. Shigolev, *Electrolytic and Chemical Polishing of Metals*, (Freund Publishing, London, 1974).
9. U. Schafer and J. Beuers, Instructions and tips concerning chemical-mechanical polishing, *Metallography*, **18** (1985) 319.
10. R. Venkatachalam, S. Mohan, and S. Guruviah, Electropolishing of stainless steel from a low concentration phosphoric acid electrolyte, *Metal Finish*, **89**(4) (1991) 47.
11. J. A. Knapp, D. M. Follstaedt, and B. L. Doyle, *Nucl. Instrum. Method*, **87–88** (1985) 38.
12. A. Boschi, C. Ferro, G.Luzzi, and L. Papagno, *J. Vac. Sci. Technol.*, **16** (1979) 1037.
13. H. C. Hseuh and X. Cui, "Outgassing and desorption of the stainless-steel beam tubes after different degassing treatments, *J. Vac. Sci. Technol.*, **A7**(3) (1989) 2418.
14. N. Yoshimura, T. Sato, S. Adachi, and T. Kanazawa, Outgassing characteristics and microstructure of an electropolished stainless steel surface, *J. Vac. Sci. Technol.*, **A8**(2) (1990) 924.
15. J. R. Young, Outgassing characteristics of stainless steel and aluminum with different surface treatments, *J. Vac. Sci. Technol.*, **6**(3) (1969) 398.
16. R. W. Bonham and D. M. Holloway, Effects of specific surface treatments on type 304 stainless steel, *J. Vac. Sci. Technol.*, **14**(2) (1977) 745.
17. R. O. Adams, A review of the stainless steel surface, *J. Vac. Sci. Technol.*, **A1** (1983) 12.
18. *SEMATECH Test Method for AES Analysis of Surface and Oxide Composition of Electropolished Stainless Steel Tubing for Gas Distribution System Components (Provisional)* SEMASPEC 91060573A-STD. Sematech, 2706 Montopolis Dr., Austin, TX 78741.
19. *SEMATECH Test Method for ESCA Analysis of Surface Composition and Chemistry of Electropolished Stainless Steel Tubing for Gas Distribution System Components (Provisional)*— SEMASPEC 90120403A-STD. Sematech, 2706 Montopolis Dr., Austin, TX 78741.
20. H. Tomari, H. Hamada, Y. Nakahara, K. Sugiyama, and T. Ohmi, Metal surface treatment for semiconductor equipment: Oxygen passivation, *Solid State Technology*, **34**(2) (1991) S1.
21. D. Verma, Surface passivation of AISI 400 series stainless steel components, *Metal Finish*, **86**(2) (1988) 85.
22. Earl C. Groshart, Pickling and acid dipping, *Metal Finishing Guidebook Directory 1994* (Elsevier, New York, 1994), p. 153.
23. N. Kaufherr, A. Krauss, D. M.. Gruen, and R. Nielsen, Chemical cleaning of aluminum alloy surfaces for use as vacuum material in synchrotron light sources, *J. Vac. Sci. Technol.*, **A8**(3) (1990) 2849.
24. P. L. Oliphant, The cleanroom enigma, *Semicond Internat.*, **15**(10) (1992) 82.
25. H. Ishimaru, Ultimate pressure of the order of 10^{-13} in an aluminum vacuum chamber, *J. Vac. Sci. Technol.*, **A7**(3) (1989) 2439.
26. H. Ishimaru, Developments and applications for all-aluminum alloy vacuum systems, *MRS Bull.*, **15**(7) (1990) 23.
27. M. Suemitsu, T. Kaneko, and N. Miyamoto, Aluminum alloy ultrahigh vacuum chamber for molecular beam epitaxy, *J. Vac. Sci. Technol.*, **A5**(1) (1987) 37.

28. A. W. Brace, *The Technology of Anodizing Aluminum* (Robert Draper Publishing, New York, 1968).
29. David Thomas, Anodizing aluminum, in *Metal Finishing Guidebook Directory 1988* (Elsevier, New York, 1988), p. 451.
30. J. K. G. Panitz and D. J. Sharp, The effect of different alloy surface compositions on barrier anodic film formation, *J. Electrochem. Soc.*, **131**(10) (1984) 2227.
31. E. D. Erikson, T. G. Beat, D. D. Berger, and B. A. Fraizer, Vacuum outgassing of various materials, *J. Vac. Sci. Technol.*, **A2**(2) (1984) 206.
32. D. J. Santeler, Estimating the gas partial pressure due to diffusive outgassing, *J. Vac. Sci. Technol.*, **A10**(4) (1992) 1879.
33. J. H. Tilley, Release agent for system cleaning, *SVC 38th Annual Technical Conference Proceedings*, SVC Publications, Albuquerque, NM, 1995, p. 457.
34. J. Winter, Surface conditioning of fusion devices by carbonization: Hydrogen recycling and wall pumping, *J. Vac. Sci. Technol.*, **A5**(4) (1987) 2286.
35. Ted Kostilnik, Mechanical cleaning systems, in *Surface Engineering*, ASM Handbook, Vol. 5 (ASM Publications, Materials Park, OH, 1994), p. 55.
36. Robert C. Mulhall and Nicholas D. Nedas, Impact blasting with glass beads, in *Metal Finishing Guidebook and Directory* (Elsevier, New York, 1994), p. 75.
37. G. P. Balcar and M. M. Woelfel, Specifying glass beads, *Metal Finish*, **83**(12) (1985) 13.
38. B. E. Durst, Non-chemical cleaning of fixtures and surfaces using plastic blast media, *Society of Vacuum Coaters 35th Annual Technical Conference Proceedings 1992* (SVC Publications, Albuquerque, NM, 1992), p. 271.
39. Mark Hanna, Blast Finishing, *Metal Finishing Guidebook and Directory* (Metal Finishing, 1994), p. 68.
40. Stanley Hirsch and Charles Rosenstein, Stripping metallic coatings, in *Metal Finishing Guidebook and Directory* (Elsevier, New York, 1995), p. 428.
41. D. M. Mattox, Surface effects on the growth, adhesion and properties of reactively deposited hard coatings, *Surf. Coat. Technol.*, **81** (1996) 8.
42. T. Spalvins, A review of recent advances in solid film lubricants, *J. Vac. Sci. Technol.* **A5** (1987) 212.
43. David B. Chalk, Classification and selection of cleaning processes, in *Surface Engineering*, ASM Handbook, Vol. 5 (ASM Publications, Materials Park, OH, 1994), p. 3.
44. V. B. Menon, Particle adhesion to surfaces: Theory of cleaning, *Particle Control for Semiconductor Manufacturing*, edited by R. P. Donovan (Marcel Dekker, New York, 1990), p. 359.
45. R. A. Bowling, An analysis of particle adhesion on semiconductor surfaces, *J. Electrochem. Soc.*, **132** (1985) 2208.
46. M. B. Renade, Adhesion and removal of fine particles on surfaces, *Aerosol. Sci. Technol.*, **7**(2) (1987) 161.
47. I. F. Stowers, Advances in cleaning metal and glass surfaces to micron-level cleanliness, *J. Vac. Sci. Technol.*, **15**(2) (1978) 751.
48. J. M. Bennett, L. Mattsson, M. P. Keane, and L. Karlsson, Test of strip coating material for protecting optics, *Appl. Optics*, **28**(5) (1989) 1018.
49. E. J. Wegener, Glass cleaning utilizing a cerium oxide solution in a high volume production environment, *Society of Vacuum Coaters 36th Annual Technical Conference Proceedings* (SVC Publications, Albuquerque, NM, 1993), p. 495.
50. L. Layden and D. Wadlow, High velocity carbon dioxide snow for cleaning vacuum system surfaces, *J. Vac. Sci. Technol.*, **A8**(5) (1990) 3881.
51. S. P. Hotaling, Adapting military technology to civilian use: Contamination removal and collection techniques, *Microcontamination*, **11**(5) (1993) 32.
52. R. Sherman and W. Whitlock, The removal of hydrocarbons and silicone oil stains from silicon wafers, *J. Vac. Sci. Technol.*, **B8**(3) (1990) 563.

53. *Etching Composition and Processes* (Chemical technology review 210), edited by M. J. Collie (Noyes Publications, Park Ridge, NJ, 1990).
54. Perrin Walker and William H. Tarn, *Etchants for Metals and Metallic Compounds*, CRC Handbook (CRC Press, Boca Raton, FL, 1990).
55. *Bibliography on Chemical Cleaning of Metals Vol. 1 (52135) & Vol. 2 (52129)*, (NACE (National Association of Corrosion Engineers) Publications, Houston, TX).
56. R. C. Sundahl, Relationship between part surface chemistry and adhesion of thin films, *J. Vac. Sci. Technol.*, **9** (1972) 181.
57. G. J. Kominiak and D. M. Mattox, Reactive plasma cleaning of metals, *Thin Solid Films*, **40** (1977) 141.
58. R. Hamilton, Volatile organic compound destruction using thermal processing, *Solid State Technol.*, **34**(9) (1991) 51.
59. P. Taylor, M-Pyrol—A practical replacement solvent, *Plat. Surf. Finish.*, **76** (1989) 42.
60. V. Wang and A. N. Merchant Metal-cleaning alternatives for the 1990s, *Metal Finish*, **91**(4) (1993) 13.
61. C. D. D'Ruis, *Aqueous Cleaning as an Alternative to CFC & Chlorinated Solvent based Cleaning*, (Noyes Publications, Park Ridge, NJ, 1991).
62. G. M. Schneider, Physicochemical principles of extraction with supercritical gases, *Angew. Chem. Int. Ed. Engl.* **17** (1978) 716.
63. E. Bok, D. Kelch, and K. S. Schumacher, Supercritical fluids for single wafer cleaning, *Solid State Technol.*, **35**(6) (1992) 117.
64. Gerald J. Cormier, Alkaline cleaning, *ASM Handbook*, vol. 5: *Surface Engineering* (ASM International, Materials Park, OH, 1994).
65. V. L. Morris, Cleaning agents and techniques for concentrating solar collectors, *Solar Energy Materials*, **3** (1979) 35.
66. Werner Kern, ed., *Handbook of Semiconductor Wafer Cleaning Technology* (Noyes Publications, Park Ridge, NJ, 1993).
67. L. Holland, The cleaning of glass, in *The Properties of Glass Surfaces* (Wiley, New York, 1964), chap. 5.
68. Cleaning of substrate surfaces, in *Coatings on Glass*, H. K. Pulker (Elsevier, New York, 1984), chap. 4.
69. R. R. Sowell, R. E. Cuthrell, R. D. Bland, and D. M. Mattox, Surface cleaning by ultraviolet radiation, *J. Vac. Sci. Technol.*, **11** (1974) 474.
70. J. R. Vig , UV/ozone cleaning of surfaces, *J. Vac. Sci. Technol.*, **A3**(3) (1985) 1027.
71. J. M. Frank, Vacuum processing equipment for quartz crystal oscillators, in *Proc. 35th Annual Frequency Control Symposium* (IEEE Publications, New York, 1981), p. 40.
72. H. Norstrom, M. Ostling, R. Buchta, and C. S. Petersson, Dry cleaning of contact holes, *J. Electrochem. Soc.*, **132** (1985) 2285.
73. R. Bouwman, J. B. van Mechelen, and A. A. Holscher, Surface cleaning by low temperature bombardment with hydrogen an AES investigation on copper and Fe-Cr-Ni steel surfaces, *J. Vac. Sci. Technol.*, **15**(1) (1978) 91.
74. L. Holland, Wetting glass surfaces, in *The Properties of Glass Surfaces*, (Wiley, New York, 1964), chap. 6.
75. M.A. Baker, Plasma cleaning and the removal of carbon from metal surfaces, *Thin Solid Films*, **69** (1980) 359.
76. H. F. Dylla, A review of the wall problem and conditioning techniques for TOKAMAKS, *J. Nucl. Mat.*, **93, 94** (1980) 61.
77. H. Yamada, "Low temperature surface cleaning method using low-energy reactive ionized species, *J. Appl. Phys.*, **65**(2) (1989) 775.
78. R. P. H. Chang, C. C. Chang, and S. Darack, Hydrogen plasma etching of semiconductors and their oxides, *J. Vac. Sci. Technol.*, **20**(1) (1982) 45.

79. M. K. Balazs, A summary of new methods for measuring contaminants in ultrapure water, *Microcontamination*, **5**(1) (1987) 35.
80. Vicki L. Rupp and Ken Surprenant, Solvent cold cleaning and vapor degreasing, in *ASM Handbook*, 5: *Surface Engineering* (ASM International, Materials Park, OH, 1994), p. 21.
81. B. Kanegsberg, Cleaning systems for low-flashpoint solvents, *Precision Clean*, **3**(3) (1995) 21.
82. Joseph Scapelli, Enclosed vapor degreasing systems, *Metal Finishing Guidebook and Directory* (Elsevier, New York, 1995), p. 144.
83. Kenneth R. Allen, Ultrasonics — A Practical Approach, in *Metal Finishing Guidebook and Directory* (Elsevier, New York, 1995), p. 152.
84. Jeff Hancock, Ultrasonic cleaning, in *ASM Handbook*, Vol. 5: *Surface Engineering* (ASM International, Materials Park, OH, 1994), p. 44.
85. F. J. Fuchs, Ultrasonic cleaning, in *Metal Finishing Guidebook and Directory* (Elsevier, New York, 1992).
86. J. H. Schleckser, Process control ultrasonics, *Society of Vacuum Coaters 35th Annual Technical Conference Proceedings (1992)* (SVC Publications, Albuquerque, NM, 1992), p. 15.
87. J. Halbert, Using ultrasonic techniques for wet-process cleaning, *Microcontamination*, **6**(11) (1988) 36.
88. S.M. Goho, An Auger electron spectroscopy study of the oxidation and mechanical degradation of Ta thin film protective overlayers at microelectronic solid/liquid interfaces, *J. Vac. Sci. Technol.*, **A8**(3) (1990) 3026.
89. T. Ohmi, H. Mishima, T. Mizuniwa, and M. Abe, IPA vapor-drying technology: Developing contamination-free cleaning and drying technologies, *Microcontamination*, **7**(5) (1989) 25.
90. A. E. Walter and C. F. McConnell, Direct displacement wet processing: How it affects wafer surface phenomena, *Microcontamination*, **8**(1) (1990) 35.
91. B. Kanegsberg, Cleaning systems for low-flashpoint solvents, *Precision Clean*, **3**(3) (1995) 21.
92. Robert J. Good, Contact angle, wetting, and adhesion: a critical review, in *Contact Angle, Wettability and Adhesion*, edited by K. L. Mittal (VSP Publishers, Utrecht, The Netherlands, 1993), p. 3.
93. A. K. Furr, *CRC Handbook of Laboratory Safety*, 3rd ed. (1989).
94. W. Mahn, *Fundamentals of Laboratory Safety*, (Van Nostrand Reinhold, New York, 1991).
95. D. J. Young, How to use materials safety data sheets in PWB manufacturing, *Electri•onics*, **32**(5) (1986) 40.
96. *SEMATECH Guide for Contamination Control in the Design, Assembly and Delivery of Semiconductor Manufacturing Equipment*, #92051107A-STD (July 10, 1992). Sematech, 2706 Montopolis Dr., Austin, TX 78741.
97. Bennie W. Goodwin, Cleanroom garments and fabrics, in *Handbook of Contamination Control in Microelectronics*, edited by Donald L. Tolliver (Noyes Publications, 1988), chap. 3.
98. *Annual buyers guide: A guide to ultraclean suppliers and products*, Micro (formally *Microcontamination*) Magazine (Canon Publishing, Santa Monica, CA).
99. *Handbook of Contamination Control in Microelectronics*, edited by Donald L. Tolliver (Noyes Publications, 1988).
100. *Particle Control for Semiconductor Manufacturing*, edited by R. P. Donovan (Marcel Dekker, New York, 1990).
101. *Compendium of Standards, Practices, Methods etc. Relating to Contamination Control* (IES-C-CC-009-84-T) available from the Institute of Environmental Sciences.
102. G. A. Harvey, James L. Raper, and D. C. Zellers, Measuring low-level nonvolatile residue contamination on wipes, swabs and gloves, *Microcontamination*, **8**(11) (1990) 43.
103. The detection and control of organic contaminants on surfaces, Malcolm L. White, in *Clean surfaces: Their preparation and characterization for interfacial studies*, edited by George Goldfinger (Marcel Dekker, New York, 1970), p. 361.

104. J. J. Wu, D. W. Cooper, and R. J. Miller, Aerosol model of particle generation during pressure reduction, *J. Vac. Sci. Technol.* **A8**(3) (1990) 1961.
105. D. Chen and S. Hackwood, Vacuum particle generation and the nucleation phenomena during pumpdown, *J. Vac. Sci. Technol.*, **A8**(2) (1990) 933.
106. Christoph Steinbruchel, The Formation of particles in thin-film processing plasmas, in *Physics of Thin Films*, Vol. 18: *Plasma Sources for Thin Film Deposition and Etching*, edited by Maurice H. Francombe and John L. Vossen (Academic Press, San Diego, 1994).
107. *Plasma Etching: An Introduction*, edited by Dennis M. Manos and Daniel L. Flamm (Academic Press, San Diego, 1989).
108. H. Ishimaru, K. Itoh, T. Ishigaki, and S. Furutate, Fast pump-down UHV aluminum vacuum system using super-dry nitrogen gas flushing, *J. Vac. Sci. Technol.*, **A10**(3) (1992) 547.
109. G. Comsa and R. David, Dynamical parameters of desorbing molecules, *Surf. Sci. Reports*, **5** (1985) 145.
110. G. Carter, P.Bailey, and D. G. Armour, The precise deduction of desorption activation energy distributions from thermal evolution spectra, *Vacuum*, **34**(8/9) (1984) 797.
111. J. F. O'Hanlon, Thermal desorption measurement technique, *J. Vac. Sci. Technol.*, **A9**(1) (1991) 1.
112. G. W. Fabel, S. M. Cox, and D. Lichtman, Photodesorption from 304 stainless steel, *Surf. Sci.*, **40** (1973) 571.
113. P. Danielson, Understanding water vapor in vacuum systems, *Microelectron Manuf. Test*, **13**(8) (1990) 24.
114. R. P. Govier and G. M. McCracken, Gas discharge cleaning of vacuum surfaces, *J. Vac. Sci. Technol.*, **7**(5) (1970) 552.
115. *Surface conditioning of vacuum systems*, edited by Robert A. Langley, Daniel L. Flamm, H. C. Hseuh, Wen L. Hsu, and Thomas W. Rusch, AIP Conference Proceedings No. 199, American Vacuum Society Series 8 (American Institute of Physics, 1990).
116. R. Anderson, J. Behnke, M. Berman, H. Kobeissi, B. Huling, J. Langan, S.-Y. Lynn, and R. Morgan, Using COO to select nitride PECVD clean cycle, *Semicond. Internat.*, **16**(11) (1993) 86.
117. D. M. Mattox, R. E. Cuthrell, C. R. Peeples, and P. L. Dreike, Design and performance of a moveable-post-cathode magnetron sputtering system for making PBFA II accelerator sources, *Surf. Coat. Technol.*, **33** (1987) 425.
118. V. S. Dharmadhikari, R. O. Lynch, W. Brennan, and W. Cronin, Physical vapor deposition equipment evaluation and characterization using statistical methods, *J. Vac. Sci. Technol.*, **A8**(3) (1990) 1603.

PART 5

Vacuum Applications

5.1 High-Vacuum-Based Processes: Sputtering 609
- 5.1.1 Sputtering and Deposition 611
- 5.1.2 Sputter Deposition Technologies 612
- 5.1.3 Magnetron Applications 624
- 5.1.4 Future Directions in Sputtering 626

References 627

5.2 Plasma Etching 628
- 5.2.1 Introduction 628
- 5.2.2 Review of Plasma Concepts Applicable to Etching Reactors 628
- 5.2.3 Basic Plasma Etching Requirements 633
- 5.2.4 Plasma Diagnostics 641
- 5.2.5 Basic Plasma Etch Reactors 643
- 5.2.6 Advanced Plasma Etch Reactors 649
- 5.2.7 New Trends 665

References 667

5.3 Ion Beam Technology 672
- 5.3.1 Introduction 672
- 5.3.2 Ion Beam Etching 678
- 5.3.3 Ion Beam Sputter Deposition 683
- 5.3.4 Ion-Beam-Assisted Deposition 687
- 5.3.5 Ion Beam Direct Deposition 689
- 5.3.6 Conclusion 690

References 691

5.4 Pulsed Laser Deposition 694
- 5.4.1 Introduction 694
- 5.4.2 Pulsed Laser Deposition System 695
- 5.4.3 The Ablation Mechanism 698
- 5.4.4 Advantages and Limitations 700
- 5.4.5 Materials Survey 705
- 5.4.6 Future Outlook 708

References 708

5.5 Plasma-Enhanced Chemical Vapor Deposition 711
5.5.1 Introduction 711
5.5.2 Equipment and Other Practical Considerations 717
5.5.3 Process Scaleup 723
5.5.4 Conclusion 727
References 728
5.6 Common Analytical Methods for Surface and Thin Film 731
5.6.1 Introduction 731
5.6.2 The Elcctron Spectroscopies 732
5.6.3 Methods Bascd on Ion Bombardment 745
5.6.4 UHV Generation and System Considerations for Surface Analysis 755
References 757

CHAPTER 5.1

High-Vacuum-Based Processes: Sputtering

Stephen M. Rossnagel
I.B.M., T. J. Watson Research Center

Sputtering is a process by which an energetic particle strikes a surface with sufficient energy that one or more of the surface atoms is ejected from the surface. The incident particle could be an ion, an atom, an electron, an alpha particle, etc., although for most practical purposes ions — typically inert gas ions — are used. The sputter yield is defined as the ratio of the number of emitted atoms to the number of incident particles:

$$Y = \text{emitted (sputtered) atoms/incident particle}$$

The sputtering process has four basic regimes, based on the incident energy. At very low energies approaching the surface-binding energy, the yield is very low ($<10E^{-4}$), and this is known as "near-threshold sputtering." There is some debate over whether a true threshold exists for sputtering, although ion energies below the binding energy are extremely unlikely to result in the removal of surface atoms.

At moderate energies of perhaps 30 eV up to about 1000 eV, sputtering is characterized by "knock on" effects in which the incident ion collides with a surface atom, the two atoms then collide with additional atoms, and eventually this chain of events may lead back to the surface, resulting in emission of a near-surface atom or two. This regime is difficult to model or calculate because the exact point of impact (between two atoms, directly on top of one atom, etc.) results in a different chain of collisions. This requires a more Monte Carlo approach, using large

ISBN 0-12-325065-7
$25.00

numbers of trial impacts. The sputter yields in this region are generally in the range of 0.1–5 atoms per incident ion.

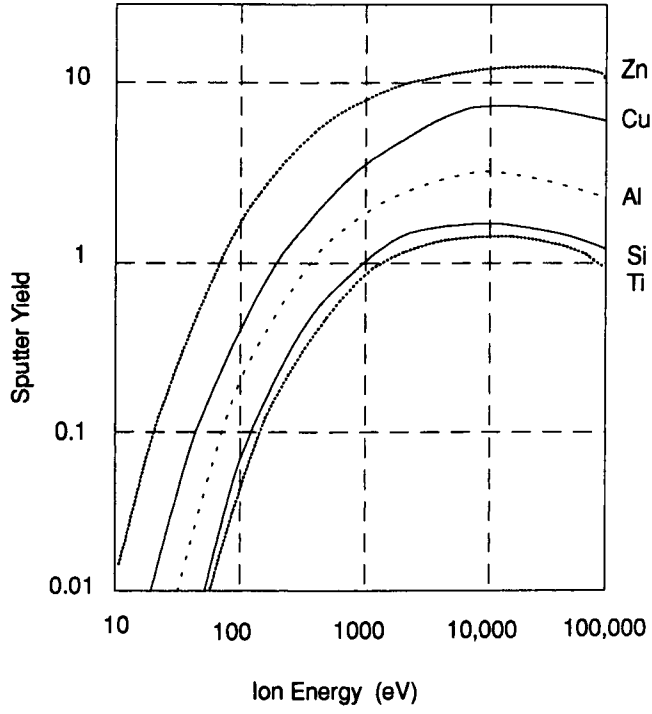
At ion energies of 1 to 50 keV or so, the incident ion has sufficient energy to break essentially all the bonds in a semispherical region at the point of impact. This "cascade" of collisions can be calculated with high accuracy by using statistical mechanics and essentially treating the collision cascade volume as a sort of gas. This approach, first developed by Sigmund in the late 1960s has been very successful at predicting sputter yields [1,2]. Yields in this region are generally two–five times higher than in the direct knock-on range, depending on the material and the incident particle species.

The final, highest-energy regime is at energies in the 50keV to several MeV range, where the incident particle travels well into the bulk of the sample before slowing down and depositing its energy. This is more characteristic of implantation, and the sputter yield falls eventually to a very low level at the MeV range.

A graph of the sputter yield as a function of ion energy for Ar bombardment of Si, Zn, Ti, Cu, and Al is shown in Figure 1.

Sputter yields for many materials have been collected by numerous authors. However, many inconsistencies appear in these published results. Reported yields

Fig. 1.



The sputter yield for Ar+ bombardment for Si, Zn, Ti, Cu and Al as a function of ion energy.

Table 1

Sputter Yields as a Function of Ar Ion Energy for Commonly Used Materials

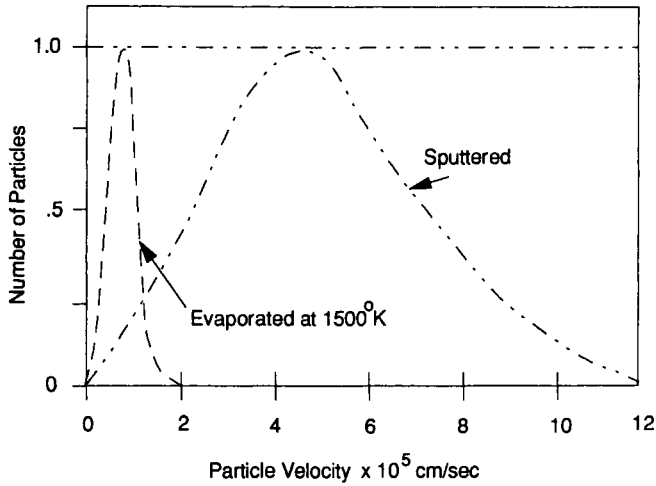
Material	300 eV	500 eV	1000 eV
Ag	1.7	2.5	3.5
Al	.6	.9	1.5
Au	1.1	1.7	2.5
B	.2	.6	1.1
Be	.3	.5	.9
C	.1	.3	.6
Cd	5	6	9
Co	.7	1	1.7
Cr	.8	1.1	1.9
Cu	1.5	1.9	2.9
Fe	.7	1	1.7
Ge	.6	1	1.5
Hf	.4	.6	.9
Mn	1.5	2	3
Mo	.3	.5	.9
Nb	.4	.6	.9
Ni	.7	1.0	1.7
Pb	2.5	3.2	4
Pd	1.5	1.8	2.5
Pt	.7	1	1.6
Re	.4	.6	1.0
Rh	.7	1	1.7
Ru	.7	1	1.7
Si	.3	.7	1
Sn	.6	.9	1.4
Ta	.3	.5	.9
Ti	.3	.5	.7
U	.5	.7	1.1
V	.4	.7	1
W	.3	.5	.9
Zn	3.7	5	7

can change by a factor of at least 2× depending on the type of experiment, the crystalline structure of the target, surface conditioning, chamber geometries, operating pressures, etc. Table 1 shows approximate sputter yields.

5.1.1 SPUTTERING AND DEPOSITION

Sputter etching and sputter deposition are differentiated simply by the observer's point of view. If the sample is bombarded with energetic ions, it is being sputter

Fig. 2.



Velocity distributions of Cu atoms evaporated at 1500°K (solid line) and Cu atoms sputtered from a Cu target with 600 eV Ar ions.

etched. If the sample is located near the surface that is being sputtered, then the emitted, sputtered atoms can be deposited in the form of a thin film on the surface. Sputter etching is described in more detail in the following chapter (5.2), on plasma etching. The remainder of this section focuses on sputter deposition.

Consistent with the fairly violent collision processes that are the basis of sputtering, the sputtered atoms generally are emitted from the surface with significantly more kinetic energy than atoms evaporated from a hot source. An example of the relative distribution of atom velocities for evaporation and sputtering is shown in Figure 2. The sputtered atom distribution has roughly a $1/E^2$ distribution on the high-energy side of the peak, which extends to fairly high energies. These high-energy sputtered atoms may actually cause secondary sputtering at the sample surface, a concept known as "intrinsic resputtering" [3]. The higher energy for the sputtered atoms can significantly change the film microstructure, grain size, stress, etc., which are discussed later.

5.1.2

SPUTTER DEPOSITION TECHNOLOGIES

Sputter deposition has been around for nearly 150 years from the earliest days of gas discharges and plasma physics. Currently, sputter deposition is practiced by a number of high-vacuum technologies. The most commonly used is magnetron sputtering, which is a form of a diode plasma. Another common technology is RF

diode sputtering, which is used primarily for dielectrics. A third variant is to use an ion beam and remove the sample or target from direct contact with a plasma. The most common technique used here is the Kaufman ion source, originally developed for space engine applications (see Chapter 5.3, on ion-beam-based processes). Historically, though, sputter deposition started with simple DC diodes, and it is best to use these as a base to explore the other, more complicated variations.

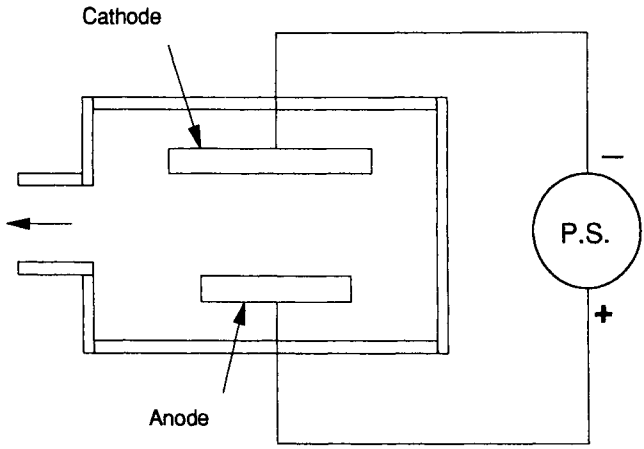
5.1.2.1 DC Diodes

A DC diode is simply an anode and a cathode placed within a vacuum system. Under suitable pressure and voltage conditions, a discharge can be formed between the two electrodes (Figure 3). The minimum voltage for the discharge is a function of the pressure and the electrode spacing [4] (see Figure 4). For purposes of this description, two electrodes of a few cm^2 in size, separated by 50 cm and in a vacuum of 100 mtorr might require 600V to form a plasma.

Ions from the plasma are accelerated to the negative electrode (the cathode), and electrons are collected at the anode. The discharge is sustained by secondary electrons that are emitted from the cathode during ion bombardment. These secondary electrons are accelerated into the plasma, gaining enough energy to make a few ions along the way, before they are collected at the anode. Because the secondary electron coefficient is typically on the order of 5%, each secondary electron needs to generate roughly 20 ion–electron pairs before it reaches the anode.

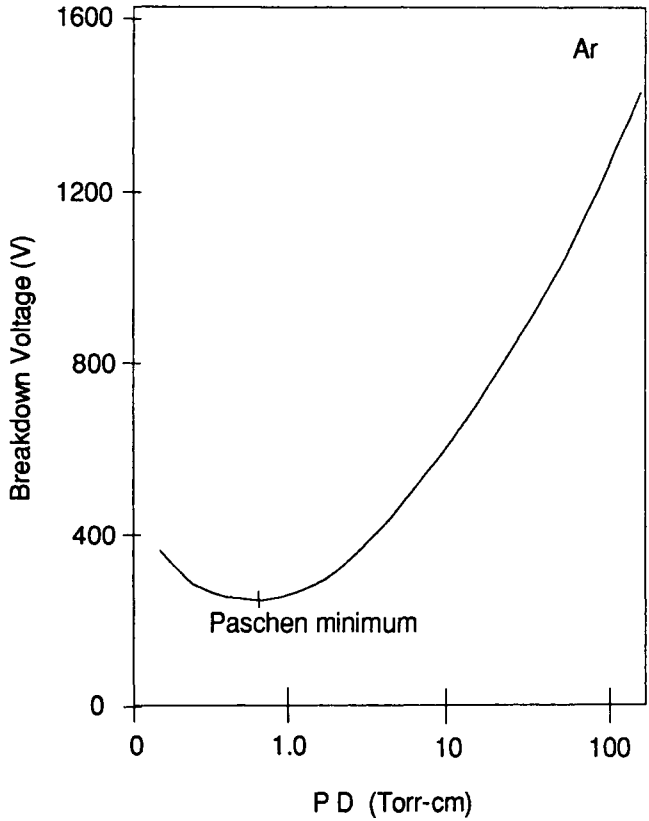
The cross section for electron impact ionization of the gas in the chamber is fairly small and energy dependent [5]. Therefore in a DC diode it is usually necessary to operate at fairly high pressures (0.2 to 2 torr) to have any reasonable ion

Fig. 3.



Typical DC diode configuration.

Fig. 4.



The Paschen curve for the breakdown voltage for a DC diode in Ar as a function of the product of the chamber pressure P and the electrode spacing D (adapted from [4]).

current. The high pressure results in poor transport of the sputtered atoms away from the cathode and subsequently a very low deposition rate. A second problem arises when the cathode is insulating. Since there is little current flow, the sputter etching and deposition process effectively stops. This would occur either if the cathode were constructed from an insulator, or if a metallic cathode was operated in the presence of sufficient oxygen. In general, DC diodes are rarely used in any practical applications.

5.1.2.2 RF Diodes

A simple variation of the DC diode results in a solution to these two fundamental problems. By applying an alternating instead of a DC voltage to the cathode, the

ion current density can be increased and concerns with the charging of insulating cathodes are eliminated. The most common AC frequency used is 13.56 MHz, which is why this technology is often known as RF diode sputtering. However, other frequencies from 60 Hz up to 100 MHz have also been used. For practical purposes, because most of the available hardware for high-power sputtering systems has been designed for 13.56 MHz, the remainder of this discussion focuses on that frequency.

The applied RF voltage to the cathode, typically on the order of several kV, results in a plasma that essentially oscillates at the same frequency. Electrons in the plasma pick up additional energy from the oscillation, in a way that has been compared to "surfing" on the applied electric "waves" [6]. In this way, energy is coupled into the electron population, which results in more ionization by the high-energy tail of the electron distribution (roughly a Maxwell-Boltzmann distribution). The higher ionization means a higher ion current at the same applied power than in a simple DC diode.

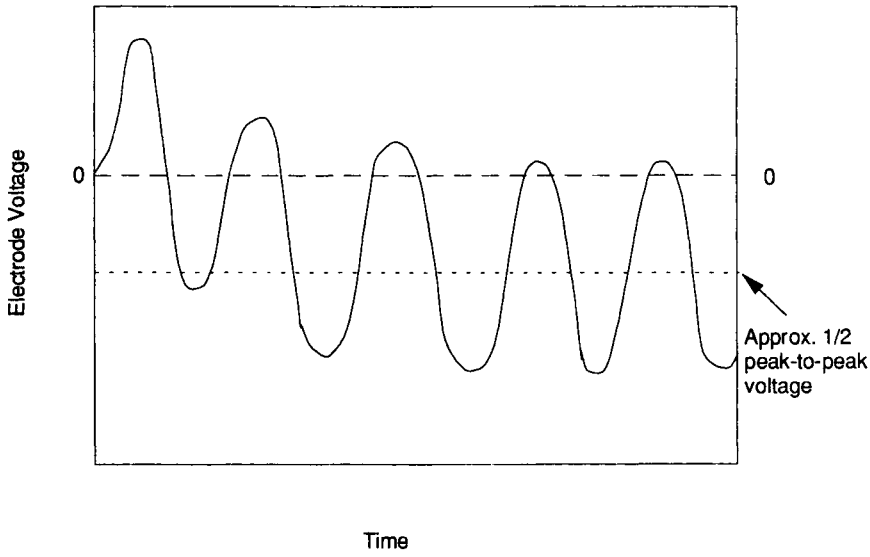
The second key advantage of RF diode sputtering is that the cathode receives no net current from the plasma. The incident ions from one part of the RF cycle are compensated for by the incident electrons from the other part of the cycle. In a sense, the cathode and anode switch places once each RF cycle, resulting in no net current or charging.

Because the electrons in a plasma move much more rapidly than the ions, the electron bombardment rate of the cathode during the positive part of the RF cycle can greatly exceed the ion flux during the negative half-cycle. If the cathode is capacitively coupled to the power supply, the net negative charge will start to look like a net negative potential on the cathode. Effectively, the DC bias of the cathode begins to drift downward with each succeeding RF cycle. As this occurs, the fraction of time that the cathode is positive with respect to the anode decreases and the net number of collected electrons decreases.

Eventually this process reaches equilibrium after just a few cycles with a net DC bias, which is just slightly less than half the applied peak-to-peak RF voltage (Figure 5). Since the ions cannot respond to the 13.56 MHz voltages (their inertia is such that they cannot respond to frequencies above a few hundred kHz), they respond to the average DC bias of the cathode. In RF diode sputtering, the ion energy is generally described by the DC bias because the potential of the plasma is usually close to zero.

RF diode sputtering, both etching and deposition, is most commonly used for the sputtering of insulating materials such as oxides (silicon dioxide, titanium dioxide, aluminum oxide, etc.) and also for the sputtering of polymers, such as polyimide. Since sputtered material is mostly emitted in the form of atoms, rather than compounds or molecules, it is often necessary to add additional reactive gas such as oxygen during the sputter deposition of oxides. Reactive sputtering is described in more detail later.

Fig. 5.



Cathode potential as a function of time showing the downward drift of the DC bias due to electron charging.

5.1.2.3 Magnetron Sputtering

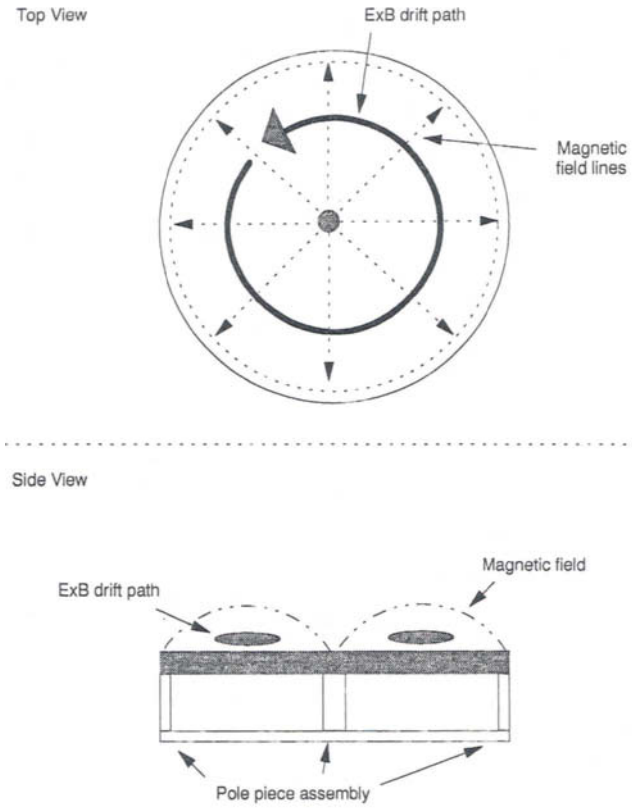
A magnetron is a cathode in a diode plasma device that can be operated either in a DC or AC mode, although the vast majority are operated DC. The magnetron uses the basic effect that electrons respond to magnetic fields by the simple relation

$$F = qv \times B \tag{1}$$

where q is the electron charge, v its velocity, and B the magnetic field. Since this is a cross product, the force is at right angles to the initial velocity, which for a free electron in the vicinity of a magnetic field results in the electron spiraling around the magnetic field line. This spiraling results in a significantly longer path length for the electron and if it is located in a plasma, a significantly higher probability that it will strike a background gas atom and perhaps cause ionization of that atom. Magnetic fields have long been used to intensify plasmas in this way, resulting in higher plasma densities and lower discharge voltages at constant applied power.

A magnetron uses a slight variation on this phenomena in which the electron now must respond to both electric and magnetic fields. If these fields are at right angles, the electron experiences a net $E \times B$ drift, which is in a direction perpendicular to the E and B fields. In solid-state physics, this is known as the Hall effect and results in a slight voltage difference across a conductor that is carrying cur-

Fig. 6.



Schematic of a magnetron cathode device.

rent in the presence of a large magnetic field. In a magnetron, though, the drift path for the electrons is configured such that it closes on itself, often in the form of a loop (Figure 6). The closed path allows electrons to travel around the drift path several [3–10] times [7], which results in a high degree of efficiency of coupling the electron energy to the plasma. This results in very low discharge voltages, high currents (amperes), and very little voltage increase as the current is increased. Magnetrons generally follow an empirical current–voltage relationship described as

$$I = kV^n \tag{2}$$

where I is the discharge current, V the voltage, k is a system-dependent constant, and the exponent n is typically in the range of 5–20.

The actual path of the secondary electron after it leaves the cathode is quite complex. The electron spirals around the magnetic field lines while it oscillates radially along the field lines, reflecting at each end of the field line as it intersects

the cathode surface. On top of these motions, the $E \times B$ drift causes a generally circular drift around the $E \times B$ drift path. In addition, since the electron is in a plasma, it is subject to a lot of interactions (noise, turbulence) with the other charged particles present.

Originally it was thought that the electron simply drifted around the drift path of the cathode until it hit a gas atom, ionized it, and then continued around the loop [8]. In reality this scenario is probably not accurate, and the electron has many more effective collisions with other electrons, resulting in a general heating of the electron distribution. Therefore, ionization by the high-energy tail of the distribution is the primary means of ionization, not direct collisions between secondary electrons and gas atoms [9].

5.I.2.4 Aspects of Magnetron Sputtering

DEPOSITION RATES AND EFFICIENCIES

In sputter deposition, high-rate, controlled, uniform deposition is generally desired. The nature of the magnetron, which is characterized by a dense plasma loop at the cathode, results in a nonuniform emission, usually from a ring or rectangular geometry. In addition, because of the system geometry, the emission profile of the sputtered atoms and the distance of the sample from the cathode, the deposition rate is difficult to predict exactly. Various studies have measured the deposition rate and profile as a function of gas pressure and throw distance. Most of these studies are only relatively quantitative in that they are system dependent, and there is little effort to calibrate the results.

There are two approaches to quantifying the deposition process with a magnetron, both effectively equivalent. The first uses the intrinsic probability that a sputtered atom emitted from the cathode will be deposited on the sample. This leads to a number between 0 and 1, with 1 being the maximum, ideal transport (all sputtered atoms deposit on the sample). An alternative approach, perhaps more practical for the operation of systems, quantifies the deposition rate in terms of "rate per watt." In this case, the general assumption is that the deposition rate increases linearly with discharge power (discharge current times discharge voltage). This assumption is fairly accurate in the region of operation of most magnetrons, where the sputter yield increases linearly with discharge voltage, and in practice the discharge voltage changes very slowly.

DEPOSITION PROBABILITY

The probability can be determined by comparing the number of atoms sputtered from the cathode with the number deposited at the sample. The number sputtered,

Table 2

Deposition Probability for Magnetron Sputtering at 1000W, 200-mm-Diameter Planar Magnetron in Large System [10]

Throw (cm)	Pressure (mtorr)	Deposition Probability
Ar Sputtering of Cu		
5 cm	5, 20, 30	0.63, 0.49, 0.54
9.5 cm	5, 20, 30	0.48, 0.47, 0.45
14.5 cm	5, 20, 30	0.39, 0.35, 0.31
Ne Sputtering of Al		
5 cm	5, 20, 30	0.80, 0.56, 0.52
9.5 cm	5, 20, 30	0.40, 0.42, 0.40
Ar Sputtering of Al		
5 cm	5, 20, 30	0.60, 0.46, 0.42
9.5 cm	5, 20, 30	0.44, 0.45, 0.35
Kr Sputtering of Al		
5 cm	5, 20, 30	0.52, 0.45, 0.38
9.5 cm	5, 20, 30	0.35, 0.27, 0.22

to a first approximation, is simply the discharge current times the sputter yield times the deposition time. This number might be accurate to 20% or so, and is limited by knowledge of the sputter yield for that particular cathode and also by the 5–10% correction caused by the secondary electron yield to the discharge current. Deposition probabilities are listed in Table 2 for a simple, 8-inch-diameter planar magnetron in an open vacuum system. Every system is different, though, and the presence of shielding, shutters, and even the chamber walls can alter these probabilities slightly. It should be noted that the best probabilities (highest) occur for low pressure, short throw distance, and a gas species that is lighter than the sputtered species. In addition, sputtering with a heavy gas, such as Kr, is often counterproductive. The sputter yield may be slightly increased, but the probability of an atom getting to the sample and making a film is reduced significantly.

DEPOSITION RATES BY POWER

These numbers are also very system dependent, based on the exact configuration of specific tools, operating conditions, etc. These rates are often shown as angstroms/min/watt. An example of this type of data is shown for a batch-type tool [11]. In this case, however, the sample pallets (37–45 cm wide) were scanned past the cathode such that the rates listed in Table 3 are less than the rates for a fixed sample directly in front of the cathode. This is a systematic problem with this type of rate measurement: The numbers apply to this system only.

It is extremely difficult to compare results such as these to another system. For

Table 3

Deposition Rate per Watt for MRC Magnetron Systems, Derived from Data by Class [11]. Units are Ång/min/watt.

Cathode (throw distance)	Focest Design 1.75 cm	Inset Design 5 cm	Planar 5 cm
Al	.29	.16	.18
Cu	.55	.31	.34
Au	.58	.32	.36
Ni–7%V	.36	.20	.22
Pt	.37	.21	.23
Ag	.83	.47	.52
Ti	.14	.07	.08
Ta	.15	.08	.09

example, present-day single-wafer semiconductor sputtering tools have a fixed sample position, which would undoubtedly mean a higher deposition rate efficiency than for a moving sample. Results for the sputter deposition of common semiconductor materials are shown in Table 4. The throw distance in this case is about 4 cm.

This general approach can be extended to specific systems in two ways. First, ignoring gas rarefaction effects (next section), the deposition efficiency of any particular system should be related to the numbers just listed by a simple, multiplicative factor. It is necessary to calibrate each system, but once done, calibration should not change without major geometry changes. Second, the dependency of the deposition rates just listed can be approximately coupled to the sputter yield, particularly for short throw distances. Therefore, when changing from a cathode included on the list to another target, a first estimate of the probable deposition rate can be made by comparing the sputter yield of the new material to those listed in the tables.

Table 4

Deposition Rates per Watt for an Applied Materials "Endura" Single-Wafer Sputter Deposition System. Cathode is 32.7-cm diameter, samples are 200-mm wafers [12].

Material	Power	Rate (Ång/min/watt)
AlCu (0.5)	12.7 kW	1 (new cathode) 0.75 (old cathode)
Ti	1 kW	0.17
TiN (nitride mode)	4 kW	0.15
Ti (collimated, 1.5:1)	7 kW	0.043

GAS RAREFACTION

One effect that was not anticipated with magnetron sputtering was the significant effect on the background gas density. This is a result of the thermalization or slowing down of the sputtered atoms due to gas-phase collisions with the Ar in the chamber. The heated Ar atoms then rarefy and reduce the local gas density in the near-cathode region [13]. Reductions of 80% or more have been observed in modest systems at pressures of 30 mtorr. This effect is strongly dependent on the sputter yield, the emitted atom velocity, the thermal conductivity of the gas, and the chamber walls. This effect may also be strongly coupled into the impedance of the discharge and the different operating voltages observed for different cathode species.

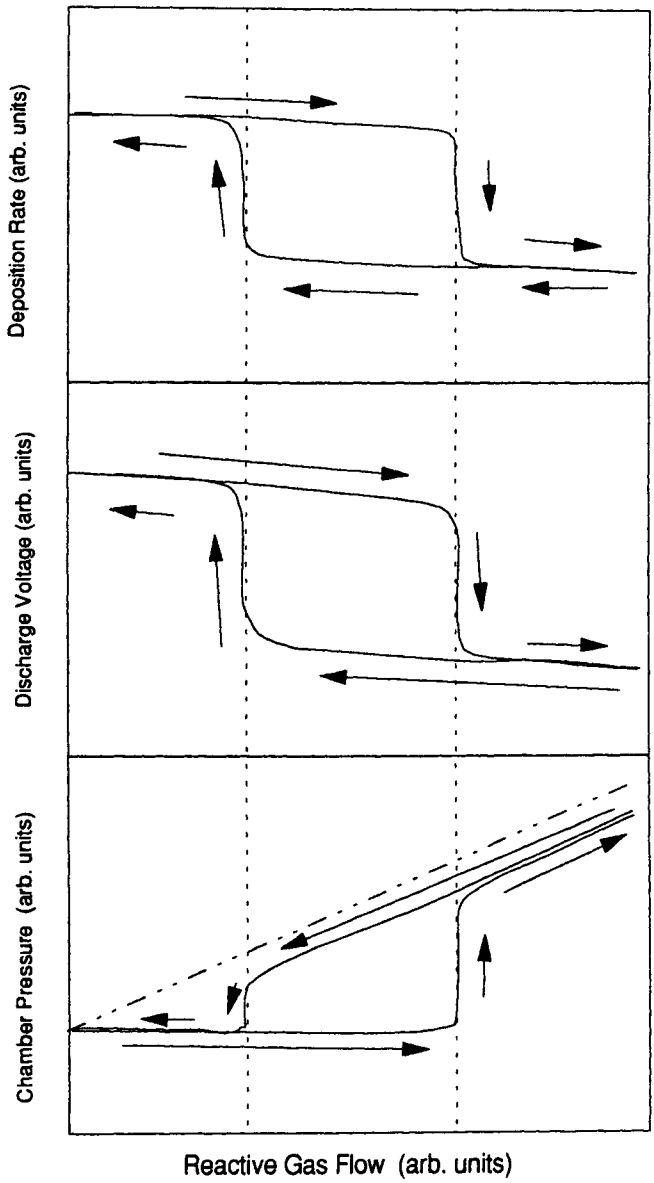
REACTIVE SPUTTERING

Reactive sputtering, and in particular reactive magnetron sputtering, is a process in which the atoms sputtered from a metallic cathode combine with background gas molecules at the sample surface to form a compound. Common examples are the sputtering of Al with oxygen present to form aluminum oxide, or the sputtering of Ti in the presence of nitrogen to form Ti-nitride. In general, the formation of a compound film occurs at the sample surface, not in the gas phase or in flight. However, this leads to a problem in that the metallic cathode is experiencing roughly the same environment as the depositing film. If there is enough of the reactive gas present to react with the metal on the film surface, generally the cathode surface also reacts.

This process can be examined phenomenologically as follows (see Figure 7): Consider an Al metal cathode being sputtered with only an inert gas species, such as Ar. The films that are deposited are purely metallic and generally the deposition rate can be rather high. If a small amount of oxygen is then added to the chamber, it is immediately incorporated into the depositing film. It is actually very difficult to observe the effects of introducing small amounts of oxygen: The deposition rate is unchanged, the discharge voltage on the cathode does not change, and the pressure does not change: No oxygen is visible on a mass spectrometer. One's first response is to check to make sure the flow controller is working, because it seems as if nothing is happening. The films, which are very reactive, are effectively gettering all the reactive gas present.

If the reactive gas flow is increased, up to a critical flow virtually all the reactive gas is still incorporated into the depositing films. In Figure 7, the flow can be increased up to the point of the second vertical dashed line. The films are at this point becoming saturated with the reactive species and are approaching their terminal (and usually desired) composition. At the critical flow, however, the films

Fig. 7.



Plots of the deposition rate (top curves), the discharge voltage (middle curves) and the chamber pressure (lower curves) for the case of increased reactive gas flow during reactive sputtering. The rightmost vertical dotted line is the critical flow point for the transition from the metallic mode to the reacted mode. The leftmost vertical dotted line is the flow point for the transition back to the metallic mode.

become completely saturated and any additional reactive gas introduced into the chamber serves to saturate the magnetron cathode, changing it from a metallic mode to a compound mode, which generally has a much lower sputter yield. This results in a rapid reduction of the sputter rate, and as a result a rapid reduction in the ability of the deposited films to getter the reactive species. The system goes through an irreversible transition at this critical flow: The deposition rate drops sharply, the discharge voltage usually changes due to the changing surface of the cathode, and the reactive species is now no longer totally absorbed by the deposited films because the supply rate of sputtered atoms has dropped significantly. Increased reactive gas flow results in relatively little change to the system: The cathode is fully reacted, and additional reactive gas atoms are irrelevant. Operation to the right side, or high-flow side, of this critical transition flow is known as the *oxide mode* for reactive sputtering in oxygen, or the *nitride mode* for reactive sputtering in nitrogen.

Reducing the flow of the reactive gas does not result in an immediate recovery of the metallic mode at pressures just below the critical flow. That is because the cathode is still fully reacted and the yield is still low. It is necessary to go to a much lower flow (the leftmost dotted line in Figure 7) to revert back to the metallic mode. These curves are often described as hysteresis curves due to their similarity to curves describing the magnetization of materials.

5.1.2.5 Bias Sputtering and Unbalanced Magnetrons

The presence of a plasma during sputter deposition allows the use of ion bombardment of the sample during deposition by ions present in the plasma. In conventional DC or RF diodes, this effect is known as *bias sputtering*. The imposition of a sample bias, either DC or RF, results in energetic ion bombardment of the growing film, which can cause several effects. From a materials point of view, the depositing film can be compressed and densified by the ion bombardment, resulting in densities close to bulk density, fine grain size, and modification of the film stress. This technique has been described as "ion peening." Obviously, too much bombardment can lead to negative effects on the film, ranging from high compressive stress to actual resputtering of the depositing film. In this latter effect, due to the angular dependence of the sputter yield, jagged or pointed topographies sputter at a faster rate than flat surfaces, and the eventual result can be increased planarization of the deposit. Reactive sputtering can also be strongly enhanced by the presence of ion bombardment, particularly for reactive chemistries such as the nitride system (AlN, TiN, etc.), which require elevated sample temperature of another energy source to push the reaction along.

Because of the good confinement of the plasma close to the cathode in the magnetron, there is usually very little ion flux available at the sample position

during magnetron sputtering. A solution to this problem was proposed by Windows [14], who developed a magnetron with unbalanced magnetic fields. This allows electrons to stream away from the cathode plasma, which tended to pull ions along toward the sample. Large substrate bias currents can then be measured at the sample, consistent with bias sputtering or reactive sputter deposition.

5.I.3 MAGNETRON APPLICATIONS

Magnetrons are the most commonly used sputtering tool for the sputter deposition of thin films. Compared to RF or DC diodes, the voltages are more approachable (500V, not several keV) and the discharge currents can be huge. In addition, the technology scales to very large dimensions: cathodes of lengths of several meters are not uncommon. These cathodes might operate at many 10s of kW, with ion currents of 25 to 100 amperes.

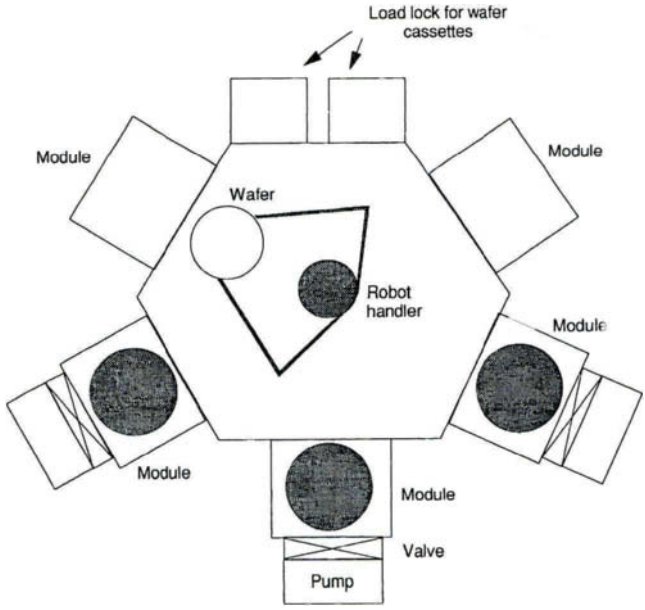
Magnetrons are used for mostly metal film deposition on a variety of substrates. One industry that relies heavily on sputter deposition is the semiconductor industry, in which a large fraction of the metal layers used for interconnects (on-chip wiring) use sputtering. The type of sputtering system used is described as a "single-wafer tool" in which only one wafer is present in the deposition chamber at a time. This is contrasted to a batch tool in which there might be many wafers experiencing the same process concurrently. Single-wafer processing has become the norm for the semiconductor industry partly due to the increasingly large size of the wafers (200 mm and heading up) and the value of the sample, which might approach $50,000 per wafer as the wafer nears completion. If anything went wrong with a batch process, potentially $1 million in wafers could be lost in a single run, whereas with a single-wafer tool the damage is contained to a single sample.

Single-wafer tools generally operate with multiple process chambers and a central handler and load-lock. An example of a single-level system is shown in Figure 8. The process modules might contain sputter deposition processes (magnetrons), or they could also be etch or Chemical Vapor Deposition (CVD) chambers.

A second large user of sputter deposition technology is the automotive industry. With the advent in the 1970s and 1980s of plastics in automobiles, sputtering has been broadly used for the metallization of these parts. Most of this is decorative, but in some cases it is also used to protect underlying parts from corrosion.

A third large user of sputtering technology is the glass-coating industry. Magnetron sputtering is used to put reflective, darkened, or decorative coatings on the

Fig. 8.



Schematic of a multiple-level integrated processing sputter tool.

large windows used in office buildings. The cathodes in these cases can be many meters long. The control of the film thicknesses in these applications is very critical: The human eye is very sensitive and can pick up intensity changes due to thickness variations on the order of tens of angstroms.

Another large application for sputtering is the deposition of metal films on plastic sheets. There are numerous uses for these metallized plastics, perhaps the most common being on the insides of potato chip or peanuts bags. The volumes of material here are very significant, and rolls of plastic sheet with lengths of many miles are typically loaded into the vacuum system (known as a web coater) for a single run.

Other applications of sputter deposition include decorative coatings for jewelry, watch bands, glasses frames, etc. These applications often use reactively sputtered Ti-nitride because of its hardness and attractive golden color.

Another application is in the field of optical coatings for very-high-quality optics applications. An application here is with laser ring gyroscopes, which are used in inertial guidance systems. Other optical applications might include spacecraft windows and the various lithographic applications used in patterning semiconductors.

5.I.4 FUTURE DIRECTIONS IN SPUTTERING

In many ways, sputtering and sputter deposition are mature technologies. Sputtering is widely used in a number of applications because it is versatile, inexpensive, and well understood. Two recent areas of sputtering technology have been developed that may have implications for the future applications of sputtering. These are (1) the ability to amplify or increase the sputter yield selectively for certain types of cathodes, and (2) the in-flight ionization of sputtered atoms and the subsequent deposition of films from ions, rather than neutrals.

5.I.4.1 Amplified Sputter Yields

Physical sputtering is a dynamic, collisional process. As such, what usually matters is the mass of the target atom and the mass and energy of the projectile. Amplified sputtering [15] uses this effect to advantage. The target is composed of a composite of the desired, to-be-sputtered species and generally a much heavier element, such as Pt. Incident ions, usually Ar or Ne, are strongly reflected from these heavy atoms, and rebound toward the near surface with most of their kinetic energy. This results in a sort of forward sputtering that can increase the emission rate for the lower-mass species by as much as an order of magnitude. This is particularly useful for sputtering carbon, which has a very small yield. It is also useful for other light atoms, such as Al or Si. Although there are no commercial applications for this technology yet, it creates an interesting capability for selectively increasing the sputter yield.

5.I.4.2 Ionized Sputter Deposition

Sputter deposition consists almost entirely of the deposition of neutral particles. In most system geometries, this results in a wide angular distribution of the arriving particles: good for coating over steps and bumps on the surface, but poor for filling up deep- or high-aspect ratio features. A recent development using the deposition from mostly ions overcomes this limitation. By combining a dense, inert gas plasma with a sputtering source, a large fraction of the sputtered flux can be ionized in flight [16–18]. A simple bias on the sample then results in an electrical acceleration of the sputtered ion species to the surface with controlled energy and direction. This technique has been developed primarily for the semiconductor industry, but other applications, primarily in the area of low temperature reactive sputter deposition, are being developed.

REFERENCES

1. P. Sigmund, *Phys. Rev.*, **184** (1969) 383; **187** (1969) 768.
2. P. C. Zalm, *Surface and Interface Analysis*, **11** (1988) 1.
3. D. Hoffman, *J. Vac. Sci. Technol.*, **8** (1990) 3707.
4. J. L. Cecchi, in *Handbook of Plasma Processing Technology*, edited by S. M. Rossnagel, J. J. Cuomo, and W. D. Westwood (Noyes, Park Ridge, NJ, 1989), chap. 2.
5. H. R. Kaufman and R. S. Robinson, in *Operation of Broad Beam Ion Sources* (Commonwealth Scientific, Alexandria, VA, 1987).
6. B. Chapman, *Glow Discharge Processes* (Wiley, New York, 1980).
7. S. M. Rossnagel and H. R. Kaufman, *J. Vac. Sci. Technol.*, **A5** (1987) 88–91.
8. J. A. Thornton and A. S. Penfold, in *Thin Film Processes*, edited by J. L. Vossen and W. Kern (Academic Press, New York, 1978).
9. S. M. Rossnagel and H. R. Kaufman, *J. Vac. Sci. Technol.*, **A5** (1987) 2276–2279.
10. S. M. Rossnagel, *J. Vac. Sci. Technol.*, **A6** (1988) 3049–3054.
11. W. H. Class, *Thin Solid Films*, **107** (1983) 379–385.
12. Darryl Restaino, IBM East Fishkill, private communication (January 1996).
13. S. M. Rossnagel, *J. Vac. Sci. Technol.*, **A6** (1988) 19–24.
14. B. Windows and N. Savides, *J. Vac. Sci. Technol.*, **A4** (1986) 196.
15. S. Berg, I. V. Katardjiev, C. Nender, and P. Carlsson, *Thin Solid Films*, **241** (1994) 1–8.
16. M. S. Barnes, J. C. Forster, and J. H. Keller, U.S. Patent 5,178,739 (January 12, 1993).
17. S. M. Rossnagel and J. Hopwood, *Applied Physics Letters*, **63** (1993) 3285–3288.
18. S. M. Rossnagel and J. Hopwood, *J. Vac. Sci. Technol.*, **B12** (1994) 449–453.

CHAPTER 5.2

Plasma Etching

V. Patel, Ph.D.
Sarnoff Corporation

5.2.1

INTRODUCTION

The concept of using plasma processes for etching materials was first applied to semiconductor processing in the 1960s for removing organic photoresist materials. Over the years, plasma etch processes have replaced most selective wet etching processes in semiconductor processing because of its environment friendliness, cleanliness, uniformity, ease of automation and ability to allow faithful transfer of submicron circuit patterns from the photolithographic masks to the surface of semiconductor wafers. Today, plasma etch processes are critical to the fabrication of semiconductor integrated circuits. However, the development of plasma etch processes has progressed on empirical grounds without much scientific foundation, because of intense production demand and the complexity involved.

This chapter is devoted to the plasma etch technology with focus on various reactor configurations. In addition, some of the basic plasma concepts related to etching and the basic plasma etching requirements are also reviewed.

5.2.2

REVIEW OF PLASMA CONCEPTS APPLICABLE TO ETCHING REACTORS

5.2.2.1 Basic Physical Phenomena Occurring in RF Discharges

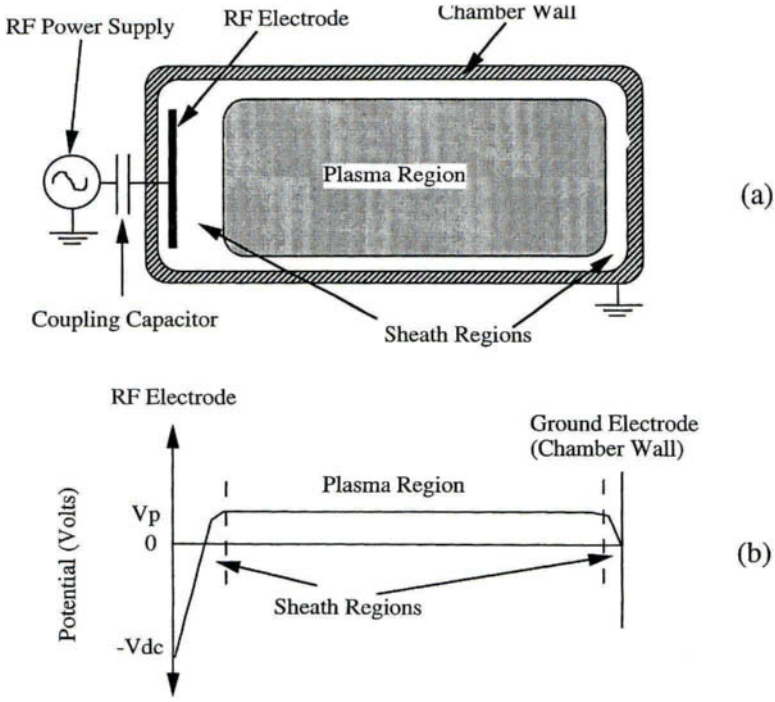
In plasma etching, a glow discharge is used to create chemically reactive species from the feed gas that reacts with the material to be etched and creates volatile re-

ISBN 0-12-325065-7
$25.00

action by-products. A glow discharge is a plasma (defined as a partially ionized gas composed of ions, electrons, and a variety of neutral species) that exists in the pressure range of interest for plasma etching (0.1 mtorr to 10 torr) and contains approximately equal concentrations of positively charged particles (positive ions) and negatively charged particles (electrons and negative ions). A particular useful aspect of plasma is its ability to enable high-temperature-type chemical reactions to be performed at low substrate temperatures.

A simplified configuration of a basic plasma reactor used for etching application is illustrated in Figure 1(a). The semiconductor substrate(s) is normally placed on the electrode to which the RF power is applied. The chamber walls are kept at the ground potential and act as a grounded electrode in many configurations. A radio frequency (RF) field of sufficient magnitude is applied across the electrodes (typically through a coupling capacitor) to cause the feed gas to break down and become ionized. A central bright region is created that is known as the plasma region. The plasma region maintains charge neutrality (equal number of positive and negative ionic species) and is equipotential (no electric field in this region).

Fig. 1.



A simplified configuration of a typical plasma etch reactor in (a) and a typical time-averaged potential profile across the reactor in (b).

The potential of this region is known as plasma potential (V_p) and is the highest potential in the system. The region that separates the plasma region from the electrodes (also chamber walls) is known as sheath or "dark space." Almost all potential drop occurs in this region. In capacitively coupled systems, a negative DC potential (V_{DC}) is induced on the powered electrode in order to balance the time-averaged electron and ion current on electrodes. A simplified diagram of the time-averaged potential profile across the reactor is shown in Figure 1(b).

In a plasma etching apparatus, the plasma is initiated by accelerating some free electrons within the chamber volume by the means of applied electric or magnetic field. The feed gas is weakly ionized as a result of inelastic collisions between neutrals and energetic electrons. Electrons have smaller mass and can respond to the RF field. Electrons can gain energy directly from acceleration in the applied RF field and continue to gain energy through elastic collisions until an inelastic collision occurs. Electron impact ionization produces electron–ion pairs that sustain the discharge against electron loss, which is normally dominated by diffusion from the plasma volume to the electrodes and the reactor walls. Electrons have sufficiently high energy and mobility to escape the plasma volume. However, the coulombic forces between the electrons and positive ions create a field, referred to as the "ambipolar field," resulting in a self-limiting process and thereby maintaining charge neutrality in the plasma. This ambipolar field (assuming electropositive plasma with negligible amount of negative ions) causes the positive ions to diffuse much faster than the free diffusion condition and at the same time slows the diffusion of the electrons.

Several gas-phase processes — such as impact ionization, dissociative ionization, dissociation, recombination, relaxation, electron attachment, and resonance charge transfer — also occur in the plasma. The plasma chemistry responsible for creating reactive radicals and ions through these processes is largely determined by the electron energy distribution function (EEDF), because it is a quantity on which the rates of the electron impact processes depend [1]. Acceleration of secondary electrons emitted from the wafer/electrode surface by the sheath and energy gained by electrons in the oscillating sheath edge by so-called surf-riding mechanism can also play significant role in determining the exact nature of EEDF [2]. The steady-state electron energy distribution is a balance between energy gain and energy loss as a result of inelastic collisions or in certain cases electron beam plasma interaction.

The extensive use of plasmas to activate etch processes derives from two major features of low-temperature nonequilibrium discharges. The first of these is the existence of energetic electrons with average energies in the range of 2 to 10 eV ($\sim 10^4$ to 10^5 Kelvin temperature) in the plasma volume. These electrons break bonds to form chemically active etchant species or their precursors. The electrons are also responsible for ionization, which sustains the discharge and creates ions. These ions are often essential to the etch process. The second important feature of plasma for etching applications is the acceleration of ions at the plasma boundary

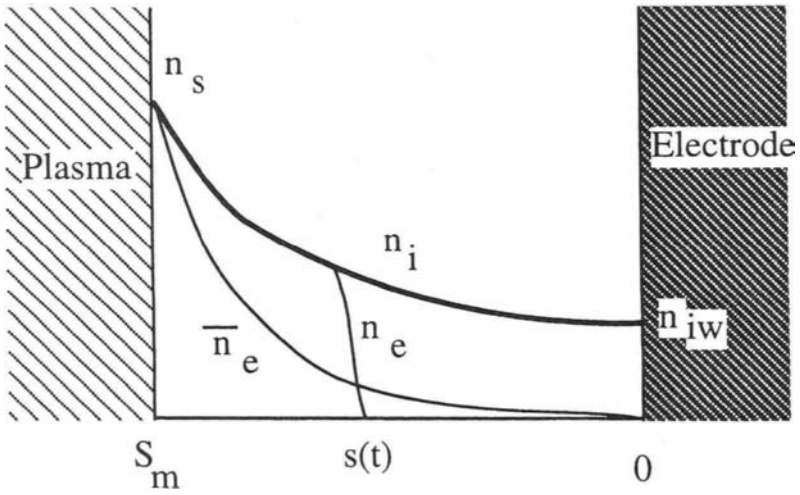
known as sheath. In many configurations, the ions are accelerated by the electric field in the sheath region between the plasma boundary and wafer. This field can accelerate ions, normal to the wafer surface, with typical energies in the range of 50 to 1000 eV. The ion bombardment often results in mechanisms that allow lithographic patterns to be etched anisotropically with little or no lateral removal of material. This feature is essential to realization of increasingly high device density found in modern integrated circuits, and is perhaps the major reason why plasma-activated etching is used so extensively. It should be noted here that ions are rarely the etchant, and neutrals are responsible for almost all reactive etching [3].

SHEATH FORMATION

An unique feature of plasma is its ability to shield out internal charge concentrations from externally applied electric fields. Sheaths (defined as regions in which the quasi-neutrality condition does not hold) are formed at all grounded and powered surfaces representing boundaries for the plasma. The sheaths insulate the plasma from electron loss. In RF discharges, the dynamics of the sheaths are more complex than for DC discharges and generally are not understood. The earliest studies on the RF sheaths were made by Butler and Kino [4] and since their pioneering work, numerous papers have been published outlining improved mathematical treatments [5,6].

An approximate description of a RF sheath based on physical reasoning is shown in Figure 2. The ions cannot respond to the instantaneous field. Instead,

Fig. 2.



The sheath structure showing the positive ion density (n_i), the time-averaged electron density ($\bar{n}_e$) and the instantaneous electron density (n_e) at one time in an RF cycle.

they move through a quasi-neutral presheath and acquire a directed average velocity that is of the order of ion sound speed (see the next section) at the sheath edge S_m where the electron density is equal to the ion density at all times. The region between the wall and the S_m is called the ion sheath, because it is a region of net positive charge on time averaged basis, although this does not hold during the entire RF cycle [7].

As positive ions move to the wall, they gain momentum in the average sheath field and conservation of ion flux causes the ion density to fall toward the electrode/wall to a value, n_{iw}, at the wall. Since the response of an ion to the RF field can be neglected, the ion density in the sheath is stationary. In contrast, the electrons respond to the changing voltage across the sheath, and their motion can be characterized most simply by the assumption of a moving sheath edge. The density profile of the electron at the moving sheath is a steplike profile [8]. Since the average ion and electron currents to the wall in capacitively coupled systems must balance, the sheath collapses for a small part of the RF cycle (when the RF voltage is maximum) enabling the electrons to reach the electrode. After the field reversal of the applied RF voltage at the electrode, the electron sheath edge moves away from the electrode. Thus, the RF field causes the sheaths to expand and decay, modulating the sheath voltage and length. An electron in the glow near the sheath boundary comes under the influence of the repulsive field at the edge of the moving sheath and can be reflected back into the plasma. This phenomenon is often termed as the "surf-riding" mechanism [2] or stochastic heating of electrons [9]. Electron reflection occurs at both electrode sheaths, and the electron velocity after reflection depends on the sheath thickness and the applied voltage.

BOHM CRITERION FOR EXISTENCE OF A PRESHEATH

A smooth transition region exists between the quasi-neutral plasma and the sheath and is known as the presheath region. This region has a relatively weak electric field, compared to that of the sheath, that accelerates the ions from the quasi-neutral plasma toward the sheath. This implies that the ions arriving at the sheath will acquire a velocity, which is of the order of the ion acoustic speed in the presheath region. The existence of this enhanced velocity at the sheath edge and the resulting criterion requiring the existence of a presheath was first demonstrated by Bohm [10] based on energy and momentum conservation arguments. The necessity for an ion-accelerating presheath has become known as the *Bohm sheath criterion*. Chen [11] demonstrated the physical significance of this criterion by showing that the acceleration of ions and repulsion of electrons in a typical sheath must be such that the ion density decreases less rapidly than the electron density across the sheath. In these classical theories, an electropositive gas was normally considered with equal concentration of electrons and positive ions in the bulk of the plasma. The existence of negative ions has generally been ignored.

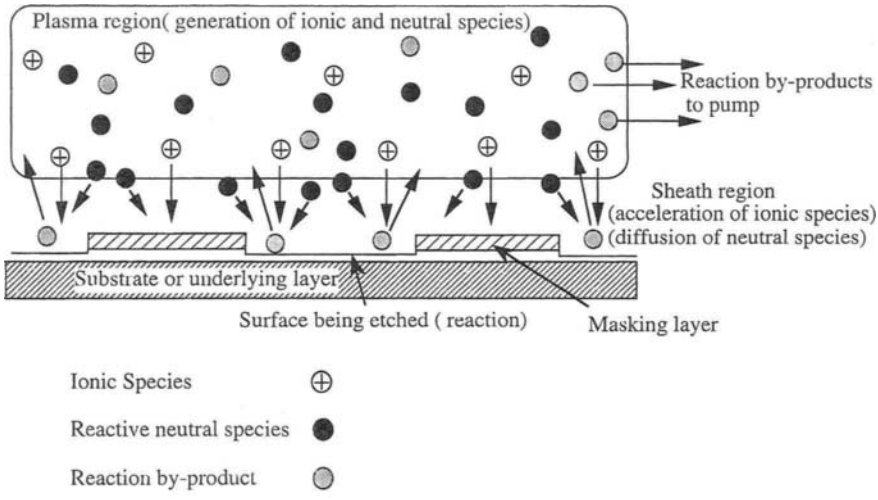
For discharges containing electronegative gases, as is the case for most plasma processing applications, the negative ions are present in copious amounts and their concentration can exceed that of electrons [12]. This causes significant deviation in the mathematical relations derived for electropositive discharges. The concept of sheath, however, is not significantly changed for electronegative plasma, because negative ions experience only the time-averaged field and are repelled from the sheath field. The negative ions do not play a significant role in the sheath dynamics. However, the presheath region in the electronegative plasma can be significantly different from the presheath region in electropositive plasma because of the influence of negative ions on the presheath electric field. Braithwaite [13] showed that as the ratio of negative ion concentration to electron concentration increases, the potential variation in the presheath becomes more rapid. For a critical value of the ratio, an abrupt potential drop occurs between the sheath edge and the plasma. This is accomplished by the formation of a "double layer" at the plasma–sheath boundary [14]. This double layer plays an important part in many plasma processing applications.

Negative ions are reflected by the sheath, and their generation and loss is dominated by volume processes such as attachment balanced by detachment and ion–ion recombination rather than by recombination at the wall. In an electronegative plasma, the electron density (n_e) can be much lower than the positive ion density (n_i), and thus a much larger fraction of the random electron flux incident on the sheath must overcome the sheath barrier in order to equalize the positive ion flux to the electrode. This is ensured by a much lower sheath potential, and hence induced DC bias.

5.2.3 BASIC PLASMA ETCHING REQUIREMENTS

Various physical and chemical processes occurring during plasma etching are extremely complex and not well understood. However, the basic processes responsible for plasma etching can be described into five basic steps as shown in Figure 3. These steps are as follows: (1) generation of reactive neutral and ionic species in plasma; (2) diffusion of the reactive neutral species and acceleration of positive ions through the sheath to the surface being etched; (3) reaction of these species on the surface resulting in a volatile by-product; (4) desorption and diffusion of the by-product from the surface into bulk plasma; and (5) removal of the by-products from the etching chamber by vacuum pumps. No etching takes place if any of these steps fails to occur. Note here that many reactive species can react readily with a solid surface, but unless the by-product has a reasonable vapor pressure, the etching cycle ceases. Also, the chemical reaction on the surface can be strongly influenced by the mass, energy, and flux of the ions impinging on the surface.

Fig. 3.



Primary processes occurring in a plasma etch process.

The layers to be etched for fabrication of semiconductor integrated circuits can be defined in three broad categories: conductors, dielectrics, and masking layers. Doped polysilicon, silicides, and various metal layers such as aluminum, titanium, titanium nitride, titanium tungsten, and indium tin oxide (ITO) are very commonly used conductor layers. Silicon oxide and silicon nitride are very commonly used dielectric layers. The masking layer is often a photoresist layer, which is patterned using standard photolithographic techniques.

For plasma etching of these layers, the feed gas is selected so as to generate species that react chemically with the material to be etched, and whose reaction product with the etched material is volatile. At the same time, removal of the underlying or masking layers may be negligible. A large variety of fluorine-, chlorine-, bromine-, and oxygen-based etching plasmas with a profusion of gas additives are used in the semiconductor industry. In most cases, a specific gas mixture is based on a great deal of empirical evidence obtained for a particular application rather than real fundamental understanding of the relevant plasma chemistry. Nevertheless, certain basic insights have proven to be helpful in formulating gas mixtures. The commonly used gas chemistries for etching various materials are listed in Table 1.

A large variety of other primary etchant and additive (such as polymer-forming gases) gas mixtures can also be used, and the specific formulation can differ from one user or tool to another [15].

For any particular etch process, a number of requirements must be met such as throughput (high etch rate), high selectivity, uniformity of etch rate, control of the etched profile (degree of anisotropy), control of critical dimensions (CD control),

Table I

Trends in Plasma Etch Chemistry

Material Being Etched	Chemistry
Poly-Si	Cl_2, HBr, SF_6/O_2, CF_4/O_2, BCl_3, $SiCl_4$
Single-crystal silicon	Cl_2, SF_6, HBr, $SiCl_4$
Al, Al-Si, Al-Si-Cu	Cl_2, BCl_3, $SiCl_4$, $CHCl_3$ (N_2 or He diligent)
TiW, TiN, Ti	CF_4, SF_6, Cl_2
WSi, TiSi, CoSi	CF_4, SF_6, Cl_2
SiO_2	CHF_3, C2F6, CF_4/H_2
Si_3N_4	CF_4/O_2, SF6/O_2, CHF_3
GaAs	$SiCl_4/SF_6$, $SiCl_4/NF_3$, $SiCl_4/CF_4$
InP	CH_4/H_2, HI

low damage to the devices, low particulate or contamination, residue-free surface after etch, process repeatability, and safety.

5.2.3.1 High Etch Rate

High etch rate is usually needed to keep the throughput (number of wafers etched in a unit time) of the plasma etch system high. There is usually a tradeoff between etch rate and other parameters such as selectivity and damage. For example, higher etch rate can be obtained by increasing the pressure so that more free radicals are available for etch. However, high pressure may cause more undercut or polymer formation. Higher power generally also results in higher etch rate, but it can create more damage to the substrate and may result in a less selective process.

5.2.3.2 Etch Profile Control or Anisotropy

Controlling the shape and dimension of the etched wall is crucial for realizing very highly dense submicron circuit patterns. The term *anisotropic etching* generally refers to etching a feature with vertical side walls, with no lateral removal. The degree of anisotropy (Figure 4) can be defined by relating the etching in vertical and lateral directions as

$$A = 1 - l/v \tag{1}$$

where v is the thickness of the layer etched in the vertical dimension
l is the lateral distance etched underneath the mask

Majority of the etch processes used today require the patterns to be etched to have much smaller dimensions than the thickness of the film. This requirement necessitates the etch process to be anisotropic. Anisotropy can be induced by en-

Fig. 4.

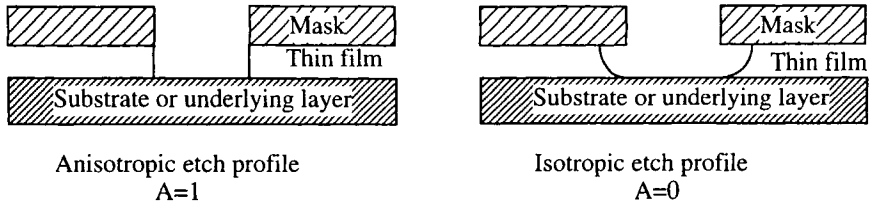


Illustration of anisotropic and isotropic etch profiles.

ergetic ion-bombardment-enhanced reaction, by inhibitor layer formation on the side walls, by lowering the pressure, or by reducing the temperature of the wafer.

ENERGETIC ION BOMBARDMENT ETCHING

A good example of ion-bombardment-enhanced etching is the etching of Si in chlorine. An undoped single-crystal Si surface is not etched by Cl_2 or Cl atoms at room temperature [16]. When the surface is simultaneously exposed to high-energy ion flux, silicon chlorides are formed in a rapid reaction. Normally, higher the energy of the ions bombarding the surface, the etched profiles are more anisotropic. However, higher ion bombardment energy can result in poor selectivity and damage.

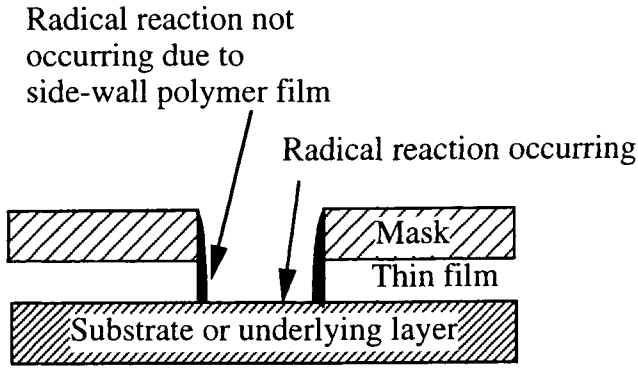
INHIBITOR ION-ENHANCED ETCHING

The inhibitor ion-enhanced etching requires two conceptually different species: etchants and inhibitors. The inhibitor species form a thin film on the surfaces being etched. Only the bottom surfaces experience energetic ion bombardment, due to directionality of ions accelerated by the sheath. The film is continuously removed from the bottom surfaces by sputtering, and etching takes place. Vertical side walls experience little or no ion bombardment, and no etching takes place on vertical side walls. The resultant etch profile is anisotropic. In fluorocarbon plasmas, CF_2 radicals can be used to form thin polymeric inhibitor films [17]. The concept of inhibitor-ion-enhanced etching or "side wall passivation" for achieving high anisotropy is illustrated in Figure 5.

LOW-PRESSURE ETCHING

Lowering the pressure of the reactor, in general, improves the degree of anisotropy. This is due to the fact that the mean free path increases with decreasing pressure. As the mean free path is increased, the scattering in ion directionality and energy in the sheath region is minimized and better anisotropy is achieved.

Fig. 5.

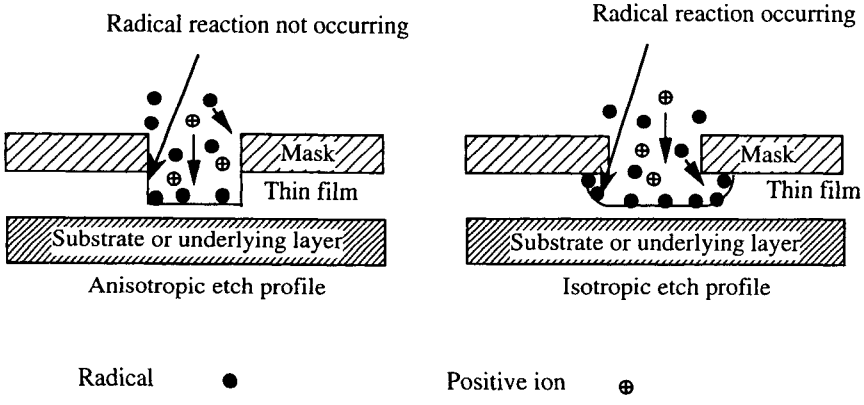


Anisotropy by polymeric inhibitor films.

EFFECT OF WAFER TEMPERATURE

Any basic etch process can be categorized in two simple components: chemical reaction by reactive radicals (isotropic) and ion-assisted reaction (anisotropic). By lowering the temperature of the substrate and hence the surface on which reaction takes place, the chemical component of etch can be minimized. Only the bottom of the features that experience directional ion flux are etched. With little or no ion bombardment on the sidewall, the resultant lateral etch rate is very slow, because the chemical etch rate is reduced. The concept of achieving anisotropy by lowering substrate temperature is illustrated in Figure 6.

Fig. 6.



Anisotropy at lower temperature.

The concept of low-temperature etch is similar to the side wall inhibitor film approach. However, the chemistries or the feed gas mixture that can be used are much simpler. The disadvantages of the low-temperature etch are lower etch rates, condensation on the wafer, and complexity involved in the design and operation of the wafer holder for maintaining high throughput of the system.

5.2.3.3 Selectivity

Selectivity is defined as the etch rate ratio of the material to be etched to the overlaying masking layer or the underlying material. Having a process with high selectivity is critical in cases where a significant overetch (additional etch time after the nominal thickness of the layer is etched away) is required to remove all materials from all topological features. For example, during formation of contacts to the substrate and polysilicon layers, the glass layer can have different thickness in different regions in a single masking step. The etch process must be able to etch the thickest layer without damaging the underlying layers. High selectivity is also critical during etching of gate stacks where the stopping layer may be a very thin oxide (~100Å) [18].

In an ideal chemical etching, the only function of the plasma is to maintain a supply of gaseous etchant species. Chemical etching is the most selective mechanism, because unwanted reactions will not take place at all when the thermodynamics are unfavorable. In reality, a combination of chemical and physical (such as ion-bombardment-enhanced) etch mechanisms is used to obtain the desirable etch rate, selectivity, and anisotropy.

In many etch processes, higher selectivity is obtained by adding nonetchant gases to the basic etchant chemistry. A classical example of using gas additives to obtain desired selectivity is the addition of O_2 or H_2 to the CF_4 feed gas. The addition of small amounts of O_2 to CF_4 plasma is known to increase the F radical concentration in the discharge, dramatically increasing both the etch rate of Si etching and selectivity of Si to oxide etching. This is caused by the reaction of oxygen with CF_x radicals, forming CO, CO_2, and COF_2 and producing more free fluorine by reducing the recombination of F atoms with CF_3 [19].

When H_2 is added to CF_4, the Si etch rate decreases monotonically as the percentage of H_2 is raised, because of formation of HF by gas-phase reaction between atomic H and atomic F [20]. Also, adding H_2 increases selective deposition of polymer on Si surface. In this case, the etch rate of the underlying layer is determined by a balance between polymer deposition and etching. Similar results can also be obtained by using C_2F_6/H_2 or CHF_3 gas chemistries, and these gases are widely used to etch SiO_2 more selective to Si.

Chlorine- and bromine-based chemistries are also widely used for Si and Al etching. Chlorine and bromine atoms have relatively high selectivity to the under-

lying oxide layers; chlorine is more widely used because it has much higher vapor pressure, is less corrosive, and tends to form somewhat less toxic by-products. O_2 is widely used for photoresist stripping without damaging underlying layers and for cleaning polymers from the reactor walls.

5.2.3.4 Etch Uniformity

Good etch uniformity (across the wafer and also from wafer to wafer) is a very desired feature for any etch process. Uniformity problems often worsen as wafer size increases or new high-etch-rate processes are developed. Nonuniform etching necessitates extensive overetching and creates a stringent demand for a high-selectivity process.

Nonuniformity can be caused by depletion of etchant concentration or gradients in ion bombardment flux and energy [21]. For an etch process to be uniform, diffusion of the reactant species and the ion flux to the wafer surface must be uniform. The removal of the volatile by-product by the pumping system also must be uniform. In many reactor configurations, nonuniform etching tends to develop at the boundary of the wafer. This is commonly known as the "bull's-eye" clearing pattern, where the etch rate monotonically decreases from the wafer periphery to its center. The other nonuniformity-related common problem is known as microloading or aspect ratio dependent etching (ARDE) [22], which results from variation of etch rates across the wafer due to the feature size and density. Normally, the smaller and denser the feature becomes, the slower it etches. This phenomenon arises from the fact that the mean free path of the etchant species can be comparable to the geometry being etched. The flux of etchant species to the surface and diffusion of the volatile by-product back into plasma can be disturbed by the high-aspect-ratio (ratio of the width to the depth) features.

In practice, pressure, power, interelectrode distance, and gas flow are interactively adjusted for a specific application until an acceptable uniformity is achieved. This empirical approach is taken because the underlying mechanisms responsible for nonuniform etching are not known quantitatively.

Another term that is used very often in production to assess uniformity is *critical dimension* or *CD control*. In production, the line widths of etched profiles are measured routinely, and any significant statistical variations in critical dimension are often related to a machine or a process problem. Control of line width is crucial for high-yield manufacturing. For example, any major deviations in line width of a etched gate layer can cause significant variations of the leakage or drive currents of the metal-oxide-semiconductor (MOS) transistors.

Also, note that uniform control of wafer temperature is required for many applications for maintaining uniformity and reproducibility of the etch rate and etch profiles [23,24]. In many reactor configurations, this is achieved by clamping the

wafer to the chuck and by applying He gas on the backside of the wafer to provide thermal conductivity between the wafer and the temperature-controlled chuck.

5.2.3.5 Damage

For high-yield integrated-circuit manufacturing, it is desirable to have an etch process that causes no or minimal damage to the wafers. In plasma etch processes, positive ions from plasma volume are accelerated by the sheath voltage. These ions can impinge on the surface of the semiconductor wafer with energies in excess of a few hundred eVs and can cause physical damage by sputtering. In general, the bias on the wafer is optimized to obtain desired etch results with minimal damage. Also, the high level of UV radiation present in the plasma can generate defects in the semiconductor oxides [25].

Various features on semiconductor wafers are usually exposed to plasma directly and collect charges from the plasma. The induced electrostatic stress on isolated features due to any nonuniformity in plasma can also degrade the quality of gate oxide. The stress can cause surface states at SiO_2–Si interface as well as trapped charges in the oxide, decreasing the breakdown voltage and deforming the CV characteristics of the gate oxide [26].

The exact mechanisms responsible for plasma-etch-induced damage are not well understood at present. Special antenna test structures are commonly incorporated on semiconductor wafers to aid process development and to minimize yield loss due to plasma-induced damage [27].

5.2.3.6 Particulate or Contamination Control

Particulate or contamination is a major concern in modern high-density integrated-circuit fabrication facilities. Particulates can severely affect the yield and reliability of the devices. With increasing size of dies and wafers, this can translate into a major financial concern. Plasma processes are among the "dirtiest" processes among the various processes (such as implantation, diffusion, photolithography, and so on) used for semiconductor circuit fabrication [28]. In plasma etching apparatus, the by-product films deposited on chamber wall and electrodes can fracture and can generate particulate. Often, a plasma clean step is used between processing of every wafer, or of a selected number of wafers, to minimize particle generation. Particles can also be formed by gas phase reactions within a plasma volume [29]. Particles present in plasma can acquire negative charge from electrons and can be suspended at the sheath boundary by electrostatic repulsion. As the discharge is turned off, these particles can fall on the wafer. Sputtering of materials from the reactor walls by the discharge is another concern.

In recent years, the reactor designs have been significantly modified to minimize particles. The reactor materials are selected to be inert to process reactants and with high resistance to sputtering. Particle traps such as the grooved electrode design are also being incorporated in some reactor designs to electrostatically trap the charged particles [30]. "Cleaner" plasma chemistries are also preferred, to minimize particle formation.

Another source of contamination is the polymers formed on the side walls of the features on the wafers itself. In many cases, the wafers are subjected to a wet chemical clean followed by a plasma etch step to remove these polymers.

During the etching of Al-based metallization, the chlorine residue from the etch process can cause serious postetch corrosion and can affect yield, failure rate, device performance, and its long-term reliability. Many techniques, such as water or solvent rinse [31], passivation using fluorocarbon plasma [32], annealing or oxidation [33], and nitric acid treatment [34] are generally used to reduce postetch corrosion.

5.2.4 PLASMA DIAGNOSTICS

Over the years, various plasma diagnostic techniques have been used to elucidate some fundamental properties of the plasma processes. Some of these basic techniques are summarized next.

IN SITU ETCH RATE MEASUREMENTS

The two major techniques used for *in situ* measurements of etch rates are laser reflection interferometry and ellipsometry. In laser reflection interferometry, a laser beam is reflected back from an unpatterned area on the wafer. The interference pattern originating from the thickness change of the layer being etched is used to measure the etch rate [35]. Standard ellipsometric techniques have also been applied to measure *in situ* the etch rate of the film being etched.

PLASMA IMPEDANCE MEASUREMENTS

Radio frequency current and voltage probes are often incorporated on various reactors for measuring plasma impedance and true power dissipated [36,37]. Since these probes are installed externally, they are nonevasive, making the measurements useful for process monitoring and tool fingerprinting. For inferring the plasma characteristics from the externally measured parameters, it is necessary to account for stray impedances between the point of measurement and the plasma–electrode interface.

OPTICAL EMISSION SPECTROSCOPY

Emission spectroscopy has played an important role in plasma diagnosis [38,39]. In optical emission spectroscopy, the optical spectrum emitted by etch plasma is analyzed. This is generally accomplished by a detector and an optical system consisting of a monocrometer. Various species present in the plasma can be identified by analyzing the spectrum and the concentration of various species can be inferred qualitatively.

LASER ABSORPTION SPECTROSCOPY

Various laser-based absorption techniques have been used to determine concentrations of various species in the plasma [40]. This is achieved by using wavelength-tunable lasers and by monitoring absorptance at a characteristic wavelength of a specific molecule. Various frequency modulation schemes have also been applied to enhance the detection limit of this technique [41].

MASS AND ENERGY ANALYZER

Mass and energy sampling of glow discharges has yielded a rich source of information for developing some fundamental understanding of overall discharge behavior [42,43]. A mass/energy analyzer normally samples the discharge through a small orifice and provides information about mass and energy of various ionic and neutral species in the plasma. These analyzers are normally differentially pumped when used in conjunction with process plasma.

MICROWAVE INTERFEROMETRY

Microwave interferometry has been used to infer electron properties in process plasma by measuring interference to the applied microwaves propagating through the plasma. In this technique, the applied frequency is selected to be higher than the electron plasma frequency, and the resulting phase shift, which is proportional to the average electron density, the path length, and the wavelength, is measured [44].

LANGMUIR PROBES

Electrical measurements using a Langmuir probe is a powerful and experimentally simple means of determining key internal discharge parameters, such as charged particle concentrations, plasma potential, and electron energy distribution function (EEDF). Since the pioneering work of Irving Langmuir nearly seventy years ago [45], numerous papers and excellent reviews of the subject have

been published [46,47]. Basically, an electrical probe consisting of one or more small metallic electrodes is inserted into the plasma. The I-V characteristic of the probe is then used to deduce important plasma parameters.

LASER-INDUCED FLUORESCENCE

Laser-induced fluorescence has been used by many researchers to yield information on various species present in the plasma [14]. Normally, a laser is used to excite the species, and the resultant fluorescence is detected. Measurement of fluorescence spectrum is used to identify laser-excited species. The state, temperature, and density of these excited species can be inferred from the spectrum.

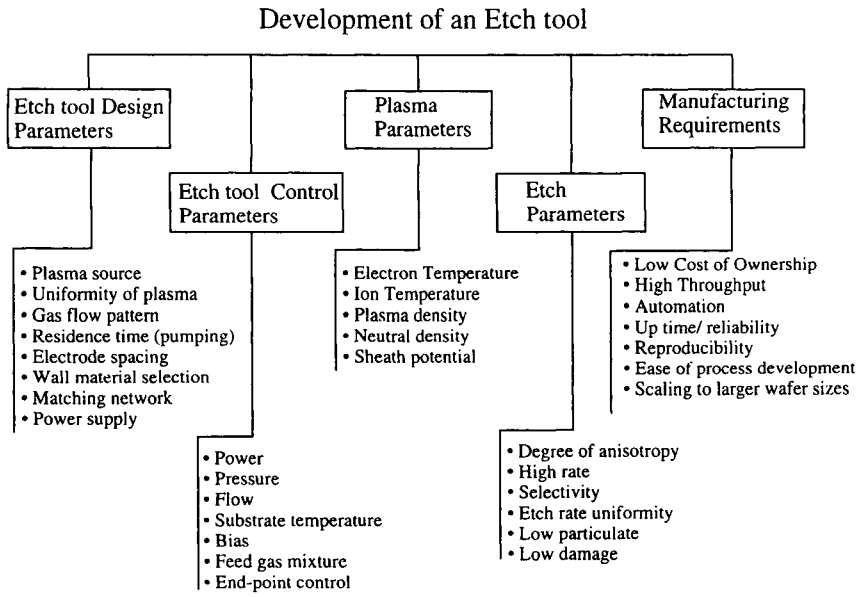
5.2.5 BASIC PLASMA ETCH REACTORS

5.2.5.1 Introduction

The ultimate goal for an etch tool would be to translate the etch requirements into proper settings for the tool control parameters—e.g., power, pressure, feed gas mixture, flow, bias voltages, etc.—and then proceed to etch with closed-loop control. Plasma-activated processes are, however, far more complex, involving a myriad of nonequilibrium mechanisms. The resultant etching rates and profiles depend on the feed gas, gas flow, power, pressure, substrate temperature, reactor geometry, and many other parameters in a fairly complex and interrelated manner. In a practical etching process, the end results are optimized through an artful balance of these parameters. The underlying mechanisms in plasma processes are not well understood, and there is much less understanding of how the tool control parameters relate to the plasma discharge characteristics that actually determine the etch process. These effects further add to the complexity in determining the connections among the tool control parameters, the plasma discharge parameters, and the etch characteristics. There is, at present, only an empirical understanding of how to set the tool control parameters to achieve a given set of etch characteristics. The development and optimization of various reactor configurations have been based on this empirical understanding. The basic considerations and interactions for development of an etch tool are illustrated in Figure 7.

The requirement for larger wafers and thinner oxides for modern high-density high-yield high-throughput integrated-circuit manufacturing is also putting stringent demands on manufacturing requirements such as cost of ownership (COO), up-time, reliability, process control, ease of process development, and automa-

Fig. 7.



Interactive parametric considerations for an etch tool.

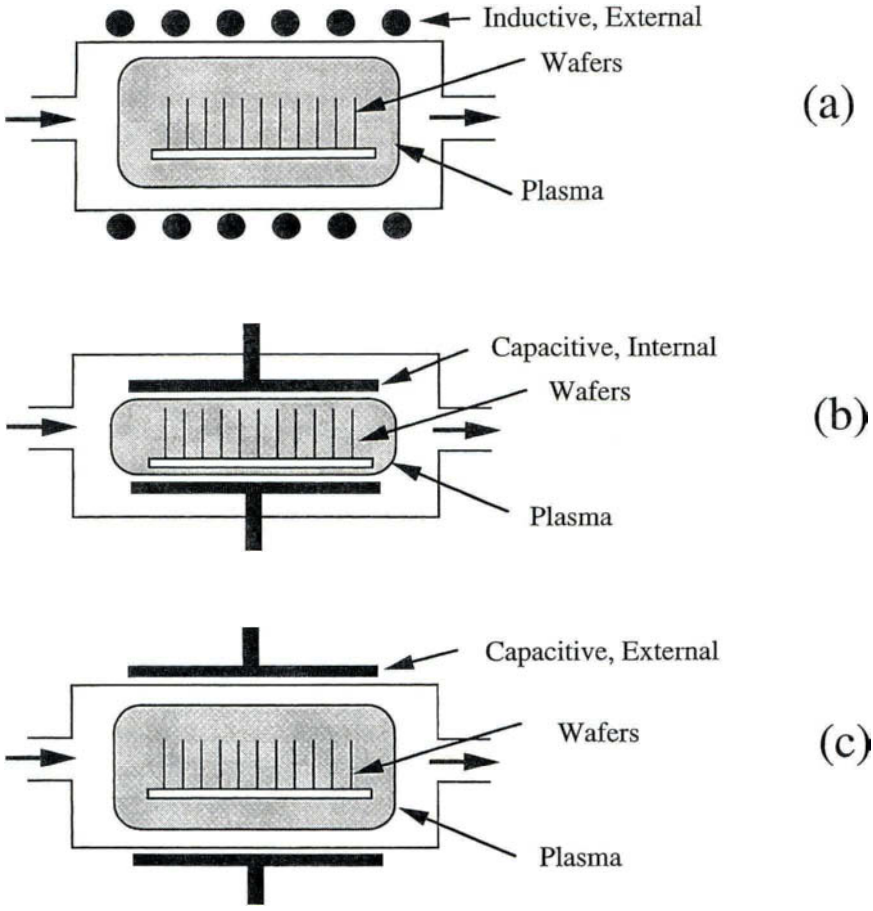
tion. A reactor configuration designed for production is required to meet all these manufacturing demands.

Production etching systems can have a variety of configurations depending on the parameters of a process that must be controlled as well as on the specific applications of the system. The configurations can vary in terms of excitation frequency (from few kHz to few gigahertz), operating pressure, and electrode arrangements (internal, external, capacitive, inductive), as well as spacing and plasma confinement schemes (magnetic or nonmagnetic).

5.2.5.2 Barrel Reactors

The most basic type of reactor is a barrel reactor. It is characterized by a high operating pressure (>1 torr) and the placement of wafers in the glow discharge (usually on a quartz holder). This type of reactor is designed to etch a batch of wafers (multiple wafers) simultaneously. The RF power is normally applied to the reactor via internal or external capacitive or inductive coupling. Various configurations of the barrel reactor are illustrated in Figure 8. This type of reactor was mainly used for plasma ashing of resist and was very widely used in the 1970s. The major limitation of this type of reactor operating at high pressure without any

Fig. 8.



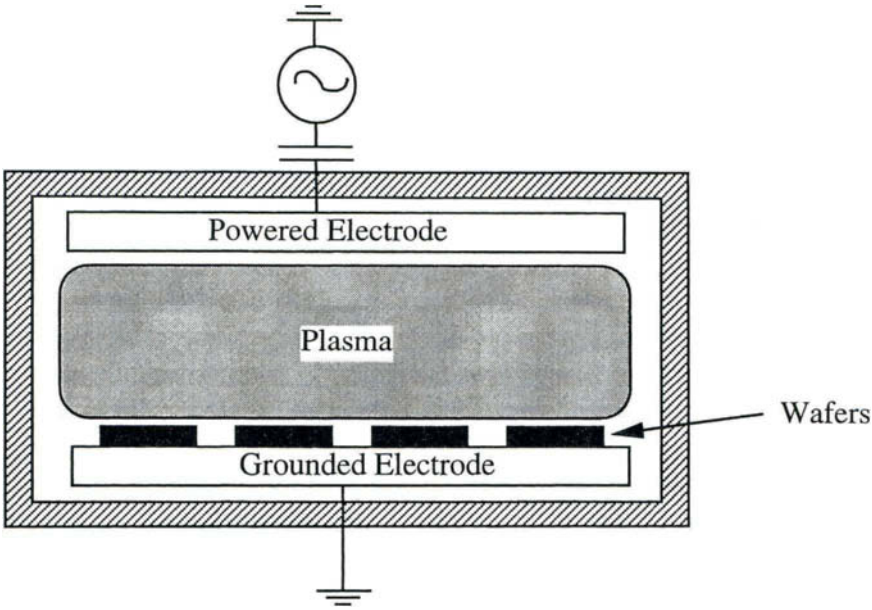
Basic barrel etch reactor configurations.

applied bias on the wafers was dimensional control. Most of the barrel reactors produce isotropic etch profiles, and their use became limited with shrinkage of the integrated-circuit dimensions.

5.2.5.3 High-Pressure Diodes

Another type of very popular reactor configuration is known as high-pressure planar diodes, parallel-plate diodes or Reinburg reactors [48]. These reactors are characterized by fairly high operating pressure (300 mtorr to 10 torr). The wafers were placed on the grounded electrode, and the RF power was capacitively cou-

Fig. 9.



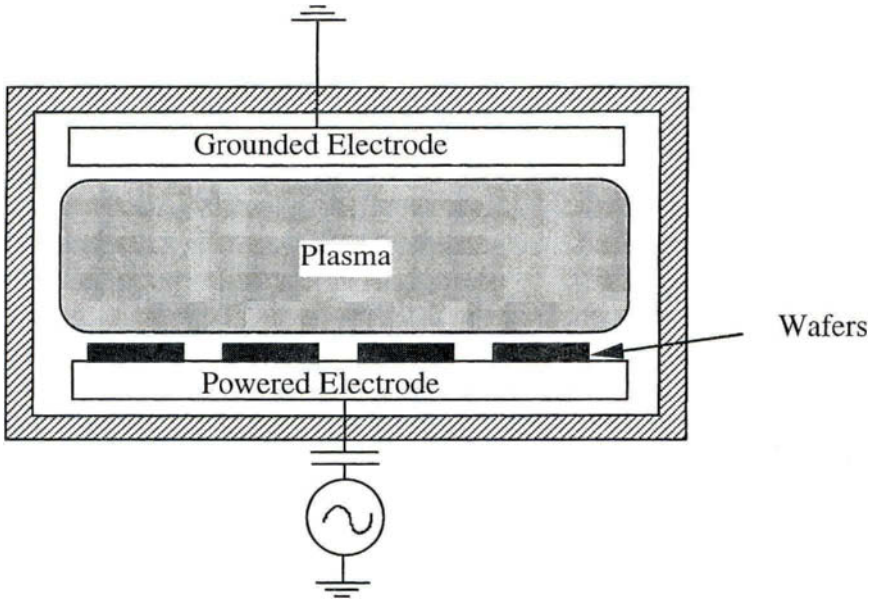
Schematic diagram of a high-pressure diode system.

pled to the other plate of the diode via a matching network. Often the gas flow is through the center of the grounded electrode, to produce a uniform flow pattern. Majority of these configurations were still designed for batch etching. They can produce both directional and anisotropic etch profiles. The schematic diagram of the reactor is shown in Figure 9.

5.2.5.4 Low-Pressure Diodes or Reactive Ion Etchers

In the next generation of plasma etching machines, the role of electrodes was reversed. The wafers were placed on the electrode to which the RF power was applied. These type of systems were known as low-pressure diodes or reactive ion etching (RIE) reactors. The name "reactive ion etching" became very popular despite the fact that the contribution of ions in the chemical etching reaction was negligible. This configuration is characterized by low operating pressure (30 mtorr to 300 mtorr). Low-pressure diode configurations were widely used for etching applications because of their simplicity and ability to direct energetic ions normal to the surface being etched. The etching in this type of reactor is almost always inherently directional. This type of configuration is the most studied configuration to date and was widely used in production till the late 1980s. In one of the com-

Fig. 10.



Schematic diagram of a reactive ion etching (batch) reactor.

monly used configurations, the RF electrode was designed in a hollow hexagon shape, and the reactor was commonly known as the HEX reactor [49]. The RIE diode configuration was developed initially for batch etching applications and was converted to single-wafer etching applications later on. Because this configuration formed the basis for developing the new generation of reactors, it will be useful to review some of the basic concepts of this configuration.

A schematic representation of the configuration is shown in Figure 10.

One electrode of the diode configuration is capacitively coupled to the RF power supply through an impedance-matching network, and the other electrode is grounded. In capacitively coupled schemes, the coupling capacitor forces the time-averaged conduction current at the driven electrode to be zero. The frequency of the RF power supply can range from several kHz to several mHz [50]. The most common commercially used frequency is 13.56 mHz, which is an industry standard reflecting Federal Communication Commission (FCC) regulations. At 13.56 mHz frequency, the ions cannot follow the electric field, because of their heavy mass, and respond only to a time-averaged electric field. The electrons can respond to the instantaneous field and gain sufficient energy to sustain the discharge by volumetric ionization. Secondary electron emission may no longer be the essential process for maintaining the discharge. Unlike DC and low-frequency discharges, continuity of current across the electrode sheaths is ensured by the alternating displacement current.

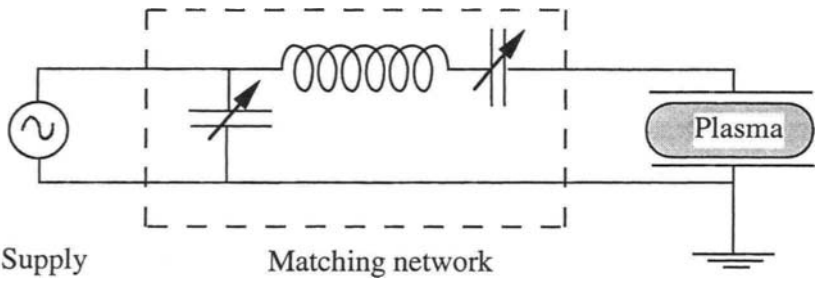
MATCHING NETWORK

Majority of the RF power generators are designed to drive a purely resistive load of 50 ohms. The RF discharge normally has a capacitive component due to sheaths. To match the discharge impedance with the output of the generator for coupling maximum power to the discharge, a matching network is placed very close to the discharge. A typical matching network configuration is shown in Figure 11. The reactive components of the network are tunable and are adjusted until the impedance of the matching network is matched to the plasma impedance to give a 50-ohm termination for the power generator. In this case, the reflected power goes to zero. Normally, the current, voltage, and phase are sensed by sensing elements and feed back to tuning elements for autotuning the network to match the plasma load. The autotuning must be achieved in only a few seconds, because the total time to etch can be less than a minute.

INDUCED DC BIAS

In capacitively coupled RF discharges, the time-averaged conduction current to the powered electrode must be zero, and this requirement is accomplished by formation of a self-induced DC potential. The formation of the induced DC bias, V_{dc}, on the RF electrode can be explained as follows. Total conduction current (of the electrons and ions) to the electrode wall must be zero. However, in an RF sheath, this conduction is applicable only on a time-averaged basis, i.e., averaged over a cycle. Due to the smaller mass of electrons, the electron current to the electrode is much larger than the ion current during an RF cycle if there is no DC bias to the electrode. A negative DC bias is self-induced on the electrode to balance the electron and positive ion current averaged over an RF cycle. The potential drop in the sheath on a time-averaged basis is the difference between the time-averaged plasma potential and the self-induced DC bias on the electrode. The large potential drop in the sheath accelerates ions toward the electrode and decreases the electron current to the electrode. Because of their mass, ions respond

Fig. 11.



A typical RF matching network.

only to the time-averaged sheath field and flow to the electrode at nearly a steady rate. On the other hand, for most of the cycle the wall potential is strongly negative with respect to plasma, because of the self-induced bias, the electrons are repelled from the wall, and negligible current flows to the wall. Only near the peak of the RF cycle, when the RF electrode is the most positive, the sheath for electrons become thin and its potential drop is small, which allows a short burst of electrons to reach the wall. The negative charge carried by this short burst of electrons equals the positive charge of ions flowing to the wall continuously throughout the cycle.

In plasma processing, the DC bias (V_{dc}) is normally used as an energy indicator for ions impinging on the powered electrode. The value of the induced bias depends on the power, pressure, nature of gas, area ratio of electrodes, and electrode surface conditions. In most commercial plasma reactors, V_{dc} is measured at a point as close as possible to the electrode via a chock coil (for attenuating RF voltage) and a low-pass filter.

LIMITATION OF THE DIODE CONFIGURATION

The major limitation of the diode configuration is the strong coupling between the plasma generation and the induced bias. It is very difficult to optimize all the etching characteristics simultaneously using conventional parallel-plate etch-tool configurations. To increase etch rate, for instance, the gas density (pressure) can be raised; but, this may reduce the average ion energy and directionality at the wafer owing to collisions in the sheath and can negatively impact the anisotropy.

The other major limitation of the diode configuration is the operating pressure range. Generally, the performance of parallel-plate diode configuration degrades at lower pressures, that is, below operating pressures of 30 mtorr [51]. The plasma density drops to unacceptable levels as observed by very low etch rates, and the induced bias on the electrodes rises to unacceptable levels, enhancing the possibility of wafer damage and contamination.

5.2.6 ADVANCED PLASMA ETCH REACTORS

5.2.6.1 Basic Requirements for Advanced Plasma Etch Reactors

The major limitation of the batch etching machines were wafers-to-wafer and batch-to-batch uniformity and reproducibility. To achieve better control and reproducibility of the etch process, the newer reactors were designed to etch only one wafer at a time. They are commonly known as single-wafer etchers (SWE). The need for single-wafer etching has forced the industry to develop many new

SWE configurations capable of high-rate etching. Another desirable feature of a plasma etching tool is its ability to independently control many of the etching characteristics. In addition, it should possess, if possible, the simplicity of the parallel-plate configurations and should be easily adaptable to the existing production lines. Beside high-rate etching, uniformity, and reproducibility, the development of the advanced reactor configurations has been driven by the following two requirements: (1) low-pressure operation, and (2) independent control of bias from plasma generation.

NEED FOR LOW-PRESSURE ETCHING

The stringent requirements of anisotropy for submicron circuit fabrication demand that the etching is performed at lower pressures (<100 mtorr). This is due to the fact that the scattering in ion directionality and energy in the sheath region is minimized, and better anisotropy is achieved, as the mean free path is increased at lower pressure. Due to inherent anisotropy at lower pressure, cleaner chemistry can be used, thus reducing the effects of chemistry-induced contamination [52]. Similarly, at lower pressure, etch reaction by-products are more volatile, and the microloading or aspect ratio dependent etch (ARDE) nonuniformity is greatly reduced.

The linear decrease in gas density with pressure requires that the ionization and dissociation efficiency of the reactor must be high to maintain adequate etch rates at lower pressures. Lower pressure operation also affects the concentration of negative ions in the discharge. Generally, the relative concentration of negative ions in the plasma is decreased due to reduced attachment cross sections at low pressure [53]. The overall kinetic behavior of low-pressure plasma differs significantly from that at higher pressure.

LOW-PRESSURE OPERATION BY MAGNETIC ENHANCEMENT

To extend the operation of an etch tool to very low pressures (<10 mtorr), some form of magnetic enhancement can be devised. Typically, magnetic fields are used to increase the collision probability between electrons and neutral feed gas molecules and thereby generating dense, stable plasma at lower pressures.

When a magnetic field is applied, the electrons are trapped in magnetic field lines and gyrate around it. This radius is smaller where magnetic field density is higher. Electrons also travel along the magnetic field line, and the resultant path is helical. Because of the negative potential of the walls and electrodes with respect to the plasma potential and gradient in the magnetic field, the electrons are reflected back and forth between points where the lines enter and leave the wall surface (referred as the "mirror effect") [54]. Electrons experience drift due to Lorentzian forces. Electrons also jump from one magnetic field line to the next as

they collide with the neutral gas molecules. At higher pressures, these kinds of collisions dominate and magnetic field have little or no effect on electron confinement. However, at low pressures, the electrons are trapped in the magnetic field, and their path length and lifetime in the system are thereby greatly increased.

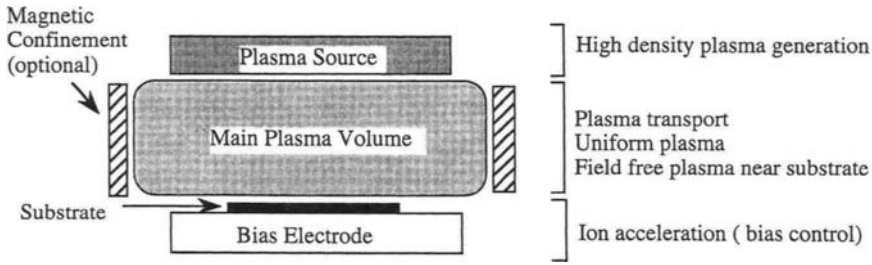
The ions also meet the same forces, but because of their greater mass and slower speed, their motion is quite different from that of electrons. Because of their greater mass, ions have much larger gyration radius than that of electrons, and ions are poorly confined in magnetic fields.

Using a proper magnetic confinement scheme can enhance ionization and dissociation efficiency of the plasma source at low pressures. There are two popular magnetic enhancement schemes. In one scheme, magnets are placed to produce magnetic field parallel to electrodes. In this scheme, it is very difficult to obtain uniform magnetic field, hence the plasma uniformity is degraded. The other scheme uses magnets in a multipolar bucket form. The multicusp arrangement is created by having rows of permanent magnets with alternating polarity. These multipolar devices are characterized by a relatively field-free region in the center and a high-intensity periphery field that "shields" the walls. Effectively, the peripheral multipolar field causes magnetic confinement of the plasma electrons and provides more efficient gas ionization at low pressures. The cusp fields also lead to much improved plasma homogeneity.

DECOUPLING OF SUBSTRATE BIAS FROM PLASMA GENERATION

To obtain faithful pattern transfer, it is desirable to operate reactors at low pressures to minimize, for example, scattering of ions while producing a dense uniform plasma to maximize etch rates [55,56,57]. An ideal plasma etch reactor should allow independent control of the substrate bias from plasma generation. In this case, the plasma generation (density, uniformity) and the substrate bias, which can cause damage, can be optimized separately. This concept is illustrated in Figure 12.

Fig. 12.



Concept of decoupling substrate bias from plasma generation.

In new reactor designs, this objective is achieved by using two powered electrodes. One electrode, on which the wafers rest is normally biased with a capacitively coupled RF power supply. The other electrode functions as the source for plasma generation. This electrode can be powered using various schemes such as capacitive, inductive, or electron cyclotron resonance excitation. Various magnetic schemes can also be applied for confining and optimizing the plasma properties at lower pressures.

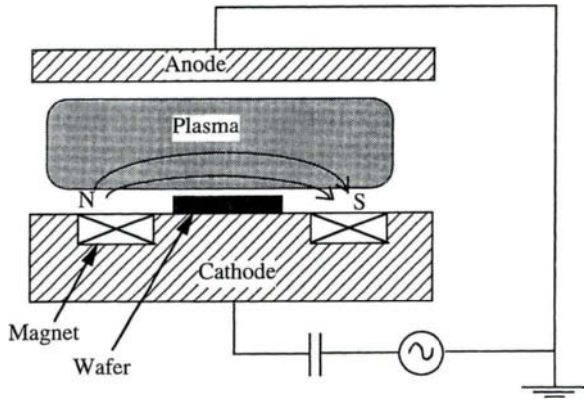
5.2.6.2 Magnetically Enhanced Diode

One of the most common approaches for magnetic enhancement is the magnetron. Magnetron sources are widely used in sputtering, and a number of publications outline the geometry of various magnetron schemes [58,59]. A comprehensive review of this subject is given by Vossen and Kern [60]. Arrangements similar to those used for magnetron sputtering have also been widely used for etching applications. In a magnetron-based etcher, a highly uniform magnetic field and planner electrodes are essential for achieving acceptable etch-uniformity. Figure 13 shows one of the most common magnetron configurations, the planar magnetron, in which a static magnetic field is applied parallel to the electrode to create an intense plasma zone [61]. This configuration is widely known as the magnetically enhanced reactive ion etcher (MERIE). A nearly parallel magnetic field is formed over the hollow region of the magnet where the wafer is normally kept. The electric field formed by the plasma sheath is perpendicular to the magnetic field, increasing the ionization efficiency near the electrode surface. The magnetic field reduces the mobility of electrons near the electrode, and the induced bias in a magnetron is generally much lower than a conventional diode system for similar operating conditions. However, the configuration shown in Figure 13, in common with all magnetron-based reactors, suffers from poor plasma uniformity and increased damage at lower pressures. To improve magnetic field uniformity, several dynamic magnetron apparatuses have also been used in commercial systems.

Figure 14 shows one of the commercially available configurations in which electromagnets are used to rotate the magnetic field at 0.5 Hz parallel to the electrode surface. The magnetic field intensity can be adjusted depending on the application. The advantage of a rotating magnetic field is that it reduces the problem of nonuniform intensification of plasma associated with the static field configuration. In another scheme, permanent magnets arranged in a racetracklike belt are mechanically scanned over the wafer surface [62].

In general, magnetron schemes result in higher etch rate and lower induced bias and can be operated at much lower pressures. Improvement in ionization efficiency and etch rate generally depends on the applied magnetic field intensity. However, magnetron schemes generally suffer from the fact that it is extremely

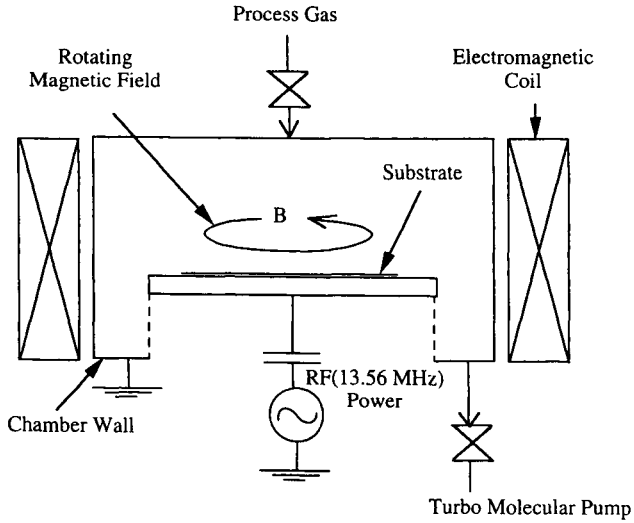
Fig. 13.



Schematics of a planar magnetron etching system with static magnetic field.

difficult to obtain very uniform magnetic field across the wafer and substrate bias is not completely decoupled from the plasma generation. Also, local nonuniformity in plasma density near the wafer surface can degrade the quality of gate oxide and device performance [63,64].

Fig. 14.



Schematic diagram of a commercial magnetically enhanced diode etching system with rotating magnetic field.

5.2.6.3 Triodes

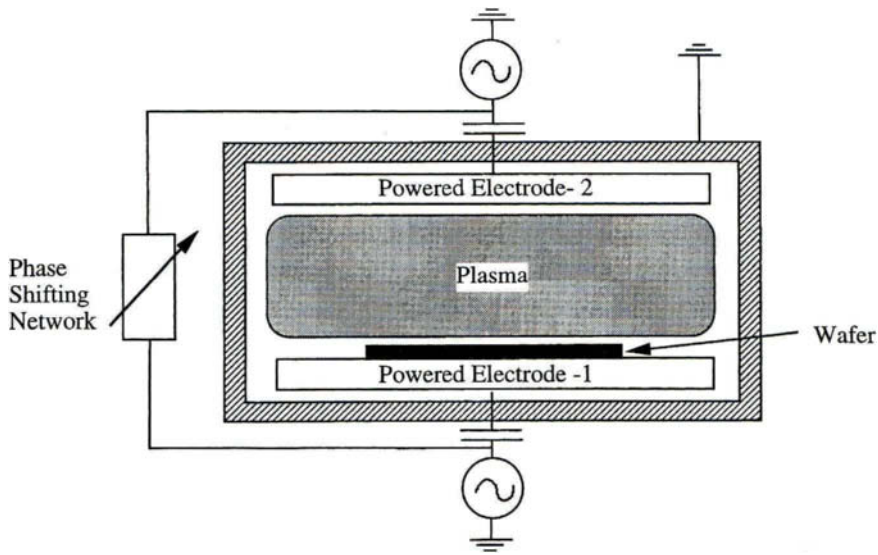
PARALLEL-PLATE TRIODES

Triodes, as their name implies, are trielectrode discharge systems. In triodes, two of the three electrodes are powered while the third electrode is normally at the ground potential. A schematic diagram of the basic parallel-plate triode system is shown in Figure 15. The plasma parameters (ion energy, electron temperature, and charged particle concentrations) can be somewhat independently varied in a triode relative to an equivalent diode system [65]. In some cases, 13.56 mHz frequency is applied to both powered electrodes. When the frequency of both electrodes is similar, the relative phase between the two electrodes can significantly affect the plasma properties and distribution [66,67].

HOLLOW-ANODE-ENHANCED TRIODES

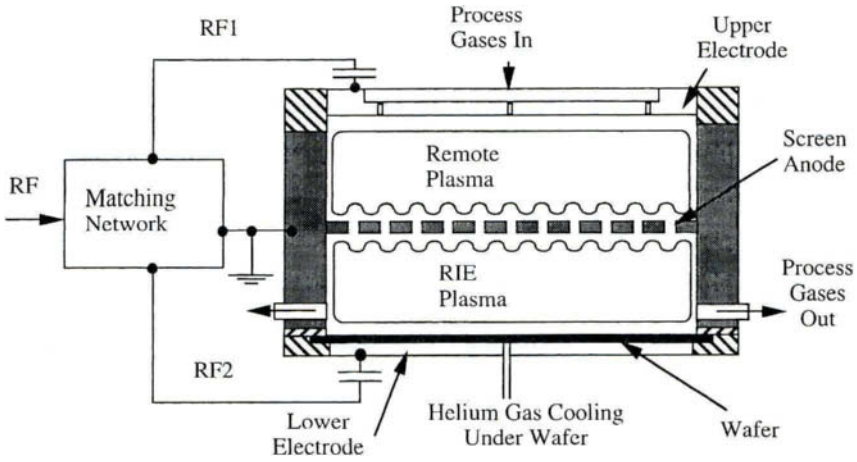
Another commercially available triode configuration uses a screened, grounded anode between the upper and lower electrodes of a parallel-plate triode to isolate plasma generation from the substrate environment [68,69]. This configuration is known as a hollow-anode-enhanced triode configuration. A schematic diagram of

Fig. 15.



Schematic diagram of a parallel-plate triode system.

Fig. 16.

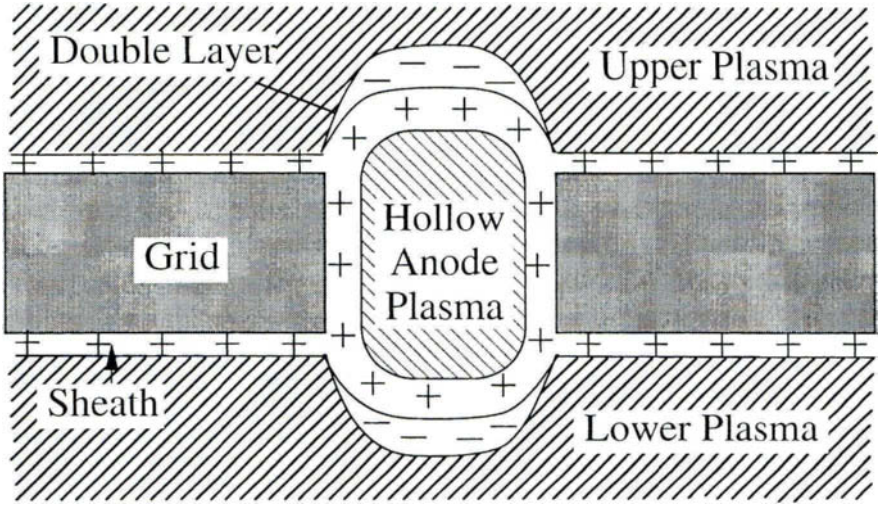


Hollow-anode-enhanced triode reactor configuration.

this reactor configuration is shown in Figure 16. In this configuration, a grounded, screened grid (referred to as an anode grid) separates the plasma created by the upper and lower electrodes arranged in a conventional triode configuration where the chamber ground is the third electrode. This configuration produces two plasma regions, which allow the substrate bias to be decoupled from the plasma generation. The upper electrode plasma can also serve to enhance the etchant species yield at the wafer surface, and the lower plasma is used to produce a controllable substrate bias. An interesting aspect of this configuration is that a very intense plasma region is formed in the apertures of the separation grid. It produces an extremely high-density plasma with considerable increase in etchant species, possibly caused by charge exchange, as well as other effects. This phenomenon is known as hollow anode effect. The hollow anode phenomenon appears to arise from the interaction of anode wall sheaths in a cavity geometry such as an aperture at ground potential in a RF discharge. The hollow anode phenomenon appears to provide a unique approach to simultaneously intensifying and decoupling the plasma generation in etching reactors. This high-density plasma in the apertures of the grid is separated from surrounding lower-density regions by a double-layer sheath at both ends of the column. The plasma density in the hollow anode region can be higher than the surrounding plasma by a factor of 50.

The structure of the hollow anode plasma is illustrated in Figure 17. The overall structure is very similar to that occurring at the constriction between two plasmas as described by Andrews and Allen [70] or to the plasma region occurring in the compression electrode in a duoplasmatron ion sources [71]. The hollow anode phenomenon is caused by the combination of electrostatic confinement at bound-

Fig. 17.



Model of a hollow anode plasma.

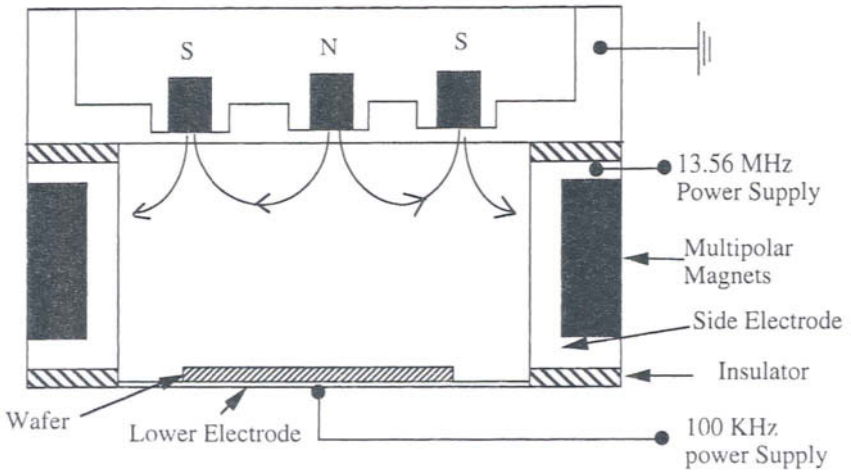
ary sheath and particle acceleration in the double sheath. A double layer is formed on both side of the aperture in order to maintain continuous discharge current through the aperture. As a result, the electron density is higher inside the aperture than it is outside. The combination of ion sheaths of the grounded walls of the aperture and the double sheaths forms an electrostatic potential well. Ions from the aperture plasma are accelerated toward the bulk plasma because of the electrostatic field in the double sheath, and the ratio of the ion loss rate to the creation rate is greater inside the aperture than outside. Thus, for the plasma inside the aperture to be quasi-neutral, the electrons flowing into the aperture from the bulk of the plasma must have more energy, to increase the ionization rate. This is achieved by the acceleration of electrons from the bulk plasma toward the aperture plasma as they traverse through the double sheath. For molecular gases commonly used in plasma etching, the increased ionization in the aperture plasma is also associated with an increase in reactive radical formation. The plasma potential of the aperture plasma self-adjusts to maintain charge neutrality.

The presence of the anode grid significantly increases the etch rates as compared to a conventional triode (formed by removing the grid) under same conditions of substrate bias, total power, and pressure.

MAGNETICALLY ENHANCED TRIODES

The ideal magnetic confinement scheme should be simple and compatible with the existing parallel-plate etching tools. The preferred configuration should have

Fig. 18.

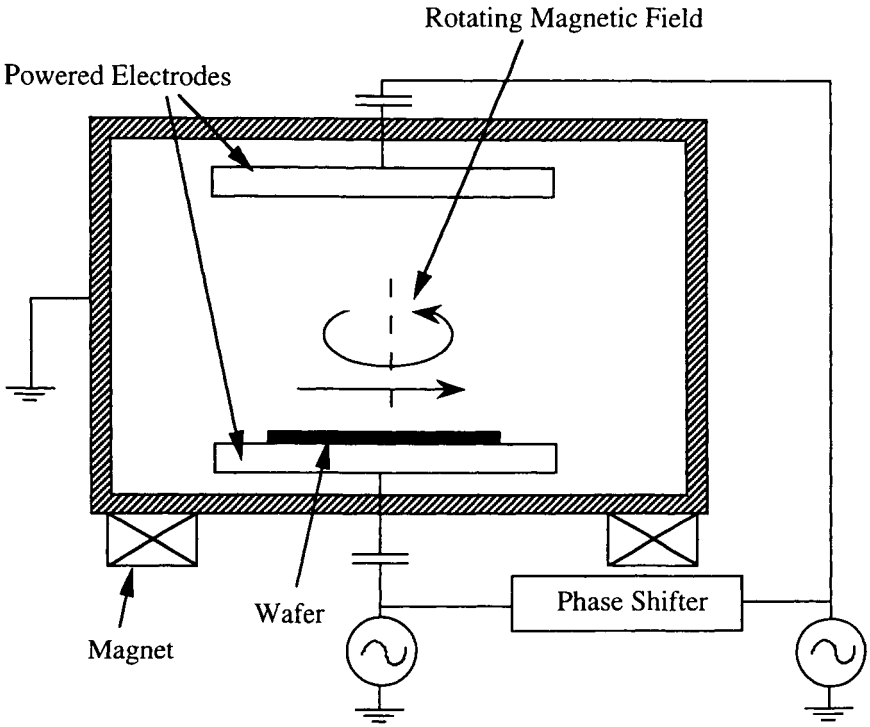


Schematic of a commercial tri-electrode reactor with magnetic confinement.

all the attributes of magnetic enhancement, namely low-pressure operation, high plasma density, low substrate bias voltage, very low magnetic field near the wafer surface, and decoupling of plasma generation (reactive species concentration) from the substrate bias (ion bombardment energy). Many approaches have been proposed to decouple plasma generation from substrate bias using triode configuration while using magnetic enhancement to increase etch rates at lower pressures. One commercially available approach uses a dual-frequency magnetically confined triode configuration and is shown in Figure 18 [72]. This configuration includes two RF-powered electrodes (13.56 mHz and 100 kHz) and a grounded electrode. 100 kHz power is applied to the wafer electrode and is used to control the energy of the ions incident on the wafer. The 13.56 mHz power is applied to the side wall electrode and independently controls the amount of ionization and reactive species generation. The upper electrode is grounded in this reactor geometry. Permanent magnets surrounding the chamber reduce the loss of charged particles to the reactor walls and thereby enhance the plasma density. The magnetic field is away from the wafer and does not directly influence etching at wafer surface.

Another approach uses a single-frequency (13.56 mHz) excited, phase-controlled, triode configuration (dubbed *supermagnetron*) with mechanically rotating permanent annular magnets to generate magnetic field parallel to the electrode surface and is shown in Figure 19 [73]. This approach represents complex experimental arrangement and also suffers from the nonuniform etching and high substrate damage due to intense magnetic field near the wafer surface. In another approach, the magnetic field is created by a multipolar bucket as an extension of

Fig. 19.



Schematics of a supermagnetron plasma etcher.

the ground shield surrounding either electrode [74]. This approach seems to produce very uniform, field-free plasma region near the wafer surface at low pressures.

5.2.6.4 Inductively Coupled Plasma Sources

Inductively coupled plasmas (ICP) have been known for over 100 years. A detailed review of the subject is given by J. Hopwood [75]. An inductively coupled plasma source uses an inductive circuit element adjacent to or immersed inside the discharge region to couple energy from an RF power source to the plasma. The inductive circuit element is typically a helical or spiral-like conductor. An external circuit is normally used to generate a resonance condition at the driving frequency and causes large currents to flow through the inductive element. These currents generate RF magnetic flux, which penetrates through the discharge volume. The time-varying magnetic flux induces a solonoidal RF electric field that accelerates free electrons in the discharge and sustains the plasma.

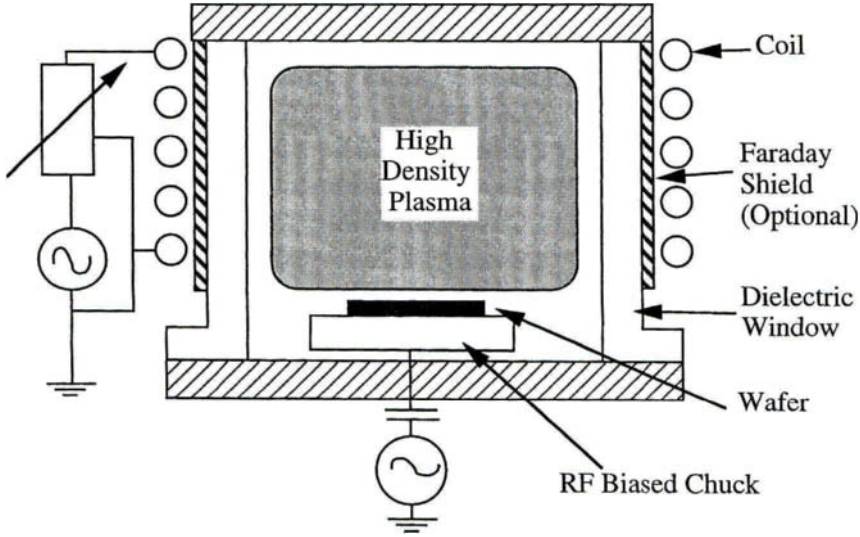
Inductively coupled plasma systems were initially used for semiconductor pro-

cessing in a configuration known as a barrel reactor (see Section 5.2.5.2). Recently, the trend toward high-rate, low-pressure, single-wafer processing has motivated the development of low-pressure ICP configurations. Very high plasma densities (in the range of 10^{12} ions/cm^3) have been reported in various ICP configurations [76]. The three most common ICP configurations used today are helical inductive couplers, helical resonators, and spiral inductive couplers.

HELICAL INDUCTIVE COUPLERS

A schematic presentation of a helical inductive coupler-type ICP reactor is shown in Figure 20. The plasma is generated within a dielectric cylinder, to which the coil is wound. The magnetic field lines generated by this configuration are parallel to the central axis of the cylinder. The induced electric field is azimuthal and forms closed loops about the axis. The induction field is maximum near the wall of the cylinder and decreases monotonically toward the center. At lower pressure, diffusion processes dominate the bulk recombination processes and the plasma is more uniformly distributed. Since a very high capacitive axial electric field exists between the two ends of the inductive coupler owing to high potential difference, no ICP is entirely inductively coupled. The degree of capacitive coupling in ICP discharges is high at lower power and decreases as the ICP power is increased [77,78]. At high power, the contribution of capacitive coupling to the total discharge is negligible compared with the inductive coupling component.

Fig. 20.



A schematic presentation of helical inductive coupler-type ICP reactor.

In many ICP configurations, a conducting Faraday shield is placed around the chamber to short out the axial electric field [79]. Note that the design of the shield should be such that it does not prevent the coupling of induced electric field to the discharge.

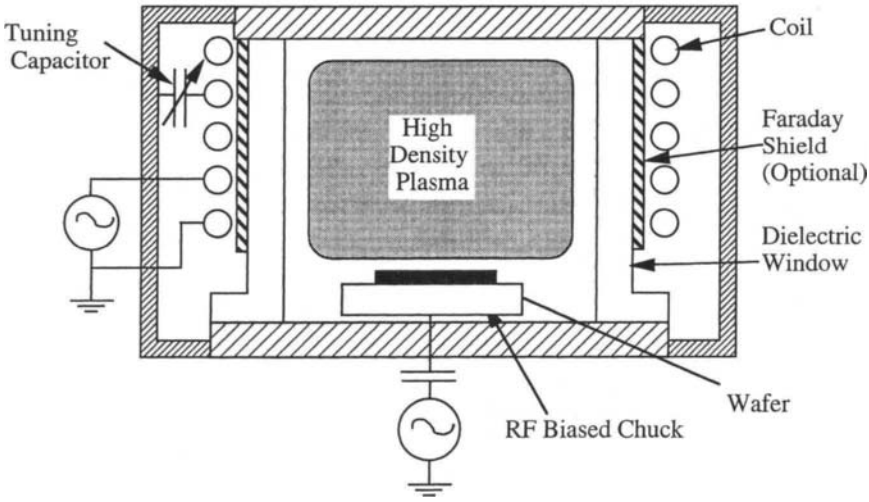
One of the earliest applications of this type of configuration in plasma processing was the barrel reactor. In barrel reactors, the wafers are loaded in the center of the cylinder. This is not an optimized configuration, because the wafers and wafer holder can significantly disturb the induction fields.

For this type of configuration designed for the low-pressure applications, the wafers are normally placed in a remote downstream chamber on a chuck with an independent RF bias. This configuration relies on the diffusion of the plasma generated by the ICP source.

HELICAL RESONATORS

The helical resonator plasma source differs from the conventional helical coupler source in the coil design. The coil is positioned around a cylindrical chamber similar to the helical coupler source. However, the coil has an electrical length of $(\lambda/4 + n\lambda/2)$ or $(\lambda/2 + n\lambda/2)$ where $n = 0, 1, 2, \ldots$ and λ is the wavelength of the excitation frequency. When the electrical length of the coil is $(\lambda/4 + n\lambda/2)$, the configuration is a quarter-wave resonator and when the electrical length of the coil is $(\lambda/2 + n\lambda/2)$, the configuration is a half-wave resonator. The coil is enclosed in a metallic (conductive) cavity as shown in Figure 21. This configuration

Fig. 21.



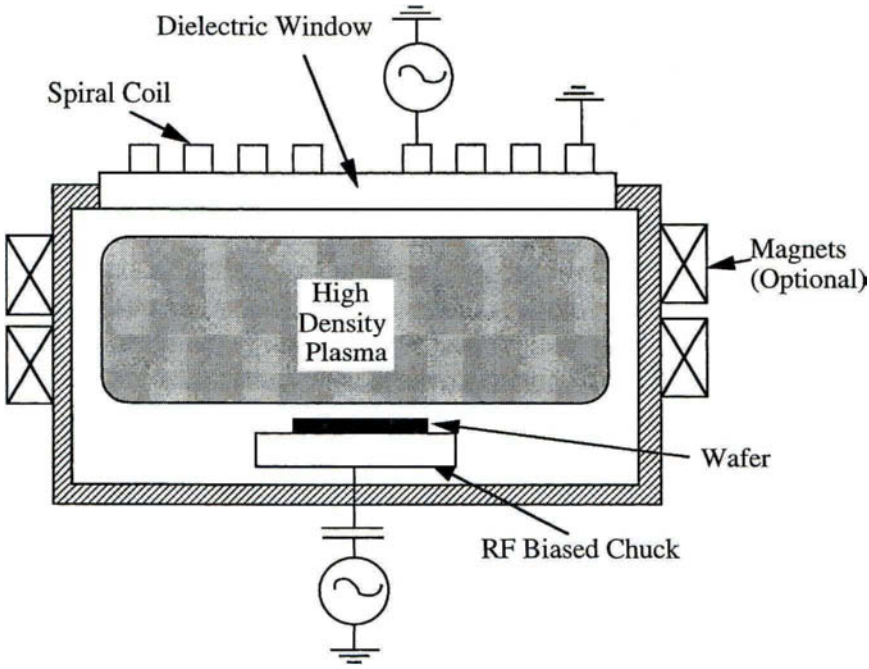
A schematic presentation of helical resonator plasma reactor.

provides a parasitic capacitance from the coil to the ground. Because the plasma conditions can affect the resonance condition of the coil, a tunable capacitor is normally added between the coil and ground to maintain resonance.

SPIRAL INDUCTIVE COUPLERS

In plasma processing, it is very desirable to generate uniform, dense plasma over a large area. One configuration that is widely used to achieve this objective uses spiral-like planar coil [80]. A typical planar design is shown in Figure 22. The planar coil is separated from the chamber by a dielectric window. The matching network is designed to generate resonance in the coil. The large currents flowing in the coil due to the resonance condition generates an oscillating magnetic field around the coil that penetrates into the plasma and induces an azimuthal electric field. In an ideal reactor, the azimuthal electric field is zero on the periphery as well as the center, peaking in an annular region at nearly half the radius. To maximize the uniformity of the plasma generated by the spiral-like planar coil, the coil geometry (spacing, size, shape) is normally optimized for a given reactor configuration and a given plasma process.

Fig. 22.



Schematic representation of a spiral-like planar ICP reactor.

Since the skin depth of a 13.56 mHz RF induction field is approximately 1–2 cm in plasmas with electron densities in the order of 10^{11} ions/cm^3, the substrate may be placed in a close proximity to the inductive coil. Typically, the coil is separated from the plasma by a 1- to 3-cm-thick dielectric window and the substrate is positioned 5–10 cm below the window. The induced electric field, and hence the ion generation rate, decay exponentially with the distance, and normally a tradeoff is made on wafer-positioning distance between maintaining high plasma density near the substrate and allowing enough diffusion distance for plasma uniformity.

MATCHING NETWORK FOR ICP SOURCES

The matching network for inductively coupled sources should be able to create a resonance condition forcing large current to flow through the coil. At the same time, the design should not allow generation of very high voltages at nodal points. It should also be capable of easily igniting the plasma and should have a wide match window for a large range of parameters. The matching network should also be able to handle high power, since the plasma density does not saturate with power in excess of 2–3 kw for ICP sources.

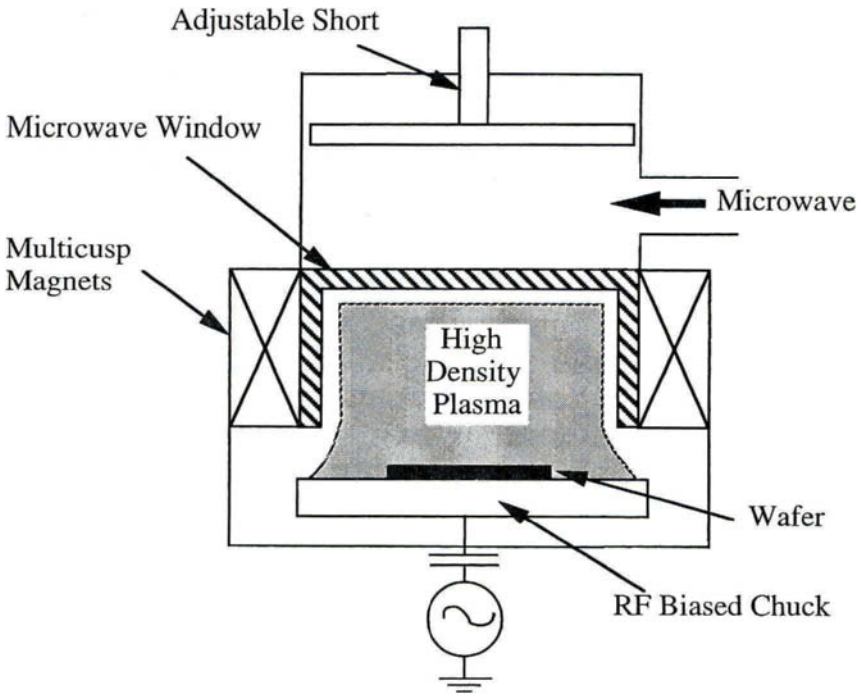
5.2.6.5 Electron Cyclotron Resonance Sources

Electron cyclotron resonance (ECR) plasma sources have received a great deal of attention due to their ability to produce very intense discharge at very low pressures. ECR sources use microwave excitation schemes (predominantly 2.45 mHz excitation) to produce low-pressure plasma. ECR sources are characterized by very low pressure of operation (0.1 mtorr to 10 mtorr). In various ECR schemes, the microwave power is generated remotely and is coupled to the etch chamber through a waveguide and a dielectric window. A magnetic field of proper magnitude is applied around the chamber to create a resonance condition. The microwave field oscillates in resonance with the electron cyclotron motion about the magnetic fields. The power absorption is localized to the resonance zone. There are two commonly used ECR configurations: plasma disk source and diffusion source.

PLASMA DISK SOURCE

In the plasma disk source approach (Figure 23), the microwaves are directly coupled and tuned into the plasma chamber via internal adjustments and confined by a set of electromagnets or permanent magnets in a multicusp array [81,82]. The source is then tuned internally via a sliding short to produce a resonance cav-

Fig. 23.



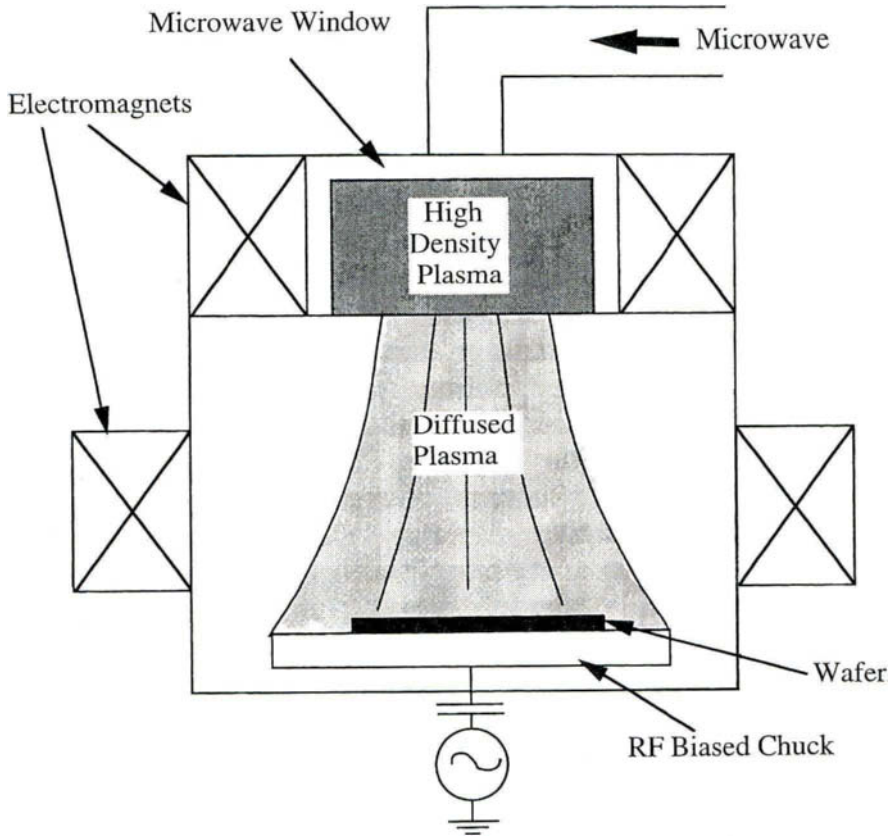
A schematic diagram of a plasma disk ECR source configuration.

ity condition. The magnets are normally arranged such that the accelerating field diverging the plasma toward the substrate is minimal, thus maintaining the low energy and directionality of the species from the source to the substrate.

DIFFUSION SOURCE

In the diffusion source configuration (Figure 24), high-density plasma is generated in a source region where the electron cyclotron resonance is maintained [83,84]. The wafer is normally placed in a downstream region, and plasma from the source region is allowed to diffuse into this region. Ion energy and flux to the substrate is normally controlled by applied RF bias to the wafer electrode. Note that the plasma density in the source region is extremely high. However, the density of the diffused plasma near the substrate can be an order of magnitude lower. Since the uniformity and directionality of charged particles near the wafer surface are strongly affected by the local magnetic field, these schemes have the potential problem of achieving high uniformity over larger-diameter substrate. Other prob-

Fig. 24.



A schematic diagram of a diffusion ECR source reactor.

lems associated with these sources are the inherent high cost and noncompatibility with existing parallel-plate production etchers.

MATCHING NETWORK FOR ECR SOURCES

The microwave systems usually employ one to four cylindrical tuning stubs inserted into the waveguide between the power source and the load [85]. These tuning stubs can be adjusted either automatically or manually. For automatic tuning to minimize reflected power, the sensors inserted in the waveguide measure the standing waves in the waveguide and feed back the information to the signal processing unit for calculation of the impedance in the chamber. The tuning stubs are adjusted accordingly to minimize reflected power.

5.2.6.6 Summary of Various Reactor Configurations

Various reactor configurations are summarized in Table 2.

5.2.7 NEW TRENDS

Some of the recent trends influencing plasma etch technology are briefly described in this section.

Table 2

Summary of Various Reactor Configurations

Category	Excitation Scheme	Plasma Density (ions/cm^3)	Pressure Range (mtorr)	Comments
Barrel	Inductive or Capacitive 13.56 mHz	$<10^{10}$	>1000	Batch etcher Isotropic, mainly used for ashing
Diode	Capacitive 13.56 mHz	$<10^{10}$	300–10,000	Batch etcher Isotropic or directional
RIE	Capacitive 13.56 mHz	$<10^{10}$	30–300	Batch etcher or SWE Anisotropic Excellent uniformity Easy to scale up
MERIE	Capacitive 13.56 mHz	10^{10}–10^{11}	10–50	SWE, anisotropic, acceptable uniformity, difficult to scale up
Hollow-anode-enhanced triode	Capacitive 13.56 mHz	10^{10}–10^{12}	10–100	SWE, high localized plasma density, excellent uniformity, easy to scale up
Magnetically enhanced triode	Capacitive single or dual freq., 100 kHz, 13.56 mHz	10^{10}–10^{11}	10–50	SWE, anisotropic, acceptable uniformity, difficult to scale up
ECR	Microwave 2.45 gHz	10^{10}–10^{12}	0.1–10	SWE, very-low-pressure operation, high density at source, difficult to scale up
ICP	Inductive 13.56 mHz	10^{10}–10^{12}	1–100	SWE, high plasma density, relatively easy to scale up

END POINT DETECTION

End point detection of a given etch process has become very critical because very thin layers (<100 Å) used for modern VLSI and ULSI circuits cannot tolerate overetching. Laser reflection interferometry [86] was used in early days for end point detection. At present, optical emission spectroscopy (OES) technique is being widely used in production for end point control. In OES, the end point is determined by following time dependence of a selected wavelength of the optical spectrum emitted by the plasma representing either a reactant or a by-product. By monitoring the change in optical signal associated with completion of an etch reaction, the end point can be called and the etch process can be terminated.

Presently available OES end point detection equipment are reaching a limit for the etch applications involving etching of very little exposed areas (<2%) on the wafer (such as via etch), due to very low signal-to-noise ratio. A major emphasis is being given to optimization of the detection capability of the OES equipment [87,88]. In parallel, several other techniques such as ellipsometry [89], plasma impedance monitoring [90], infrared absorption spectroscopy [91], and neural networks [92] are also being investigated.

CLUSTER TOOLS

In the past few years, many etch chambers have been integrated on the same platform in a form of a cluster tool. The concept of a cluster tool was developed to improve yield by reducing particles through automated, in-vacuum transfer between etch modules. The throughput is also increased because many operations can be done in parallel. Clustering, however, places stringent requirements on equipment reliability due to interdependence of the equipment set and the inherent lack of flexibility [93,94].

ELECTROSTATIC CHUCKS

Electrostaic chucks are being integrated on proven etch platforms by almost all equipment vendors as a replacement for conventional Helium backside or cryogenically cooled clamped electrodes [95]. The electrostatic chuck holds the wafer with no front side contact, offering significant improvement in total number of dies that can be processed. In addition, the elimination of moving clamp parts help reduce particle generation and simplifies chamber design.

ORIGINAL EQUIPMENT MANUFACTURE (OEM) SUPPORT

In past few years, the up-time and reliability of equipment has become very critical, because of financial constraints in high-volume manufacturing of semicon-

ductor circuits. The software and robotic systems used on an etch tool are constantly being upgraded. Also, in many cases the etch equipment is linked to the OEM through a telephone line so that the performance of the machine can be monitored on-line and potential problems can be diagnosed and resolved without sending out a field service engineer.

GUARANTEED PROCESSES

In the initial days of the semiconductor industry, the majority of the etch processes were developed by engineers in the fabrication facility, on an in-house-built or a vendor-supplied reactor. In the past few years, the trend has been shifted toward process development by the OEM. At present, the majority of the production machines are being sold with a guaranteed process performance for a specific application.

REFERENCES

1. A. T. Hell, Spatial distribution of electron density and electric field strength in a high-frequency discharge, *Ind. Chem. Eng. Fundam.*, **9** (1970) 160–166.
2. M. J. Kushner, Mechanisms for power deposition in Ar/SiH4 capacitively coupled RF discharges, *IEEE Trans. on Plasma Sci.*, **PS-14** (1986) 188–196.
3. D. Flamm and G. Herb, *Plasma Etching: An Introduction*, edited by D. Mannos and D. Flamm (Academic Press, San Diego, 1989), chap. 1, p. 37.
4. H. S. Butler and G. S. Kino, Plasma sheath formation by radio-frequency fields, *Phys. Fluids*, **6** (1963) 1346–1355.
5. K. Riemann, Theoretical analysis of the electrode sheath in RF discharges, *J. Appl. Phys.*, **65** (1989) 999–1004.
6. M. A. Liberman, Dynamics of a collisional, capacitive RF sheath, *IEEE Trans. on Plasma Sci.*, **PS-17** (1989) 338–341.
7. D. Vender and R. Boswell, Electron-sheath interaction in capacitive radio-frequency plasmas, *J. Vac. Sci. Technol.*, **A10**(4) (1992) 1331–1338.
8. P. M. Vallinga, P. M. Meijer, and F. J. de Hoog, Sheath properties of RF plasmas in a parallel plate etch reactor; the high frequency regime, *J. Phys. D. Appl. Phys.*, **22** (1989) 1650–1657.
9. G. R. Misium, A. J. Lichtenberg, and M. Liberman, Macroscopic modeling of radio frequency plasma discharges, *J. Vac. Sci. Technol.*, **A7**(3) (1989) 1007–1013.
10. D. E. Bohm, Minimum ionic kinetic energy for a stable sheath, in *The Characteristics of Electrical Discharges in Magnetic Fields*, edited by A. Guthrie and R. K. Walkerling (McGraw-Hill, New York, 1949).
11. F. F. Chen, *Introduction to Plasma Physics and Controlled Fusion*, vol. 1, 2nd ed. (Plenum Press, New York, 1984).
12. J. Meichsner, H. U. Poll, and K. H. Wixkleder, Negative ions in a glow discharge of tetrafluoromethane, *Contrib. Plasma Phys.*, **25**(5) (1985) 503–512.
13. N. St. J. Braithwaite and J. E. Allen, Boundaries and probes in electronegative plasmas, *J. Phys. D: Appl. Phys.*, **21** (1988) 1733–1737.

14. R. Gottscho, Glow-discharge sheath electric fields: Negative-ion, power, and frequency effects, *Phys. Review*, **A36**(5) (1987) 2233–2240.
15. Laura Peters, Plasma etch chemistry: The untold story, *Semiconductor International* (May 1992) 66–71.
16. J. W. Coburn and H. F. Winters, Ion and electron assisted gas surface chemistry—An important effect in plasma processing, *J. Appl. Phys.*, **50** (1979) 3189–3194.
17. E. Kay, J. Coburn, and A. Dilks, Plasma chemistry of fluorocarbons as related to plasma etching and plasma polymerization, *Topics Current Chem.*, **94** (1980) 1–9.
18. J. Cook and K. Donohoe, Etching issues at 0.35 μm and below, *Solid State Technology* (April 1991) 119–124.
19. R. d'Agostino, F. Cramarossa, S. DeBenedictis, and G. Ferraro, Spectroscopic diagnosis of CF_4–O_2 plasmas during Si and SiO_2 etching processes, *J. Appl. Phys.*, **52** (1981) 1259–1265.
20. L. Ephrath, Dry-etching for VLSI—A review, *J. Electrochem. Soc.*, **129** (1982) 2282–2287.
21. D. Economou, E. Aydil, and G. Barna, In situ monitoring of etching uniformity in plasma reactors, *Solid State Technol.*, **34**(4) (1991) 107–111.
22. P. Singer, Plasma etch: A matter of fine-tuning, *Semiconductor International*, **12** (1995) 65–68.
23. Y. Liu and M. Flowers, Kinetics of triode mode reactive ion etching of Si (100) wafers by chlorine plasmas: Temperature and DC self-bias effects, *Vacuum* **42**(18) (1991) 1213–1217.
24. A. Watts and W. Varhue, Low temperature etching of silicon trenches with SF_6 in an electron cyclotron resonance reactor, *J. Vac. Sci. Technol.*, **A10**(4) (1992) 1313–1330.
25. A. De Dios, E. Castan, L. Bailon, J. Barbolla, M. Lozano, and E. Lora-Tamayo, RIE-induced damage in MOS structures, *Solid-State Electronics*, **33**(11) (1990) 1419–1423.
26. D. Misra, W. Zhong, R. Bartynski, V. Patel, and B. Singh, Etch induced damage in high density inductively coupled plasma etching reactors, *Semicond. Sci. Technol.*, **11** (1996) 816–821.
27. X. Li, T. Brozek, F. Preuninger, D. Chan, and C. Viswanathan, Evaluation of Plasma Damage Using Fully Processed MOS Transistors, *J. Vac. Sci. Technol.*, **B11**(1) (1996) 571–576.
28. V. Menon, Fine Particle Society Annual Meeting, Santa Carla, CA (July 1991).
29. S. J. Choi and M. J. Kushner, The role of negative ions in the formation of particles in low pressure plasmas. *Appl. Phys. Letters*, **62** (1993) 2562–2564.
30. G. Selwyn, A phenomenological study of particulates in plasma tools and processes, *Jpn. J. Appl. Phys.*, **32**(6B) (1993) 3068–3073.
31. J. Maa, H. Gossenberger, and R. Paff, Effect of post-etch treatment on chlorine concentration of AlSi and Ti capped AlSi films, *J. Vac. Sci. Technol.*, **B8**(5) (1990) 1052–1057.
32. N. Parekh and J. Price, VLSI Multilevel Interconnection Conference Proceedings, Phoenix, AZ (IEEE, New York, 1989), pp. 506–511.
33. S. Pearson, M. Wilders, and A. Chambers, Reactive ion etching of AlSiCu alloy films, *Semiconductor International* (February 1987), pp. 97–100.
34. T. Y. Fok, *Electrochemical Society Extended Abstracts 80-1, 1980* (Electrochemical Society, Pennington, NJ, 1980), Abstract No. 115.
35. H. Busta, R. E. Lajos, and D. Kiewit, Plasma etch monitoring with laser interferometry, *Solid State Technol.*, **22**(2) (1979) 62–65.
36. V. Godyak and R. Piejak, In situ simultaneous radio frequency discharge power measurements, *J. Vac. Sci. Technol.*, **A8**(5) (1990) 3833–3837.
37. B. Andries, G. Ravel, and L. Peccoud, Electrical characterization of radio frequency parallel plate capacitively coupled discharges, *J. Vac. Sci. Technol.*, **A7**(4) (1989) 2774–2783.
38. G. Selwyn and E. Key, Spatially resolved optical emission studies of fluorocarbon RF plasma through the use of UV-transmitting optical fibers, *Plasma Chemistry and Plasma Processing*, **5**(2) (1985) 183–199.
39. T. Chau and K. Kao, Optical emission spectra of microwave oxygen plasmas and fabrication of SiO_2 films. *J. Vac. Sci. Technol.*, **B14**(1) (1996) 527–532.

40. K. Maruyama, K. Ohkouchi, Y. Ohtsu, and T. Goto, CF_3, CF_2 and CF radical measurements in RF $CHEF_3$ etching plasma using infrared diode laser absorption spectroscopy, *Jpn. J. Appl. Phys.*, **33**(7B) (1994) 4298–4302.
41. H. C. Sun, E. Whittaker, Y. Bae, C. Ng, V. Patel, W. Tam, S. McGuire, B. Singh, and B. Gallois, Combined wavelength and frequency modulation spectroscopy: A novel diagnostic toll for materials processing. *Applied Optics*, **32**(6) (1993) 885–893.
42. J. Coburn, A system for determining the mass and energy of particles incident on a substrate in a planar diode sputtering system. *Rev. Sci. Instrum.*, **41**(8) (1970) 1219–1223.
43. D. W. Grave, T. J. Knight, X. Cheng, B. Krogh, M. Gibson, and J. LaBrosse, Process control based on quadrupole mass spectroscopy, *J. Vac. Sci. Technol.*, **B14**(1) (1996) 489–493.
44. H. Hermansdorfer, Microwave diagnostic techniques, in *Plasma Diagnostics*, edited by W. Lochte-Holtgreven (North-Holland, Amsterdam, 1968).
45. I. Langmuir and H. Mott-Smith, Collected works of Irving Langmuir, Vol. 4 (Pergamon Press, Long Island City, NY (1961), pp. 99–132.
46. P. Chung, L. Talbot, and K. Touryan, Electrical probes in stationary and flowing plasma: Part 1, Collisional and transitional probes; Part 2, Continuum probes, *AIAA Journal*, **12**(2) (1974) 133–154.
47. B. E. Cherrington, The use of electrostatic probes for plasma diagnostis—A review, *Plasma Chemistry and Plasma Processing*, **2**(2) (1982) 113–140 and references therein.
48. A. R. Reinberg, Plasma processing with a planar reactor, *Circuits Mfg.* (April 1979) 25–31.
49. G. K. Herb, D. Rigger, and K. Shields, Silicon trench etch in a hex reactor, *Solid State Technology* (October 1987) 109–115.
50. V. Donnelly and D. Flamm, Effects of frequency on optical emission, electrical ion, and etching characteristics of a radio frequency chlorine plasma, *J. Appl. Phys.*, **58**(6) (1985) 2135–2144.
51. S. Schultheis, Single-wafer high rate RIE employing magnetron discharge, *Solid State Technol.*, **28**(4) (1985) 233–237.
52. R. Hill, Characterization of a low pressure, high ion density, plasma metal etcher, *J. Vac. Sci. Technol.*, **B14**(1) (1996) 547–551.
53. D. Sheehan and N. Rynn, Negative ion plasma sources, *Rev. Sci. Instrum.*, **59**(8) (1988) 1369–1375.
54. B. Singh and P. R. Denton, The magnetron: Many gains and more to come, *Optical Spectra*, **15**(10) (1981) 77–81.
55. J. W. Coburn, *Plasma Etching and Reactive Ion Etching* (American Vacuum Society, New York, 1982), and references therein.
56. C. M. Melliar-Smith and C. J. Mogab, in *Thin Film Processes*, ed. by J. L. Vossen and W. Kern (Academic Press, New York, 1979).
57. R. Kraft, T. Boonstra, and S. Prengle, Etching 0.35 μm polysilicon gates on a high density helicon etcher, *J. Vac. Sci. Technol.*, **B145**(1) (1996) 543–546.
58. R. Waits, Planar magnetron sputtering, *J. Vac. Sci. Technol.*, **15**(2) (1978) 179–187.
59. J. Thornton, Magnetron sputtering: Basic physics and application to cylindrical magnetrons, *J. Vac. Sci. Technol.*, **15**(2) (1978) 171–177.
60. J. Vossen and W. Kern, eds., *Thin Film Processes* (Academic Press, New York, 1978).
61. H. Kinoshita, T. Ishida, and S. Ohno, Highly uniform magnetron etching system using an annular permanent magnet, *Appl. Phys. Lett.* **50**(25) (1987) 1838–1841.
62. H. Okano, T. Yamazaki, and Y. Horiike, High rate reactive ion etching using a magnetron discharge, *Solid State Technol.*, **25**(4) (1982) 166–170.
63. A. Hoff, T. Esry, and K. Nauka, Monitoring plasma damage: A real-time, noncontact approach, *Solid State Technol.*, **39**(4) (1996) 139–148.
64. W. En, B. Linder, and N. Cheung, Modeling of oxide charging effects in plasma processing, *J. Vac. Sci. Technol.*, **B14**(1) (1996) 552–556.

65. G. J. Gorin, U.S. Patent No. 4,464,223 (August 1984).
66. J. H. Thomas III and B. Singh, Effect of the radio-frequency voltage phase in a balanced triode plasma etching reactor, *J. Vac. Sci. Technol.*, **A10**(5) (1992) 3070–3075.
67. V. Patel, *Characterization of Advanced Etching Reactors Using Novel Diagnostic Tools*, unpublished doctoral thesis, New Jersey Institute of Technology, 1993.
68. E. Bogle-Rohwer, D. Gates, L. Hayler, H. Kurasaki, and B. Richardson, Wall profile control in a triode etcher, *Solid State Technol.*, **28** (1985) 251–255.
69. J. F. Donohue, M. Kent, and A. Sampson, Geometric focusing technology applied to SiO2 etching, *J. Vac. Sci. Technol.*, **B14**(1) (1996) 466–470.
70. J. Andrews and J. Allen, Theory of a double-sheath between two plasmas, *Proc. Roy. Soc. Lon.*, **A320** (1971) 459–472.
71. B. Singh, unpublished doctoral thesis, *Investigation of Some of the Physical Processes Occurring in High Current Ion Sources*, University of Aston, Birmingham, UK, 1978.
72. P. Laporte, Y. Melaku, and M. Goethals, Magnetically confined dry etching for silylated resist development, *Solid State Technol.*, **34**(4) (1991) 69–73.
73. H. Kinoshita and O. Matsumoto, A new supermagnetron plasma etcher remarkably suited for high performance etching, *J. Vac. Sci. Technol.*, **B9**(2) (1991) 325–333.
74. B. Singh, J. H. Thomas, III, and V. Patel, Magnetic multipole based reactive ion etching reactor, *Appl. Phys. Lett.*, **60**(19) (1992) 2335–2337.
75. J. Hopwood, Review of inductively coupled plasmas for plasma processing, *Plasma Sources Sci. Technol.*, **1**(2) (1992) 109–116.
76. J. Givens, S. Geissler, J. Lee, O. Cain, J. Marks, P. Keswick, and C. Cunningham, Selective dry etching in a high density plasma for 0.5 μm CMOS technology, *J. Vac. Sci. Technol.*, **B12**(1) (1994) 427–432.
77. R. B. Piejak, V. Godyak, and B. Alexandrovich, A simple analysis of an inductive RF discharge, *Plasma Sources Sci. Technol.*, **1**(2) (1992) 109–116.
78. A. Perry, D. Vender, and R. Boswell, The application of the helicon source to plasma processing, *J. Vac. Sci. Technol.*, **B9**(2) (1991) 310–317.
79. H. Sugai, K. Nakamura, and K. Suzuki, Electrostatic coupling of antenna and the shielding effect in inductive RF plasmas, *Jpn. J. Appl. Phys.*, **33**(4B) (1994) 2189–2193.
80. J. Carter, J. Holland, E. Peltzer, B. Richardson, E. Bogle, H. Nguyen, Y. Melaku, D. Gates, and M. Ben-Dor, Transformer coupled plasma etch technology for the fabrication of subhalf micron structures, *J. Vac. Sci. Technol.*, **A11**(4) (1993) 1301–1306.
81. P. Singer, Trends in plasma sources: The search continues, *Semiconductor International* (July 1992), pp. 52–58.
82. S. Samukawa, Highly selective and highly anisotropic SiO_2 etching in pulse-time modulated electron cyclotron resonance plasma, *Jpn. J. Appl. Phys.*, **33**(4B) (1994) 2133–2138.
83. J. Cook, D. Ibbotson, P. Foo, and D. Flamm, Etching results and comparison of low pressure ECR and RF discharge sources, *J. Vac. Sci. Technol.*, **A8**(3) (1990) 1820–1824.
84. M. Goeckner, J. Meyer, G. Kim, J. Jenq, J. Taylor, and R. Breun, Role of contaminants in ECR plasmas, *J. Vac. Sci. Technol.*, **A11**(5) (1996) 2543–2547.
85. J. D'Ignazio, Microwave power helps ignite plasma processing, *Semiconductor International* (November 1995), pp. 99–104.
86. M. Sternheim, W. van Gelder, and A. Hartman, A laser interferometer system to monitor dry etching of patterned silicon, *J. Electrochem. Soc.: Solid-State Science and Technology*, **130**(3) (1983) 655–658.
87. M. Welch, P. Luscher, Z. Sui, and B. Lee, Breaking the 0.5 percent exposed area endpoint barrier, *Semiconductor International* (July 1996), pp. 269–273.
88. H. Litvak, End point control via optical emission spectroscopy, *J. Vac. Sci. Technol.*, **B14**(1) (1996) 516–520.

89. S. Henck, In process thin film thickness measurement and control, *Proceedings of the SPIE Symposium on Process Module Metrology, Control and Clustering*, **1594** (1991) 216–221.
90. V. Patel, J. H. Thomas, and B. Singh, Reactive ion etching endpoint determination by plasma impedance monitoring, *Applied Phys. Letters*, **61**(16) (1992) 1912–1914.
91. H. C. Sun, V. Patel, E. Whittaker, B. Singh, and J. Thomas, III, Measurements of neutral species in low pressure C_2F_6 discharge using diode laser absorption spectroscopy, *J. Vac. Sci. Technol.*, **A11**(4) (1993) 1193–1197.
92. R. Allen, R. Moore, and M. Whelan, Application of neural networks to plasma etch end point detection, *J. Vac. Sci. Technol.*, **B14**(1) (1996) 498–503.
93. G. Gibson and D. Hemker, Integrated stack etching using a helicon plasma source, *Semiconductor International* (July 1996), pp. 193–200.
94. A. S. Bergendahl, D. Horak, P. E. Bakeman, and D. Miller, Cluster tools, Part II: 16Mb DRAM processing, *Semiconductor International* (September 1990), pp. 94–102.
95. H. Shan, B. Pu, H. Gao, K. Ke, J. Lewis, M. Welch, and C. Deshpandey, Process kit and wafer temperature effects on dielectric etch rate and uniformity of electrostatic chuck, *J. Vac. Sci. Technol.*, **B14**(1) (1996) 521–526.

CHAPTER 5.3

Ion Beam Technology

Michael J. Powers
Commonwealth Scientific Corporation

5.3.1 INTRODUCTION

Ion beam processes, broadly defined, include a wide range of unique applications. The ability to control independently the beam's energy, size, and mass over many orders of magnitude gives these processes a large amount of flexibility. Beam energies can range from up into the MeV, as routinely used in Rutherford backscattering spectroscopy (RBS) for compositional analysis of materials [1], and down to just a few eV for plasma processing [2]. Beam sizes can range from micrometers (μm), in focused ion beam systems [3], up to tens of centimeters in ion implanters [4] and broad ion beam technology as discussed in this chapter. Ion masses can also range from that of a proton beam (ionized hydrogen) up to, in theory, that of the heaviest element on the periodic table. The scope of this chapter is limited to broad, low-energy ion beams (50 eV–2 keV) used for etching and deposition of thin films in research and production environments.

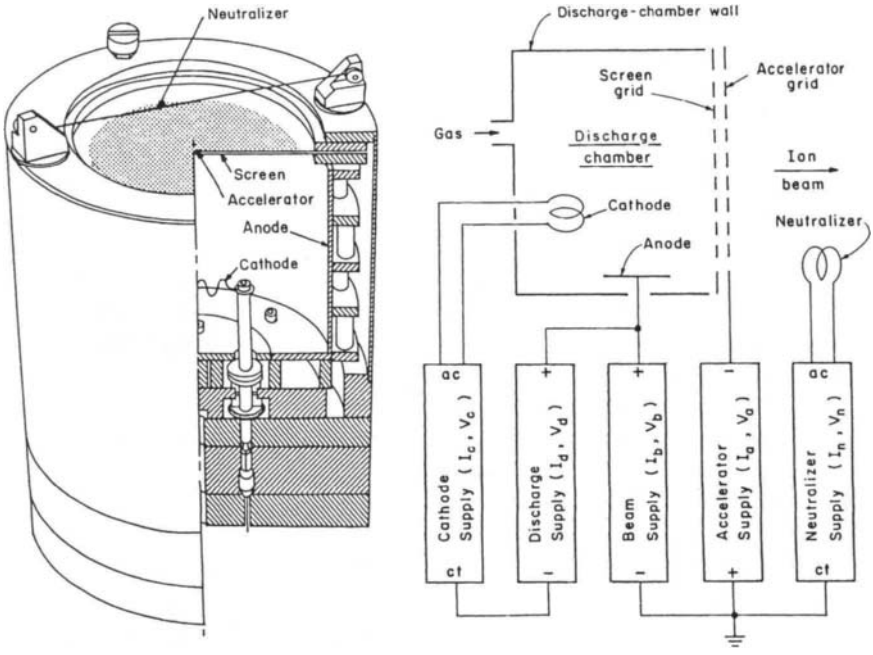
The broad beam ion source was originally conceived in NASA's electric space propulsion program [5], and active research continues for this application [6]. The thin-film industry has since adapted the technology to meet needs in a wide range of applications. Industrial broad beam ion sources consist of two main components: a discharge chamber where a plasma is formed, and some form of extraction optics that accelerates and directs positive ions out of the plasma toward a substrate or target. A major distinction between ion beam processing and other

ISBN 0-12-325065-7
$25.00

plasma processes is that the plasma is maintained remotely from the process chamber. Diameters of the optics (and the beams) in commercial sources range from 1cm up to about 40cm [7] in diameter. The two most widely used methods of plasma generation define two distinct types of ion soures: the Kaufman direct current (DC) ion source [8] and the radio frequency (RF) ion source. The optics for each are nearly identical and in many cases interchangeable.

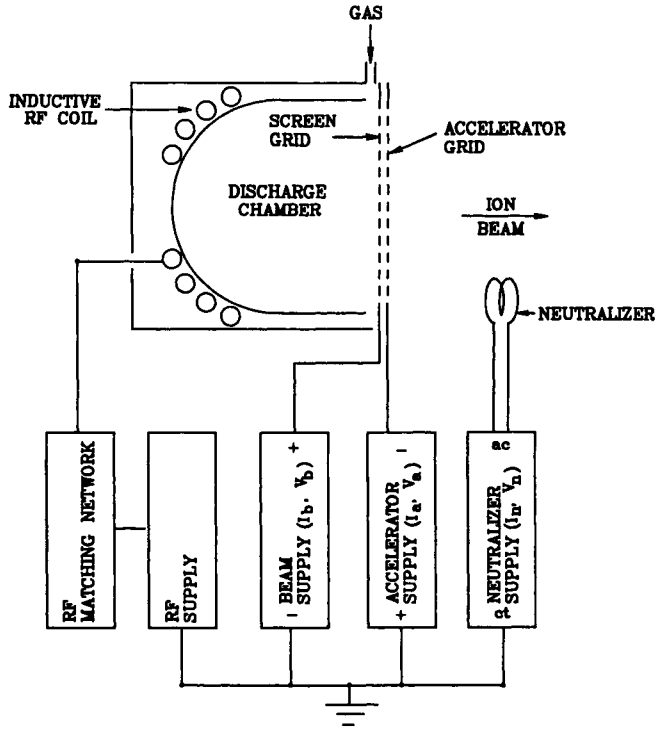
In a Kaufman ion source, as shown in Figure 1, thermionic cathodes (heated tungsten filaments) emit electrons that are accelerated within the discharge chamber toward the anode by DC voltage. A plasma is formed by electron impact ionization of gas (typically argon) atoms or molecules in the discharge chamber. Gas flow of 1–50 sccm (depending on the source size and system pumping) maintains the discharge chamber at a pressure of about 1 mtorr. The extraction optics limit the flow conductance and allow the system pressure to be in the low 10^{-4} torr range, several orders of magnitude less than typically found in plasma processing systems [2]. Electromagnets or permanent magnets are often used to form a weak magnetic field (10–100 G) within the source to enhance the plasma density and uniformity by altering the electron trajectories as they accelerate toward the anode.

Fig. 1.



(a) Kaufman DC ion source, cut-away view, and (b) schematic diagram of power supply configuration.

Fig. 2.



RF inductive ion source configuration.

Figure 2 depicts a typical configuration for an RF-type ion source. The plasma is generated by accelerating the electrons with an RF electromagnetic field instead of the anode–cathode DC voltage present in a Kaufman ion source. The frequency most commonly used is 13.56 mHz due to the wide availability of commercial RF generators at this frequency. Other frequencies are also used, from several hundred kHz to the GHz microwave regime, such as in electron cyclotron resonance (ECR) sources. In most designs, a coiled RF antenna inductively couples the RF energy to the plasma through a dielectric window. A magnetic field can also be used in these sources to enhance the plasma.

Several practical differences must be considered when choosing a DC or RF ion source. The major advantage of using RF to excite the plasma is that reactive gases can be easily handled that would otherwise poison the electron emission from the filaments in a DC source. Maintenance issues, however, are not eliminated by the fact that there are no filaments to change. Backsputtered material from the system chamber, target, and ion source components coat the dielectric window and tend to block the RF energy, making it necessary to clean the window at regular intervals. In certain applications, such as sputter depositing or etching

metals, the amount of backsputtered metal can easily make maintenance on an RF ion source equal to if not exceeding that of a DC source.

Once a plasma has been formed, two or three closely spaced grids are generally used to extract a controlled beam. The grid material is typically pyrolytic graphite or molybdenum, depending on the application. Graphite is generally used in a reactive gas environment and for smaller ion sources, but should be avoided if oxygen will be used, since the graphite will be etched quickly. The high strength and stiffness of molybdenum makes it the material of choice for larger sources in a production environment. The grids can be configured for a divergent, collimated, or focused ion beam by offsetting the hole patterns in the grids and, in the case of molybdenum, dishing the grids in a concave or convex manner to provide the desired beam shape. Divergent and collimated grids are selected for etching applications where beam uniformity is of prime concern. Focused grids are generally chosen for ion beam deposition applications where the beam is to be focused and contained on a target.

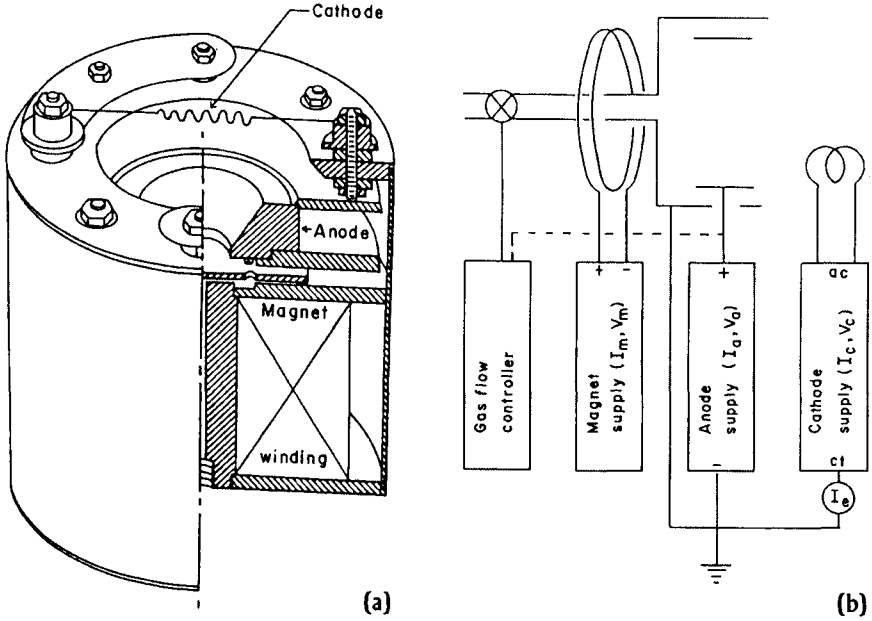
The two figures of merit that determine the major processing characteristics of an ion beam are the beam energy and current. The kinetic energy of the ion beam is the potential difference between the plasma and ground, which is set by the beam voltage power supply. The beam current or ion flux is determined by the intensity of the plasma in the source. In the DC source the plasma density is controlled by the discharge current, and in the RF source it is controlled by the RF power. It is important to note that these two parameters can be independently varied over a wide range.

5.3.1.1 Gridless Sources

Another type of broad beam ion source that is used mainly in ion-assisted deposition and precleaning applications is the end hall ion source [9] shown in Figure 3. This type of ion source was originally developed in the former Soviet Union's space propulsion research program to overcome the current density limitations at low energies imposed by the extraction optics in a gridded ion source. The accelerating potential for the ions in the end hall source is developed by a circulating electron current in a magnetic field. A high-curie point permanent magnet provides the magnetic field in current commercial designs, allowing the source and power supply to be simple and rugged. The electron source can be either a heated tungsten filament or a hollow-cathode-type electron source, which is discussed later.

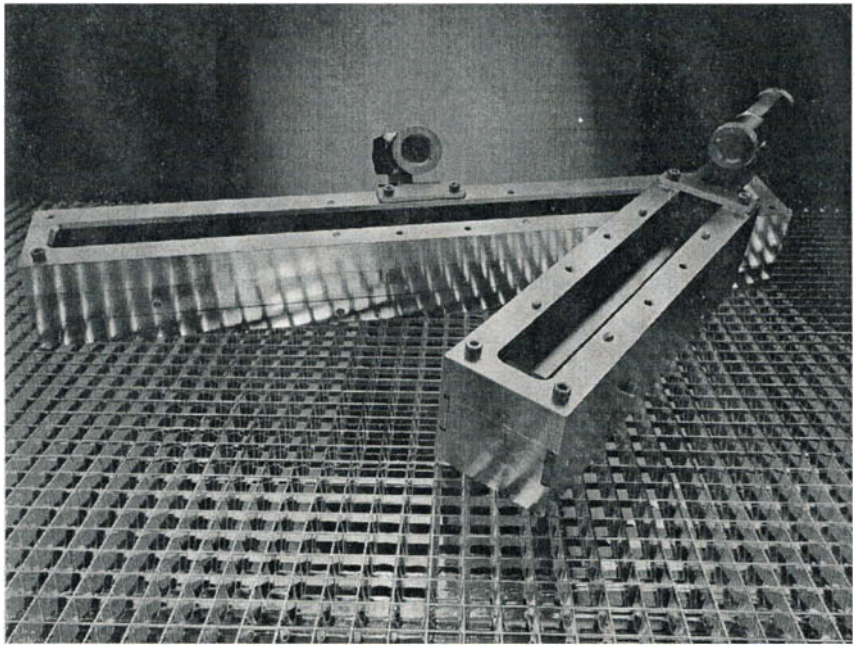
Typical beam characteristics are a highly divergent (~45° half angle) beam, high current densities (several mA/cm^2 at the source), and low ion energies (30–120 eV). In addition to the circular end hall sources as shown in Figure 3, the sources can be constructed in a linear configuration for use in continous feed applications. A 20 cm and 50 cm linear source are shown in Figure 4 and sources up to 1 meter in length are commercially available.

Fig. 3.



(a) End hall ion source cut-away view and (b) schematic diagram of power supply configuration. Note that the magnet supply is often eliminated by using permanent magnets.

Fig. 4.



50 cm and 20 cm linear end hall ion sources with hollow cathode electron sources.

5.3.1.2 Ion Beam Neutralization

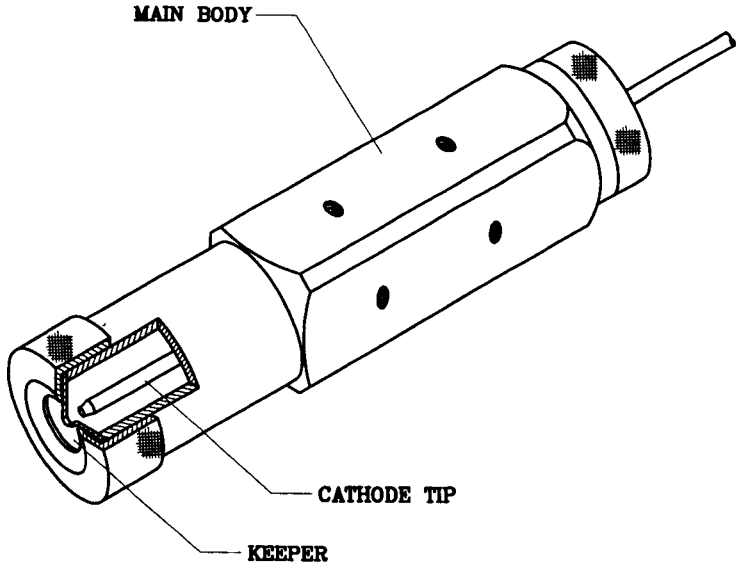
As the ion beam transits from the ion source to the target or substrate, there is an intense space charge repulsion between the positive ions which will tend to disrupt the beam. By putting a source of electrons nearby, a net charge of zero per unit volume (charge neutralization) can be achieved in the beam; ie., equal densities of positive and negative charges. In addition to charge neutralization of the beam, sufficient electrons must be supplied to achieve total current neutralization at the substrate, i.e., equal arrival rates of positive and negative charges. If the target is an insulator or is electrically isolated from ground, insufficient current neutralization will result in positive voltages building up on the surface. If an intentional source of electrons is not provided, the surrounding surfaces in the system will supply them in the form of short arcs and flashes, which can be observed through a viewport, if available. This uncontrolled arcing can cause damage, either directly if the arc is from the target or substrate, or indirectly when particles, ejected from an arc, land on sensitive surfaces. Voltage fluctuations in the ion beam due to inadequate neutralization can cause variations in the process during the run and from run to run. To assure that neutralization is occurring, the electron emission current is generally set to a value of about 10% over the ion beam current set point.

An external source of electrons is only necessary for the gridded ion sources. Since the electron emitter for the gridless ion sources is external by design, it supplies the electrons to both form the plasma and to neutralize the ion beam.

Several methods for neutralizing the ion beam are commonly used. The simplest approach is to immerse in the ion beam a heated tungsten filament that will thermionically emit enough electrons to cancel out the positive charges in the beam. The placement or symmetry of the filament in the beam is not critical, since the mobility of the electrons is orders of magnitude larger than the more massive ions. The filament is sputtered by the ion beam, which limits its lifetime, and the sputtered filament material may present a problem if contamination of a depositing film is critical. In this case two other types of neutralizers are available, the plasma bridge and the hollow cathode neutralizers.

In the hollow cathode [10], the electron emission is from a low function work function insert as shown in Figure 5. The emission is primarily thermionic due to the high temperature of the insert, although it is usually enhanced by strong electric fields at the insert surface. The insert is heated by ion bombardment from the internal plasma. The plasma couples to the surroundings through the small opening at the end of the hollow cathode. The internal pressure to the hollow cathode is in the 1 – 10 torr range, and the electron emission surface is sensitive to contamination, thus the gas flowing through the tip must be a high-purity noble gas such as argon, krypton, or xenon.

Fig. 5.



Hollow cathode electron source (HCES).

The plasma bridge concept [11] is very similar to the hollow cathode, except that the electron emitter is a heated tungsten filament enclosed inside the discharge chamber, which has a small aperture. Appropriate bias voltages are placed on the device such that a flood beam of electrons can be extracted from the plasma bridge. Both the hollow cathode and the plasma bridge devices are generally placed within a few centimeters of the edge of the beam to ensure proper neutralization coupling. Gas flows are typically on the order of a few standard cubic centimeters per minute (sccm) in both the hollow cathode and plasma bridge.

5.3.2

ION BEAM ETCHING

Ion beam sputtering is fundamentally a momentum transfer process between the incident ions and the target atoms. The sputter yield is the number of sputtered atoms per incident ion and is a function of the ion energy, ion species and mass, target atom species and mass, and incidence angle [12–15]. To eject an atom from the target surface, kinetic energy greater than the chemical binding energy of the

atom must be imparted to the atom. Ion energy thresholds to impart this amount of energy range from 10 to 100 eV [13], depending on the material. The sputter yield from material to material follows the trend of binding energy within the periodic table and increases with ion beam energy up to about 10 keV where implantation starts to dominate [12].

The etch rate $R(\theta)$, can be written as a function of sputter yield $S(\theta)$:

$$R(\theta) \cong 9.6 \cdot 10^{25} \cdot J \cdot \frac{S(\theta)}{n} \cdot \cos(\theta) \text{ in angstroms/min} \tag{1}$$

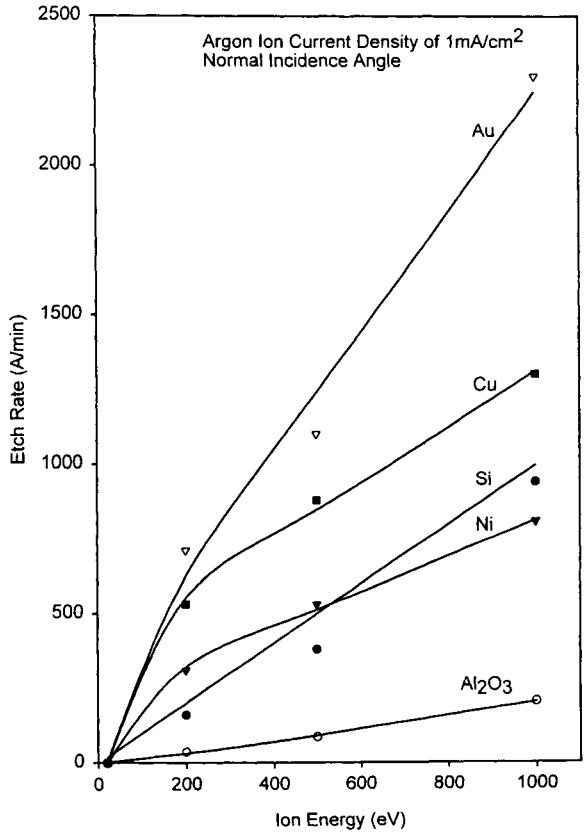
where J is the current density in mA/cm^2
n is atomic number density $atoms/cm^3$
θ is the angle of incidence from normal

As θ increases from zero, the sputter yield increases, due to the more efficient momentum transfer between the ions and the target atoms. The sputter yield as a function of angle, $S(\theta)$, largely depends on the element being sputtered, and for some materials, such as gold, the angle dependence practically disappears [14]. As θ increases beyond 50–60°, the $\cos(\theta)$ term in Equation (1), which accounts for the reduction in ion flux density on the surface, begins to dominate and the etch rate falls off to essentially zero at 90°. Surface roughness can affect the sputter yield by giving an effective incidence angle on a microscopic level, which is different from the macroscopic-level incidence. Etch rates of several materials vs. incident beam energy and incident beam angle are shown in Figures 6 and 7. Table 1 lists some approximate etch rates of common materials with an Argon ion beam.

Pattern definition with ion beam etching is accomplished by using a mask, usually made of photoresist due to the ease of application and removal. Many types of photoresist are available, with AZ 1350H and Kodak 5214 being the most common. Line widths under 100 Å can be generated with ion beam etching [16], and the mask quality is what limits the achievable resolution in all practical applications.

Since ion beam etching is fundamentally a momentum transfer process and only a small fraction of the ion kinetic energy is carried away by the sputtered atoms, cooling of the substrate is of concern and often is the limiting factor in determining the maximum etch rate in a given system. An estimate of the power density (in W/cm^2) at the substrate can be obtained by multiplying the ion beam energy (in eV) by the current density (in A/cm^2). Most patterns are delineated by photoresist, which generally has a maximum service temperature of about 100°C. Significant heat radiation ($1W/cm^2$) does not occur until a blackbody reaches 300–400°C so the heat transfer must be made by conduction or convection. Since the ions bombard the front side of the wafer, the backside of the substrate must be

Fig. 6.



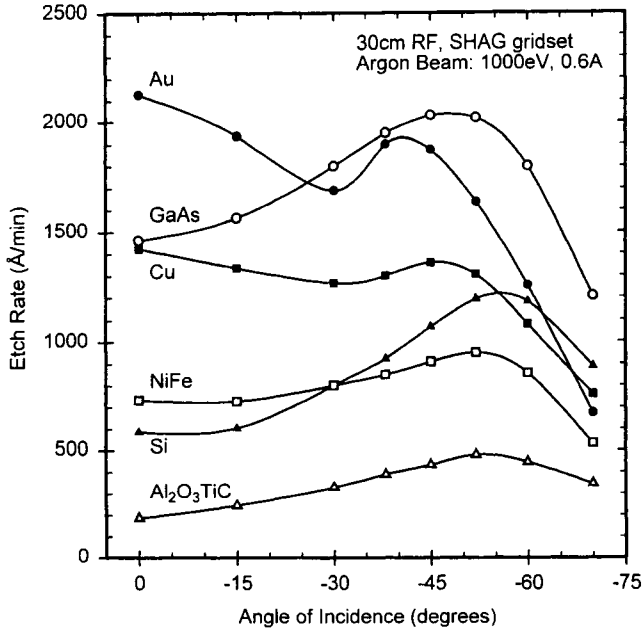
Etch rates of various materials vs. ion beam energy.

used to remove the heat. Some common methods (and approximate power levels of usefulness) are

- Thermally conductive pastes ($\sim\frac{1}{4}$ W/cm^2)
- Thermally conductive elastomer ($\sim\frac{1}{2}$ W/cm^2)
- Low-melting-temperature alloys such as indium and gallium ($\sim$1 W/cm^2)
- Static or flowing helium gas ($>$1 W/cm^2)

State-of-the-art substrate cooling is currently done with the last method, using low-pressure (5–50 torr) He gas as the heat transfer medium between the wafer backside and the water-cooled substrate holder. This technique allows power levels of over 2 W/cm^2 to be used.

Fig. 7.



Etch rates of various materials vs. ion beam incidence angle.

Table 1

Etch Rates of Some Materials with an Argon Ion Beam at Normal Incidence. The current density is 1 mA/cm^2 in all cases.

Material	200 eV	500 eV	1000 eV
C	13	44	80
Si	160	380	700
Fe	260	530	660
Ni	310	660	810
Cu	530	1000	1200
Mo	240	540	730
Au	710	1700	2300
SiO_2	—	400	520
Al_2O_3	30	85	180
GaAs (110)	780	1600	—
GaP (111)	690	1600	—

The ions do not stop at the surface but penetrate approximately 20–100 Å deep in the energy range 100–1000 eV. Damage in ion-etched semiconductors has been studied [17–19] in Si, GaAs, and other electronic materials, and damage thresholds ranging from 50 eV up to several hundred eV have been found. Although generally detrimental in semiconductor devices, the damaged layer on the surface can sometimes be beneficial in forming good ohmic contacts.

The most common gas species used for nonreactive ion beam etching is argon, due to the combination of low cost and relatively high atomic mass. Etching rates can be increased for some materials by using a reactive gas to generate the plasma and extracting the ionized reactants. This process is termed *reactive ion beam etching* (RIBE). The reactive gas can also be introduced near the substrate, chemisorbed onto the surface, and an inert ion beam can provide the energy to complete the chemical reaction. This is referred to as *chemically assisted ion beam etching* (CAIBE). In a purely chemical etching system, compound materials often cannot be successfully etched, because of the widely different reactivity and vapor pressures of the component elements. Complex gas mixtures are often used in reactive ion etching (RIE) plasmas to achieve the desired etch selectivity and anisotropy [20]. In ion beam etching, there is always a physical component to the material removal mechanism, and many of the complications found in RIE processing disappear.

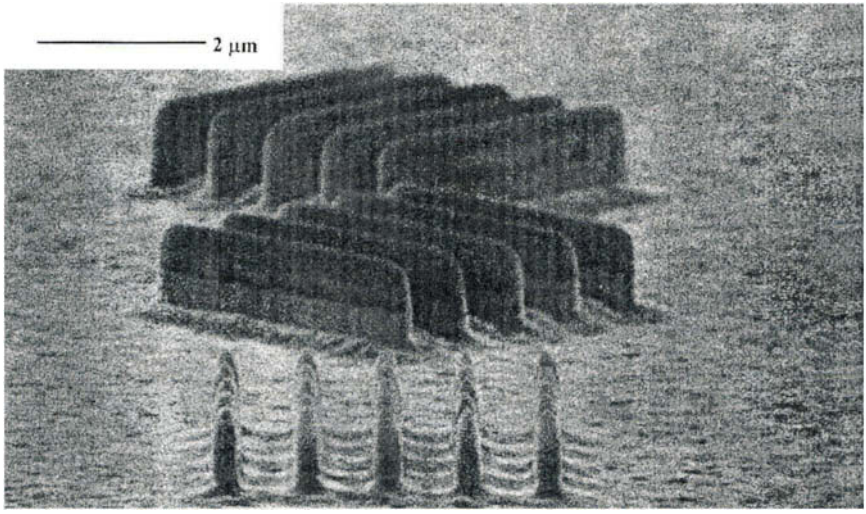
Etch rate increases of 50:1 over inert ion milling can be achieved on a variety of materials using the RIBE and CAIBE techniques [21–29]. The benefits of chemically enhanced ion milling such as higher rates, etch selectivity, and profile control have to be weighed against the added cost and complication of handling reactive gases such as Cl_2 and CF_4 as opposed to pure argon. It has been shown that submicron features in GaAs can be etched with aspect ratios of over 40:1 using an Argon ion beam and Cl_2 gas introduced at the substrate surface [25]. CAIBE techniques appear promising [30,31] in efforts to etch III-V nitrides, of interest in new optoelectronic devices, where chemical and RIE methods have met with resistance.

Argon ion beam etching has shown some unique capabilities in optical applications. These range from formation of microlenses on InP and Si for use in infrared devices [32,33] to ion beam "figuring" of ultra-high-quality optical surfaces over large areas where the limits of mechanical polishing are met [34].

Ion milling is also useful for use in etching of complex compounds such as $LiNbO_3$ waveguides [35] and Y-Ba-Cu-O superconductor devices [36,37] where it is difficult to determine the optimum process in a pure chemical etch. In addition to compound materials, ion beam milling is often used to etch multilayer stacks such as Ti/Pt/Au in a single process step without requiring parameters such as gas mixture to change from one material to the next.

A major application of ion beam sources is in the "precleaning" of substrates prior to a deposition process. The ion source allows the highly controlled removal of the top few monolayers on a surface that may contain impurities and/or con-

Fig. 8.



0.25 μm test pattern etched in gold by argon ion beam etching. Foreground of the image contains 0.25 μm square posts. Etch rate was ~1500 Å/min.

taminants. Dramatic improvements in properties such as the adhesion and electrical contact of the subsequently deposited layers have been observed [38].

Ion beam etching is used in almost every production line in the magnetic storage industry. The applications range from etching of compound magnetic materials such as Ni, Fe, Co, Cr alloys in the ever-shrinking read/write heads, to the formation of the aerodynamic air-bearing surfaces in Al_2O_3TiC composites [39–41].

Microsharpening by ion beam etching has been shown to allow the formation of ultrasharp tips in molybdenum [42] and diamond [43] with radius of curvatures less than 10 Å. These tips have found application in surface profilometer tips, phonograph needles, hardness-testing equipment, and scanning tunneling microscopes.

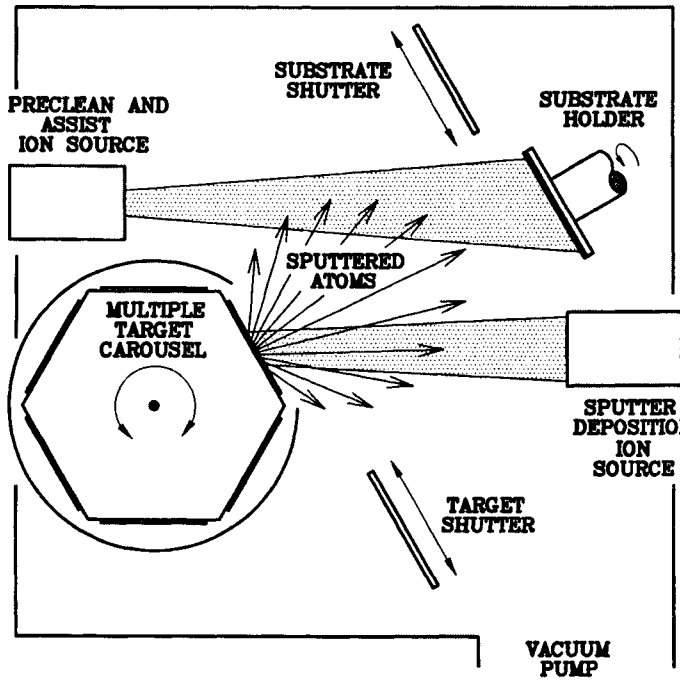
Ion beam etching is often chosen to etch gold due to its high sputter yield and low chemical reactivity. Etch rates into the 1000's of Å/min can be achieved using inert argon ion beams. Figure 8 shows a submicron test pattern of gold lines and posts etched with an argon ion beam.

5.3.3

ION BEAM SPUTTER DEPOSITION

In ion beam sputter deposition, the ion beam is directed at a target and sputtered atoms are deposited onto a nearby substrate surface. A typical ion beam deposi-

Fig. 9.

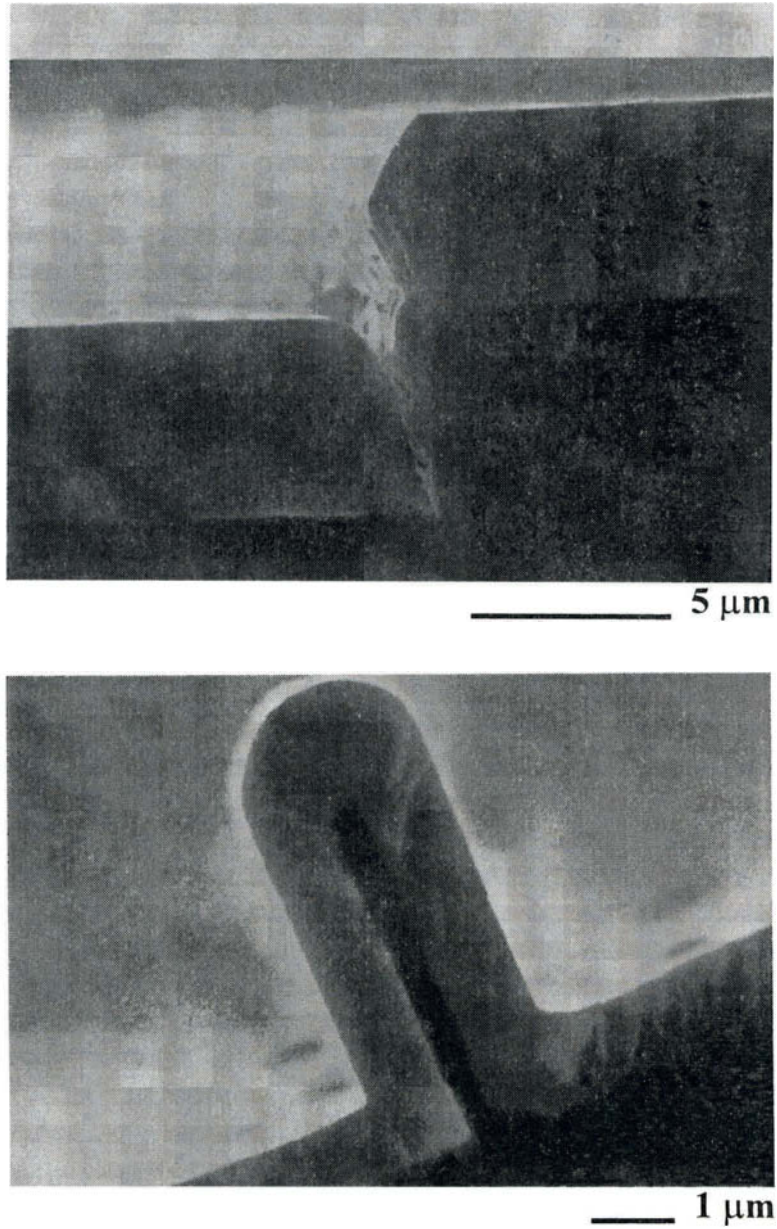


Typical configuration of ion beam sputter deposition (IBSD) system with ion-beam-assisted deposition (IBAD) capability.

tion sytem layout is shown in Figure 9. Several properties of ion beam deposition, such as the energetics involved, distinguish it from other deposition methods.

The energy of the sputtered atoms as they leave the target ranges from a few up to tens of eV [45]. Due to the low-pressure regime of ion beam processing (10^{-4} torr), where the mean free path is on order of a meter these sputtered atoms arrive at the substrate with most if not all of their original kinetic energy. This is in contrast to the higher-pressure regime of magnetron sputtering, where the sputtered atoms thermalize with the background gas before reaching the substrate. This lack of scattering has two direct consequences. First, the higher kinetic energy with which the atoms arrive imparts a high degree of mobility to the growing film, yielding good coverage at very small thicknesses, smoother films [46], and enchanced microstructure. Second, the sputtered atoms maintain their direction vector from the target to the substrate, allowing control over the atom's incidence angle at the substrate, which can give excellent coverage of high-aspect-ratio geometries. Figure 10(a) shows a typical step coverage result from magnetron sputtering and Figure 10(b) shows step coverage of the same film material deposited by ion beam sputtering.

Fig. 10.



(a) Typical step coverage for magnetron sputter deposition and (b) step coverage using ion beam sputter deposition. Substrate material is SiO_2 and deposited film is "Sendust" (85% Fe/ 10% Si/5% Al) in both (a) and (b).

Energy is also added to the growing film by reflected neutrals and has been studied extensively [45,47,48]. The higher energetics of ion beam sputtering makes substrate heating unnecessary where it is often required in other thin-film deposition systems to obtain the desired film microstructure.

The amount of energy carried to the substrate by the sputtered atoms and reflected neutrals can be controlled over more than an order of magnitude by adjusting the beam energy and the ion species. Herein lies the true elegance of ion beam processing, where the process parameters can be independently controlled over wide ranges to optimize the desired film properties. In glow discharge-based deposition systems such as magnetrons, it is difficult if not impossible to decouple these variables, which are interlinked in a complex manner [49,50].

The major sources of contamination in ion beam deposition are from lack of beam containment on the target and sputtered material from the ion source. Proper design of the system geometry and proper ion source operation [51] should result in contamination levels less than $1:10^4$ for transitional metals. Target contamination is the highest from the immersion-type neutralizer and is minimized with the plasma bridge and hollow cathode emitters.

Until recently, ion beam sputtering has often been characterized as useful only for processes requiring low deposition rates and small substrate sizes. While low deposition rates ($\sim$1 Å/s) are often desirable to control deposition of ultra-thin films, Fe deposition by ion beam sputtering has recently been demonstrated [44] at $>$20 Å/s over 200 mm diameter wafers with excellent uniformity. In general, the deposition rate from material to material follows the same trend as etch rate, because they both are derived from the sputter yield.

Al_2O_3 films deposited with ion beam deposition have been shown to have superior properties to e-beam and sputtered films [52]. These properties include film stress an order of magnitude less than RF-sputtered films, nearly double the Al_2O_3 electrical breakdown strength, and higher refractive index.

A current area of intense research is the deposition of new types of magnetic thin films for higher-density recording. Soft magnetic materials with high magnetization saturation and permeability such as Fe-N and Fe-C are of interest for use in high-flux-density write heads and have been studied using ion beam deposition [53,54] because of the advantages previously discussed, especially for conformal coverage of irregular topologies.

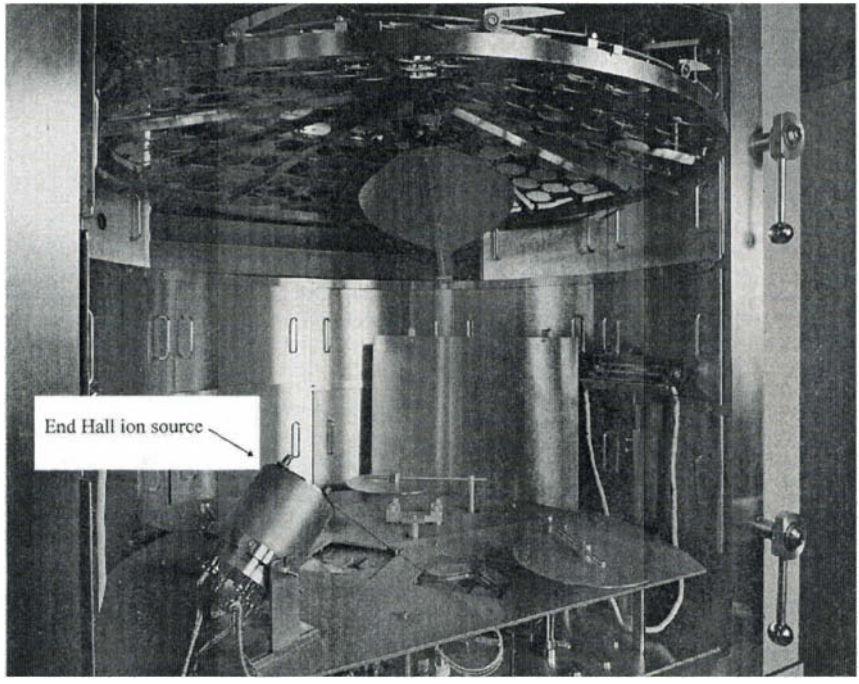
Thin-film magnetoresistance (MR) read heads are the current state of the art in digital magnetic storage and has been studied extensively by ion beam sputtering [55–57]. The MR effect is the change in electrical resistance of a material in the presence of a magnetic field and is used in the read head in magnetic storage devices. Giant magnetoresistance (GMR), dicovered in 1988 [58], has been observed in ultra-thin ($\sim$10 Å) multiple (5–20) layer stacks of magnetic and nonmagnetic materials [59–66] known as spin valves. These types of structures will likely be used in the next generation of magnetic storage devices. The ability of ion beam

sputtering to put down these multilayer stacks in a precise, controlled fashion by use of multiple targets is making it an attractive tool for research and possible production.

5.3.4 ION-BEAM-ASSISTED DEPOSITION

Many advantages due to the energetic condensation of the growing film found in ion beam deposition can also be obtained by bombarding the growing film with an energetic ion beam. This process is known as ion-beam-assisted deposition (IBAD) and can be used in conjunction with a variety of deposition methods, including ion beam sputtering, electron beam evaporation, molecular beam epitaxy, and pulsed laser deposition. A typical evaporation system layout with IBAD capability is shown in Figure 11.

Fig. 11.



End hall ion source installation in commercial e-beam evaporation system for substrate precleaning and ion-beam-assisted deposition (IBAD).

Several excellent reviews of IBAD have been recently written [67–71]. The two parameters in ion-assisted film growth that most influence the film properties are the ion energy and the ion-to-atom arrival rate ratio. As the energy and ion/atom arrival rates are varied, different mechanisms come into play during film growth. At low energies of 1–10 eV, surface desorption of adsorbed and chemisorbed background gas is dominant. At this energy level, some displacement of surface adatoms also occurs, which can affect surface morphology. The energy range most commonly used in IBAD is 50–500 eV where significant enhancement of film properties is achieved. The dominant mechanism in this regime is often termed "densification" due to the number of altered film properties that are related to film density. At still higher energies, effects from resputtering of deposited atoms and ion implantation will become more important. As the ion/atom arrival rate is varied, different mechanisms can be brought into play. For example, the amount of densification occurring at 20 eV for an ion/atom arrival ratio of 0.1 will be much less than that for a ratio of 1. In general it appears that high ion/atom ratios combined with low ion energy are most suited to minimize implantation and below-surface defect production [71].

Evaporation alone tends to produce porous films with columnar grain structure and is often supplemented with an ion assist source. Higher-temperature processing of the substrate can improve these properties, but the disadvantages associated with heating the samples to several hundred degrees Celsius make this unattractive in production systems. With the substrate at near room temperature, the arriving atoms have a sticking coefficient near unity and very low surface mobility due to their low (~0.1 eV) thermal energies. As the ions collide with the surface, loosely packed clusters of atoms will be dissociated and given enough energy to diffuse along the surface and approach a more ideal packing density.

Intrinsic stress can develop in a growing film due to chemical mechanisms such as the completion of oxidation reactions occurring below the surface and microstructure changes caused by defect generation or relaxation below the surface. The stress induced chemically can be either compressive or tensile, depending on the reaction, but the evolution of defects and microvoids generally results in tensile stress. By bombarding the growing film with low-energy ions, compressive stress can be introduced at the surface by forcing the atoms into closer proximity to each other than in their relaxed state. This is similar to the ball peening that is performed in metal working to induce compressive stress. This effect is sometimes referred to as "ion peening" and is used to counter the intrinsic tensile stress that develops in the growing film. Proper adjustment of the ion beam parameters can effectively neutralize the film stress.

This destabilization of surface atom clusters promotes step flow growth and much smoother surface morphologies can be obtained [72]. This mechanism also enhances the ability to obtain epitaxial growth at low substrate temperatures. Crystallographic texturing of the growing film can be controlled in some cases

with IBAD by preferentially sputtering away one crystalline orientation. This has recently become an important application in the deposition of superconducting films. In this process, during the deposition of a yttria-stabilized-zirconia (YSZ) buffer layer, IBAD is used to achieve in-plane texturing of the depositing YSZ. High-quality superconducting films can be grown on single-crystal YSZ, which acts as a seed layer. This texturing during IBAD allows the YSZ seed layer to be deposited on a variety of materials prior to depositing the superconductor film [73,74].

The control of film density has a large impact on properties such as refractive index, hardness, and corrosion resistance. Optical coatings produced by IBAD have higher refractive indexes and better temporal stability in air than those produced by other techniques. Electron beam evaporation coater manufacturers now offer IBAD as either standard or optional equipment on their systems [75].

When depositing oxides such as Al_2O_3, TiO_2, Ta_2O_5, or SiO_2, a low-energy oxygen ion assist can be used to maintain proper stoichiometry of the deposited film. Reactive IBAD can be used to deposit nitrides such as Si_3N_4, TiN, or BN where the atomic element is deposited with simultaneous nitrogen ion beam bombardment to complete the chemical reaction. Novel crystalline compounds such as cubic-BN [76] and C-N [77], both of which have hardnesses on par with diamond, have been studied by using IBAD.

Uniform coverage of irregular surface features such as was done in Figure 10 with ion beam deposition can also be performed with IBAD. Proper choice of ion incidence angle on the sample can redistribute the deposited atoms over the features and resputter the high spots of deposition.

In summary, IBAD allows the production of high-quality films by providing the user with a very controllable method of changing the energetics at the surface of the growing film. Both gridded and gridless ion sources are used for this application, with gridless sources ideally suited due to their high current output at low ion energies.

5.3.5 ION BEAM DIRECT DEPOSITION

The use of broad beam ion sources to deposit diamondlike carbon (DLC) has recently found application in several areas. In this process, a carbon-carrying gas such as methane is fed into the ion source and a low-energy (50–400 eV) beam is directed at the substrate, forming a DLC film [78–82]. The term "direct deposition" is used since no sputtering target is present. Table 2 summarizes some properties of DLC films grown by this method and compares them with diamond and graphite. The films generally have high hardnesses, extremely low static and dy-

Table 2

Properties of Diamondlike Carbon (DLC) Deposited by Direct Ion Beam Deposition. Values for pure diamond and graphite are listed for comparison.

Property	Diamond	DLC	Graphite
Hardness (GPa)	100	15–30	—
sp^3 Bonding (%)	100	40–50	0
Hydrogen Content (%)	0	30–50	0
Optical Gap (eV)	5.45	1.2–4.0	0
Coefficient of Friction	0.05	0.05–0.1	0.1
Refractive Index	2.4	2	—

namic friction coefficients, high transparency to the infrared and visible spectrum, and high indexes of refraction. Low deposition rates, similar to those found in ion beam sputtering, characterize this method. Since the power density seen by the substrate in this method is low, near-room-temperature film deposition is easily achieved with minimal cooling. This is often not the case in CVD plasma reactors where the substrate is immersed in the plasma. The DLC direct deposition process can be done with both the end-hall-type sources and the gridded DC and RF sources.

Currently this process is widely used to coat read/write heads for the magnetic storage industry to provide a hard, low-friction surface that can also act as a chemical diffusion barrier to prevent corrosion of the head. Since the flying heights (head–media gap) of the head in the hard disk arena are currently on the order of 500 Å, the DLC coating is on the order of 100 Å or less and getting thinner as flying heights decrease.

Another growing application for DLC coatings by direct ion beam deposition is for coatings in the ophthalmic industry. Coatings on optical components from sunglasses to supermarket scanner windows have been successfully coated with DLC to provide a highly transparent scratch-resistant coating.

5.3.6 CONCLUSION

A broad range of ion source types and sizes are available commercially for a number of ion beam processes. Due to the wide scope of ion beam technology, it is not possible to cover many of the other types of ion beam sources and applications. Broad ion beam technology applied to thin film etching allows a range of material compositions and feature geometries to be etched in the same process

due to its inherently physical nature. A reactive component can be added to this process to enhance the rate and selectivity of the etch in a RIBE or CAIBE process.

Independent control of the ion beam parameters allows the isolation of key figures of merit in thin-film deposition processes. This is in contrast to the interlinking of parameters in other methods such as RF diode sputtering, where it is difficult to determine the importance of various mechanisms. This variable decoupling allows precise control over the film properties. The increased energy of the film growth in ion beam sputter deposition and ion-assisted deposition often can yield superior results in a variety of properties over other deposition methods.

REFERENCES

1. W. K. Chu and G. Langouche, *MRS Bull.* January (1993), p. 32.
2. Donald L. Smith, Thin film deposition: Principles and practice (McGraw-Hill, New York, 1995), pp. 453–555.
3. Proceeding of 2nd Japan–U.S. Seminar on Focused Ion Beam Technology and Applications, *J. Vac. Sci. Technol.*, **B9** (1991) 2461–2732.
4. *Handbook of Ion Implantation Technology*, edited by J. F. Ziegler (North-Holland, Amsterdam, 1992).
5. H. R. Kaufman, *NASA Technical Note* TND-585 (NASA, Washington, D.C., January 1961).
6. Charles E. Garner, *NASA Tech Briefs* (May 1995), p. 48.
7. H. R. Kaufman, W. E. Hughes, R. S. Robinson, and G. R. Thompson, *Nuclear Instr. and Meth. in Phys. Research*, **B37/38** (1989) 98.
8. H. R. Kaufman and R. S. Robinson, *American Institute of Aeronautics and Astronautics J.*, **20** (1982) 745.
9. Harold R. Kaufman, Raymond S. Robinson, and Richard Ian Seddon, *J. Vac. Sci. Technol.*, **A5** (1987) 2081.
10. J. Reece Roth, *Industrial Plasma Engineering*, vol. 1 (Institute of Physics Pub., Philadelphia, 1995), pp. 167–170.
11. P. D. Reader, D. P. White, and G. C. Isaacson, *J. Vac. Sci. Technol.*, **15** (1978) 1093.
12. P. Sigmund, *Phys. Rev.*, **184** (1969) 383.
13. E. Henshke, *Phys. Rev.*, **106** (1957) 737.
14. G. Wehner, *J. App. Phys.*, **30** (1959) 1762.
15. N. Laegreid and G. K. Wehner, *J. Appl. Phys.*, **32** (1961) 365.
16. A. N. Broers, W. W. Molzen, J. J. Cuomo, and N. D. Wittels, *Appl. Phys. Lett.*, **29** (1976) 596.
17. W. D. Sawyer, J. Weber, G. Nabert, J. Schmalzlin, and H.-U. Habermeier, *J. Appl. Phys.*, **68** (1990) 6179.
18. M. Miyao, T. Sukegawa, and M. Hagino, *J. Appl. Phys.*, **48** (1977) 1383.
19. J. M. Millunchick, L. Hultman, and S. A. Barnett. *J. Vac. Sci. Technol.*, **A13** (1995) 1155.
20. H.-C. Scheer, *Rev. Sci. Instrum.*, **63** (1992) 3050.
21. Ch. Steinbruchel, *J. Vac. Sci. Technol.*, **B2** (1984) 38.
22. R. B. Jackman, *Vacuum*, **44** (1993) 239.
23. M. Geis, S. W. Pang, N. E. Efremow, G. A. Lincoln, G. D. Johnson, and W. D. Goodhue, in *Handbook of Ion Beam Processing Technology: Principles, Deposition, Film Modification and Synthesis*, edited by Jerome J. Cuomo, Stephen M. Rossnagel, and Harold R. Kaufman (Noyes Publications, Park Ridge, New Jersey, 1989), pp. 219–240.

24. J. D. Chinn, I. Adesida, and E. D. Wolf, *Appl. Phys. Lett.*, **43** (1983) 185.
25. G. A. Lincoln, M. W. Geis, S. Pang, and N. N. Efremow, *J. Vac. Sci. Technol.*, **B1** (1983) 1043.
26. S. W. Pang, M. W. Geis, N. N. Efremow, and G. A. Lincoln, *J. Vac. Sci. Technol.*, **B3** (1995) 398.
27. S. W. Pang, G. A. Lincoln, R. W. McClelland, P. D. DeGraff, M. W. Geis, and W. J. Piacentini, *J. Vac. Sci. Technol.*, **B1** (1983) 1334.
28. Haruo Okano and Yasuhiro Horiike, *Japanese J. of Appl. Phys.*, **21** (1982) 696.
29. M. A. Bosch, L. A. Coldren, and E. Good, *Appl. Phys. Lett.*, **38** (1981) 264.
30. I. Adesida, A. T. Ping, C. Youtsey, T. Dow, M. A. Khan, D. T. Olson, and J. N. Kuznia, *Appl. Phys. Lett.*, **65** (1994) 889.
31. A.T. Ping, I. Adesida, and M. A. Khan, *Appl. Phys. Lett.*, **67** (1995) 1250.
32. R. Congxin, X. Deyuan, C. Guoliang, L. Xianghuai, and Z. Shichang, *Nucl. Instr. and Meth. in Phys. Res.*, **B96** (1995) 401.
33. O. Wada, *J. Electrochem. Soc.*, **131** (1984) 2373.
34. T. W. Drueding, S. C. Fawcett, Scott R. Wilson, and T. G. Bifano, *Optical Engineering*, **34** (1995) 3565.
35. Ren Congxin, Yang Jie, Zheng Yanfang, Chen Lizhi, Chen Guoliang, and Tsou Shichang, *Nuclear Inst. and Meth. in Phys. Res.*, **B19** (1987) 1018.
36. S. Matsui, N. Takado, H. Tsuge, and K. Asakawa, *Appl. Phys. Lett.*, **52** (1987) 69.
37. J. Gao, Y. Boguslavskij, B. B. G. Klopman, D. Terpstra, G. J. Gerritsma, and H. Rogalla, *Appl. Phys. Lett.*, **59** (1991) 2754.
38. J. Baglin, in Reference 23, pp. 279–297.
39. T. Nakanishi, T. Toshima, K. Yanagisawa, and N. Tsusuki, *IEEE Trans. on Magnetics*, **15** (1979) 1060.
40. T. Nakanishi, K. Kogure, T. Toshima, and K. Yanagisawa, *IEEE Trans. on Magnetics*, **16** (1980) 785.
41. T. Toshima, T. Nakanishi, and K. Yanagisawa, *IEEE Trans. on Magnetics*, **15** (1979) 1637.
42. S. Morishita and F. Okuyama, *J. Vac. Sci. Technol.*, **A9** (1991) 167.
43. I. Miyamoto, S. T. Davies, and K. Kawata, *Nuclear Inst. and Meth. in Phys. Res.*, **B39** (1989) 696.
44. M. J. Powers, personal communication, 1995.
45. H. Coufal, Harold F. Winters, H. L. Bay, and W. Eckstein, *Phys. Rev.*, **B44** (1991) 4747.
46. O. Auciello, S. Chevacharoenkul, M. S. Ameen, and J. Duarte. *J. Vac. Sci. Technol.*, **A9** (1991) 625.
47. E. Kay, F. Parmigiani, and W. Parrish, *J. Vac. Sci. Technol.*, **A5** (1987) 44.
48. Harold F. Winters, H. Coufal, C. T. Rettner, and D. S. Bethune. *Phys. Rev.*, **B41** (1990) 6240.
49. J. A. Thornton, *J. Vac. Sci. Technol.*, **11** (1974) 666.
50. J. A. Thornton, *Annu. Rev. Mater. Sci.*, **7** (1977) 239.
51. H. R. Kaufman, *J. Vac. Sci. Technol.*, **A4** (1986) 764.
52. M. Tan, *Data Storage*, **3**(1) (1996) 35.
53. T. Kobayashi, R. Nakatani, and N. Kumasaka. *J. of Magnetism and Magnetic Mat.*, **81** (1989) 259.
54. H. Jiang, Q. L. Wu, K. Tao, and H. D. Li, *J. Appl. Phys.*, **78** (1995) 3299.
55. Yasuhiro Nagai and Tomoyuki Toshima, *J. Vac. Sci. Technol.*, **A4** (1986) 2364.
56. Masaki Ueno and Shuji Tanoue, *J. Vac. Sci. Technol.*, **A13** (1995) 2194.
57. Yasuhiro Nagai and Masakatsu Senda, *J. Appl. Phys.*, **64** (1988) 4108.
58. M. N. Baibich, J. M. Broto, A. Fert, F. Nguyen van Dau, F. Petroff, P. E. Etienne, G. Creuzet, A. Friedrich, and J. Chazelas, *Phys. Rev. Lett.*, **61** (1988) 2472.
59. C. Kim, S. B. Qadri, H. Y. Yu, K. H. Kim, B. Maruyama, and A. S. Edelstein, *J. Vac. Sci. Technol.*, **A8** (1990) 1407.
60. C. Kim, S. B. Qadri, M. Twigg, and A. S. Edelstein, *J. Vac. Sci. Technol.*, **A8** (1990) 3466.

61. K. Inomata, Y Saito, and S. Hashimoto, *J. of Magnetism and Magnetic Mat.*, **121** (1993) 350.
62. R. Nakatani, T. Dei, T. Kobayashi, and Y. Sugita, *IEEE Trans. on Mag.*, **28** (1992) 2668.
63. Y. Wang, F. Z. Cui, W. Z. Li, and Y. D. Fan, *J. of Magnetism and Magnetic Mat.*, **102** (1991) 121.
64. S. B. Qadri, C. Kim, M. Twigg, P. Lubitz, and M. Rubinstein, *J. Vac. Sci. Technol.*, **A9** (1991) 512.
65. T. Hamaguchi, H. Aida, S. Nakaqawa, and M. Naoe, *J. Appl. Phys.*, **73** (1993) 6444.
66. F. Z. Cui, Y. Wang, H. Cui, W. Z. Li, and Y. D. Fan, *J. of Phys. D: App. Phys.*, **27** (1994) 2246.
67. E. Kay and S. M. Rossnagel, in Reference 23, pp.170–218.
68. J. E. Greene, *Solid State Technology* (April, 1987), p. 115.
69. D. V. Vechten, G. K. Hubler, E. P. Donovan, and F. D. Correll, *J. Vac. Sci. Technol.*, **A8** (1990) 821.
70. G. K. Hubler, D. V. Vechten, E. P. Donovan, and C. A. Carosella, *J. Vac. Sci. Technol.*, **A8** (1990) 831.
71. J. S. Colligon, *J. Vac. Sci. Technol.*, **A13** (1995) 1649.
72. B. K. Kellerman, E. Chason, J. A. Floro, and S. T. Picraux, *Appl. Phys. Lett.*, **67** (1995) 1703.
73. X. D. Wu, S. R. Foltyn, P. Arendt, J. Townsend, C. Adams, I. H. Campbell, P. Tiwari, Y. Coulter, and D. E. Peterson, *Appl. Phys. Lett.*, **65** (1994) 1961.
74. Paul C. McIntyre, Kevin G. Ressler, Neville Sonnenberg, and Michael J. Cima, *J. Vac. Sci. Technol.*, **A14** (1996) 210.
75. Satis Vacuum Corp., 10510 Olympic Dr., Dallas, TX 75220.
76. D. J. Kester, K. S. Ailey, D. J. Lichtenwalner, and R. F. Davis, *J. Vac. Sci. Technol.*, **A12** (1994) 3074.
77. K. J. Boyd, D. Marton, S. S. Todorov, A. H. Al-Bayati, J. Kulik, R. A. Zuhr, and J. W. Rabalais, *J. Vac. Sci. Technol.*, **A13** (1995) 2110.
78. J. Robertson, *Surface and Coatings Tech.*, **50** (1992) 185.
79. M. J. Mirtich, D. M. Swec, and J. C. Angus, *Thin Solid Films*, **131** (1985) 245.
80. Dan Nir and Michael Mirtich, *J. Vac. Sci. Technol.*, **A4** (1986) 560.
81. S. Aisenberg and R. Chabot, *J. Appl. Phys.*, **42** (1971) 2953.
82. M. J. Mirtich, D. M. Swec, and J. C. Angus, *NASA Technical Memo*, TM-83743, (NASA, Washington, D.C., 1984).

CHAPTER 5.4

Pulsed Laser Deposition

Jeffrey T. Cheung
Rockwell Science Center

5.4.1 INTRODUCTION

A physical vapor deposition process must have three basic components: a source material, a substrate, and an energy supply to transport material from the source to the substrate during film deposition. Of these three components, the form of energy and its associated mass transfer mechanism are what set one technique apart from another. In fact, thin-film deposition techniques are often named after these mechanisms, such as sputtering deposition and E-beam evaporation. Pulsed Laser Deposition (PLD), as the name suggests, uses a high-energy pulsed laser as an external power source to ablate the source or target material [1,2]. Interaction is short but intense, and induces ablation via a cascade of complex events. Theoretical and experimental studies of this phenomenon have always been a subject of intense interest and have spawned new ideas in laser applications. One idea, Pulsed Laser Deposition, was first demonstrated by Turner and Smith [3] nearly three decades ago and has become a widely accepted technology of today.

Despite its early inception and demonstration, PLD development was slow and sometimes stagnant until the 1980s. The reason is twofold: the lack of a niche for this technique to compete against the more mature thin-film growth technologies, and the lack of reliable and affordable high-energy lasers. It has remained an obscure research tool for nearly two decades. Its first explosive growth was triggered by the successful preparation of high-quality superconducting $YBa_2Cu_3O_{7-x}$ films by PLD [4]. At the same time, breakthroughs were made in laser technology and

ISBN 0-12-325065-7
$25.00

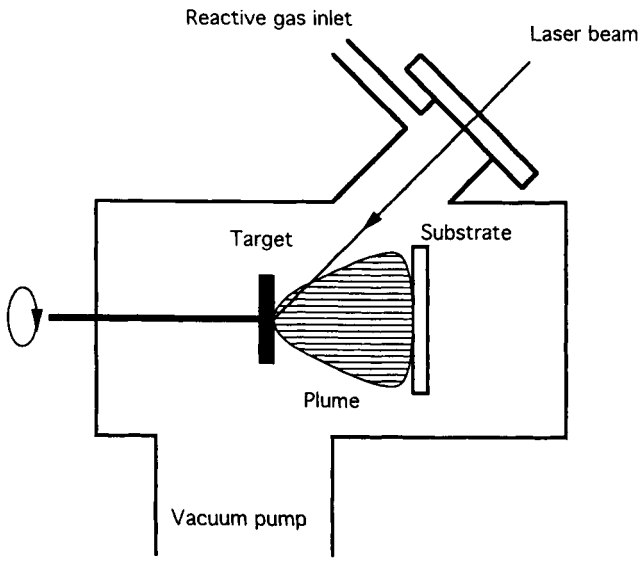
operational costs have declined steadily, especially for the UV excimer laser (which is considered the system of choice for PLD). This technological advancement played a major role in transforming PLD from a laboratory curiosity to a mainstream material growth technique.

5.4.2 PULSED LASER DEPOSITION SYSTEM

Among all physical vapor deposition processes, the PLD hardware is the simplest and most flexible. Figure 1 shows a schematic representation. The center piece is a high-energy pulsed laser placed outside the vacuum chamber.

The laser beam is focused onto the target surface. The type of laser used for PLD has been evolving with time. Ruby [3,5] and Nd:Glass lasers [6] were used in early studies. They were replaced by the more reliable Nd:YAG [7] and TEA-CO_2 lasers [8] in the 1970s. In the 1980s, the trend moved toward using short-wavelength lasers to produce ablation plumes with higher plasma temperature and thin films with better crystallinity. Consequently, UV excimer lasers have become the laser of choice. To date, the majority of PLD systems use either the 248 nm KrF excimer laser or the 193 nm ArF excimer laser. Over the last 10 years, ex-

Fig. 1.



Schematic representation of a PLD system.

cimer laser performance has improved drastically. This engineering triumph has made PLD competitive against other thin-film deposition techniques.

To avoid deep surface pitting caused by repetitive ablation, the focused beam changes location constantly, by target rotation, by beam rastering, or by a combination of both. The most popular configuration uses rotating target disks mounted on a carousel multiple holder for multiple-layer growth. Commercial target holder of this design are available from a number of vendors.

Special substrate holders are used to address the special environmental needs for growing oxide films. These holders must survive the harsh condition of high operating temperature (up to 850°C) in an oxidizing environment. The most successful designs use high-power halogen quartz lamps or nichrome wire heaters embedded inside an oxidation-resistant Inconel sheath. The latter has proven to be particularly compact and durable.

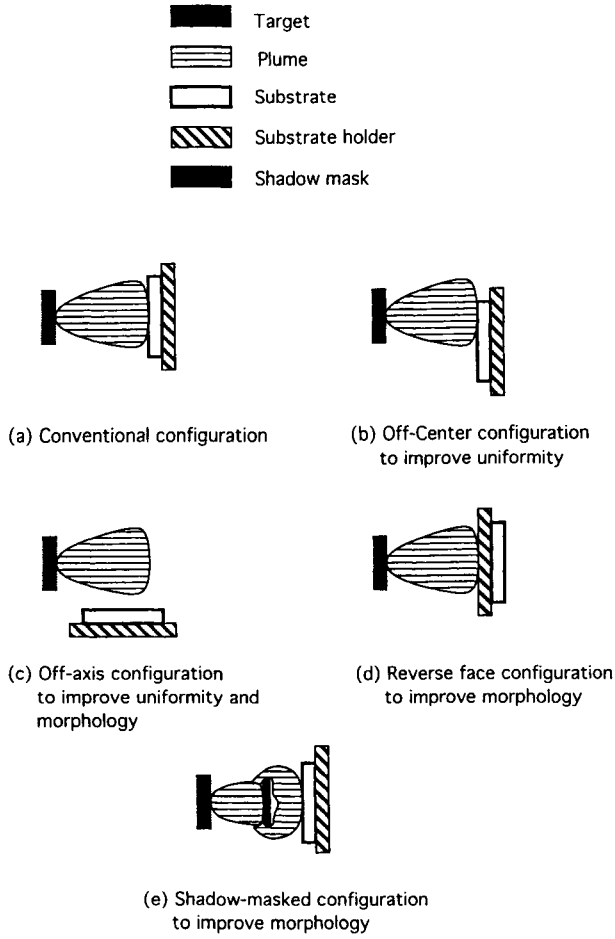
Target/substrate geometry is a key parameter and strongly affects the film quality. Normally the substrate and target are placed face to face a few centimeters apart, such that the substrate just touches the tip of the ablation plume, as shown in Figure 2(a). Other configurations are also used for specific purposes. Figure 2(b) shows an "off-center" configuration to improve thickness uniformity. Figure 2(c) shows an "off-axis" geometry where a rotating substrate is placed along the direction of the ablation plume [9]. Film deposition takes place when radially diffused ablation species land on the substrate, whereas the heavy particles continue their trajectories. Figure 2(d) represents an extreme case where the substrate is facing away from the plume to reduce splashing [10]. Finally, Figure 2(e) shows the use of a shadow mask placed between the substrate and the target to divert plume flow and prevent splashed particles from reaching the substrate [11].

An important capability of PLD is the deposition of films in a reactive environment. Gas inlets are usually located just behind the vacuum window through which the laser beam transmits. This serves the dual purpose of introducing reactive gas and preventing condensation of evaporants on the window, which can reduce its optical transmission. Unlike any other physical vapor deposition techniques, PLD can be carried out in any background condition from high vacuum to atmospheric pressure [12].

The flexibility of PLD has brought out many system modifications. Most of them are designed to activate reactive gas in order to achieve better stoichiometry and lower growth temperature. Examples include the use of a discharge ring in the ablation plume to lower the growth temperature for $YBa_2Cu_3O_{7-x}$ (YBCO) superconducting films [13], a microwave discharge hydrogen source to grow diamond film from a graphite target [14], and a microwave discharge seeded nitrogen nozzle beam to synthesize β-carbon nitride (CN_x) films [15].

Another notable example is the combined use of a Pulsed Laser Evaporation source with Molecular Beam Epitaxy to conduct "Bandgap Engineering" [16].

Fig. 2.



Five different PLD target/substrate configurations.

Composition depth resolution of two monolayers can be achieved by modulating the deposition flux density with laser repetition rate.

Experiments using two synchronized lasers [17] were also reported . In these experiments, one laser is used for target ablation and the other laser with lower energy density is used to illuminate the as-deposited film to promote nucleation and enhance crystallinity.

This list of variations of PLD is by no means complete. PLD's flexibility is still being exploited to further improve its capability.

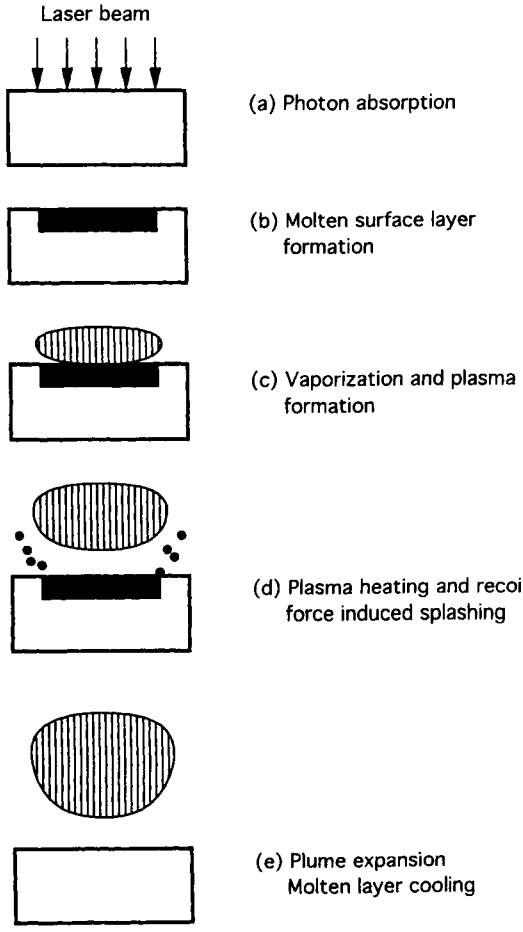
5.4.3

THE ABLATION MECHANISM

Material ablation induced by high-power laser radiation is a very complex process and can be described as a series of heterogeneous events. Sequential steps are shown in Figure 3. A full cycle takes approximately one microsecond to complete.

At first, the target material absorbs the leading edge of the laser pulse. If the energy density is above the ablation threshold, sufficient photons will be absorbed

Fig. 3.



Schematic representation of various stages of laser–solid interaction leading to ablation plume formation.

to melt and vaporize surface material. Under typical PLD experimental conditions, the energy fluence is a few J/cm^2 for 10^{-8}-second pulses. It is almost an order of magnitude higher than the ablation thresholds of most materials. The peak power density under this condition exceeds 10^8 W/cm^2. The electric field inside an absorbing medium under this condition can be estimated from the following relation:

$$E = (2\Phi/cE_0n)^{1/2} \tag{1}$$

where E electric field
Φ power density
E_0 dielectric constant in vacuum
n refractive index of the medium
c speed of light

For a material with $n = 1.5$ under a radiation power density of 5×10^8 W/cm^2, the corresponding electric field inside the bulk reaches 5×10^5 V/cm^2, high enough to cause local dielectric breakdown. Therefore, the molten surface layer vaporizes to form a dense plasma. The early stage of plasma flow is hydrodynamic and expands isoentropically into the vacuum like a supersonic nozzle beam. It absorbs the remaining portion of the laser pulse to reach a higher temperature. Plasma temperature has been determined to be in the range of 10,000K by optical emission spectroscopy.

During the rapid supersonic expansion, a shock wave is generated. Its recoil pressure exerts a force strong enough to ripple the molten surface layer. The oscillatory motion of the ripples propagates along the surface and is interrupted by the liquid–solid interface at the edge of the molten pool. The continuous propagation crashes against the solid interface and expels liquid droplets into vacuum. It is the major cause of "splashing."

As the reactive gas pressure increases, so does the frequency of collisions among various species. Multiple collisions take place and promote condensation among molecular and atomic species to form clusters ranging from dimers to submicron clusters. Their density and distribution are sensitive to background gas pressure and the distance to target surface. In general, the cluster concentration in the plume increases with background gas pressure and the distance from the target.

This description is an oversimplified version of a complex process. There have been many experiments, using an array of *in situ* diagnostic tools such as optical emission spectroscopy, laser-induced fluorescence, CCD camera, ion probes, and time-of-flight mass spectroscopy (TOFMS) to monitor and understand this process [18]. However, even with such powerful diagnostic instrumentation, the physics of laser ablation is still far from being completely understood. Despite the simple experimental setup of PLD, its ablation mechanism is no doubt the most sophisticated among all physical vapor deposition techniques.

5.4.4 ADVANTAGES AND LIMITATIONS

The unique approach of PLD has its advantages and limitations. The advantages have stimulated the rapid growth, and the disadvantages have kept it from becoming a large-scale production tool in the technological mainstream.

5.4.4.1 Advantages

When a multiple-component target is heated on its way to vaporization, individual constituent elements segregate in the solid (or melt) because of their different volatility. Elements with the highest vapor pressure will be vaporized, first followed by the less volatile ones. In other words, the source material does not melt and vaporize "congruently," and a thin film deposited under this condition does not maintain the target stoichiometry. Unfortunately, many materials behave this way, ranging from the semiconductor GaAs to the superconductor $YBa_2Cu_3O_{7-x}$. To grow thin films of these materials by simple thermal evaporation, individual sources must be used to compensate for different individual vapor pressures. In PLD, the short thermal cycle induced by laser pulse provides a simple solution to this problem. The volume heated by each pulse is characterized by the laser spot size multiplied by the penetration depth, $L = 2(D \cdot \tau)^{1/2}$, where D is the thermal diffusivity and τ is the laser pulse duration. If L is less than the thickness of material ablated per pulse, then the stoichiometry of the plume will be the same as the target, enabling the use of a single source. This criterion is satisfied by many materials.

CRYSTALLINITY ENHANCEMENT

Ablation species are energetic particles with kinetic energy up to a few keV's and potential energy in the form of high plasma temperature (i.e., electronic excitation). In the presence of background gas for reactive deposition, kinetic energy is retarded by collisions to just a few electron volts, which is ideal for enhancing film crystallinity. A comparison of the crystallinity of ZnO thin films deposited under two different target–substrate configurations gives convincing evidence of this effect. If the substrate faces the target, as in Figure 2(a), where it receives the full benefit of impingement of the energetic particles in the ablation plume, the (0002) X-ray diffraction peak of the thin film has a FWHM width of 0.21°. If the substrate is placed parallel to the plume, as in Figure 2(c), where it completely avoids direct exposure to the plume, the (0002) diffraction peak FWHM is 0.43°.

On the other hand, the role of potential energy in the form of electronic excita-

tion is not as clear. Its significance lies in activating the reactive background gas via collision-induced energy transfer.

REACTIVE DEPOSITION

The high plasma temperature of the ablation plume acts as an energy reservoir for activating the reactive background gas. Therefore, PLD bears some resemblance to magnetron sputtering. However, in magnetron sputtering deposition of oxides, film quality is very sensitive to oxygen partial pressure in the Ar/O_2 sputtering gas mixture. Excessive oxygen causes resputtering, whereas insufficient oxygen results in films with high oxygen deficiency. The optimal window of oxygen partial pressure is narrow. PLD does not have this shortcoming. In addition to oxides, PLD has also been effective in growing nitride films in a nitrogen ambient [19].

MULTILAYER GROWTH

Because of stoichiometry preservation by PLD, only a single source is needed to grow films with chemical complexity. Therefore, multiple-layered structures, such as superconductor [20] and ferroelectric [21] memory devices can be grown easily by ablating multiple sources.

FAST TURNAROUND TIME

PLD only needs a very small amount of target material as long as the target is larger than the laser spot. Therefore, it is an ideal way to grow thin films of exotic materials when only a small quantity is available. Because of the experimental simplicity, PLD is also known for its fast turnaround time when growing a thin film of a new material starting from its powder form. This technique gives material scientists a quick peek at new material properties with minimal effort and time. In fact, PLD is sometimes perceived as the material scientist's equivalent of what "rapid prototyping" has been to mechanical engineers.

5.4.4.2 Limitations and Solutions

PLD has some limitations. Following is a list of the most common ones and their solutions:

LARGE-AREA UNIFORMITY

The difficulty of achieving large-area uniformity has always been PLD's main drawback. There are several reasons for this. First, the laser ablation plume is

non-Lambertian. Its spatial profile follows a $(\cos\theta)^n$ distribution where the exponent n varies from $n = 3$ to $n = 11$ and higher. In addition to the narrow angular profile of the plume, separation between the target and substrate is only a few centimeters in order to take full advantage of the laser-induced plasma. Both limitations pose severe problems in scaling up to large area.

There have been many attempts to meet this challenge. Solutions include an off-center substrate mount, an off-axis substrate mount, and a rastering laser spot over a large target [22]. Verdicts are still out to decide as which approach is the most practical and appealling.

ROUGH SURFACE MORPHOLOGY DUE TO "SPLASHING"

A major concern of PLD is the presence of micron-size particles on deposited films due to "splashing," which is the consequence of intense laser–target interaction.

There are three major causes of splashing.

1. *Exfoliation.* During laser ablation, materials are removed in a preferential orientation along the direction of the incident laser beam. The departed material leaves deep channels and thin columns on the surface. Such microstructures are mechanically weak and are easily broken by thermal shocks [23]. Debris lands on the substrate surface, resulting in rough morphology. The signature of this type of splashing is the appearance of irregular-shaped micron-sized particles.
2. *Subsurface boiling.* Subsurface boiling was originally referred to as "true splashing" by Ready [24]. It occurs if the time required to transfer laser energy into heat is shorter than the time to evaporate a surface layer with a thickness on the order of absorption depth. Under this condition, the subsurface layer is superheated before the surface material is completely vaporized. Boiling subsurface material explodes from the surface and ejects micron-size molten globules. Tourtellote [6] estimated a threshold power density for subsurface boiling of a metallic target to be 5×10^8 W/cm^2 and higher for dielectric materials. This power density range is higher than typically found in PLD. Therefore, subsurface boiling has a negligible significance except for a few materials containing very volatile constituents such as $Hg_{1-x}Cd_xTe$ (high Hg pressure) [25].
3. *Shock-Wave-Induced Droplet Expulsion.* Unlike subsurface boiling, in which the force that expels liquid droplets comes from within the bulk, shock-wave-induced droplet expulsion is caused by the recoil pressure exerted by the shock wave of the departing plume. The signature of this type of splashing is also condensed molten globules, similar to subsurface boiling, but much smaller, typically in the submicron range. In contrast to subsurface boiling, the threshold power density for this mechanism to occur is low, nearly as low as the ablation threshold. Therefore, there is little chance of eliminating it completely. For some compounds, such as CdTe, ZnTe, and ZnSe, PLD-grown films are smooth without any

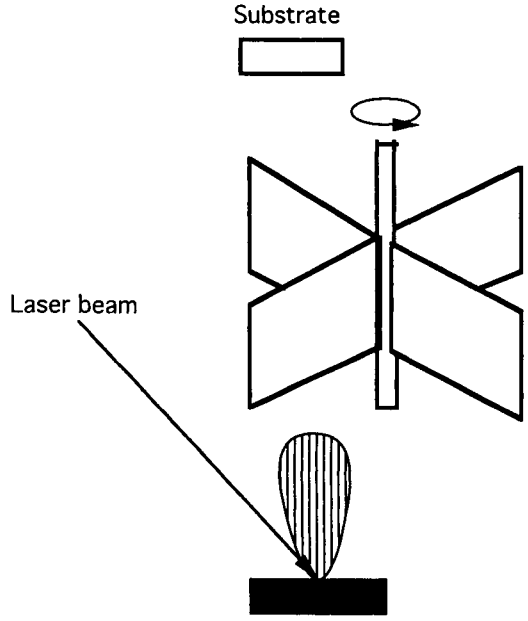
trace of splashing. This is consistent with the recoil pressure/surface melt interaction model, because these materials do not have a melt phase. The vaporization of these materials is dominated by sublimation.

SOLUTIONS TO AVOID SPLASHING

Since the inception of PLD, there has been a continuous struggle to eliminate splashing. Many experiments have been attempted. Some of them can be considered as the brute force approach, while the others rely on more subtle principles. Here, we only list the most common remedies to this problem.

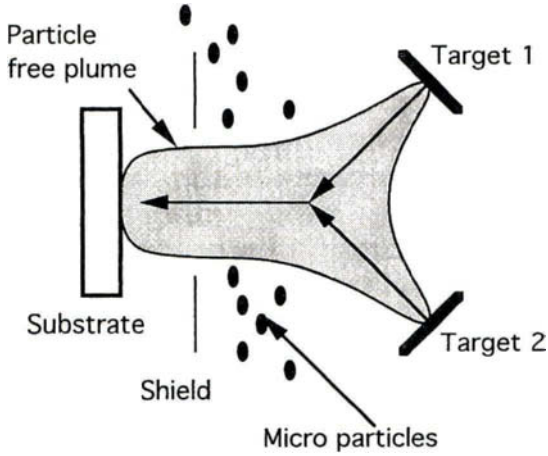
1. *Target Surface Conditioning.* The use of dense targets, mechanically polished surface, or ablation with low-fluence laser pulses prior to film growth can reduce exfoliation.
2. *Particle Filter.* The particle filter approach relies on the fact that particles in the ablation plume travel much more slowly (about 10^3 cm/sec) than atomic and molecular species (10^5 cm/sec and higher). Particles can therefore be removed by a synchronous mechanical shutter [26], rotary vane filter [27] (shown in Figure 4), or multiple-disk velocity filter [28]. These methods are effective in removing par-

Fig. 4.



A rotary vane particle filter [27].

Fig. 5.



Cross plume configuration [31].

ticles from the plume, but their practical value is limited because of their bulkiness and low transmission.

Another type of particle filter uses a second laser or a synchronous pulsed gas jet perpendicular to the plume to "blow away" the slower particles [29]. However, the efficiency of this approach is low.

3. *Special Target–Substrate Geometry.* Two special configurations have been reported to reduce particle formation: a cross plume configuration and an off-axis configuration.

a. *Cross Plume Configuration*

The cross plume configuration was first reported in 1975 [30] and later modified [31] by Gapanov. In this approach, shown in Figure 5, two plumes arrive at the same time and intersect in a region not far from the targets. Because of the high number density in the plumes, the collisional frequency inside the intersecting zone is high. Kinematic restriction directs scattered atomic and molecular species to move forward, while micron-size particles from splashing continue their trajectories because of their heavy mass and are therefore removed. This approach is applicable to all materials, but the cross section of scattered plume is too small to be practical.

b. *Off-Axis Substrate Configuration*

The flow of ablation plume is hydrodynamic and has a radial component due to the large number of collisions; however, the trajectories of heavy particles do not change their courses. Therefore, in an off-axis configuration, the radially outdiffused species can deposit onto the substrate [9]. Films free of particles can be grown this way. This technique suffers from having a low depo-

sition rate, and, most importantly, it negates some advantages offered by PLD. In ZnO deposition using the off-axis configuration, particulate-free film with a thickness uniformity of better than 5% over a 3-inch diameter area can be grown [32]; however, its crystallinity is inferior to that of films grown in the traditional face-to-face configuration. If a lattice-matching substrate is used, then the dominant driving force for nucleation and coalescence is the lattice energy between the film and the matching substrate rather than the plume impingement energy. Under such circumstances, high-quality epitaxial films can be grown using this configuration.

4. *Liquid Target.* The major source of splashing is the expulsion of molten droplets from the surface melt by the recoil pressure of the departing plume. Expulsion takes place at the liquid–solid interface near the edge of the laser spot. Therefore, one way to eliminate this effect is to make the area of the melt much greater than the laser beam spot size. By the time the oscillatory motion of the melt surface reaches the liquid–solid interface, its amplitude has been sufficiently damped to prevent any splashing. The principle was first verified by Sankur [33] in a brilliant attempt to grow Ge films by PLD using a molten Ge source. Particulate-free films could be grown even under extremely high-energy fluence. Recently Xiao [34] reported the growth of diamond films by PLD, using diffusion pump oil as a liquid target source. The approach is unique and successful. Although the use of a liquid target is scientifically sound and technically simple, it suffers from the major drawback of material availability. As a matter of fact, very few materials satisfy the criterion of having low vapor pressure at the melting temperature.

This obstacle was recently overcome by Witanachchi [35] with a dual-laser approach in which a TEA-CO_2 laser was used in conjunction with a KrF laser. The pulses are synchronized such that the target surface is melted by the CO_2 laser just prior to the arrival of a KrF laser pulse for ablation. The CO_2 laser is defocused, covering an area much larger the focused KrF laser spot. This approach is extremely promising. It can be applied to all materials of interest and has an added advantage in achieving higher plume plasma temperature and better uniformity over a large area.

5.4.5 MATERIALS SURVEY

PLD demonstrates its versatility in a wide material selection. The first list of PLD-grown materials was compiled in 1988 in a review article by Cheung and Sankur [1]. Since then, owing to the growing popularity, the list has grown considerably [2]. Here, we will only highlight some key examples representing different families of materials and point out some trends.

5.4.5.1 Metals

Most work on metal films grown by PLD was reported in the 1960s and 1970s [36]. Since PLD does not present any obvious advantage in this area over competitive techniques, studies on metal film growth by PLD have slowed. Only in areas such as thin-film coating of magnetic metal alloys [37] and other non-equilibrium alloys [38] where stoichiometry plays an important role, does PLD have an advantage over the conventional techniques.

5.4.5.2 Semiconductors

There are some advantages for using PLD to grow semiconductor films, but there is also one major disadvantage: the presence of micron particles from splashing. The only semiconductors that do not splash are the II-VI compounds, because they lack a liquid state. By no coincidence, PLD growth of II-VI semiconductors has been very successful. This work can be exemplified by the series of studies on $Hg_{1-x}C_xdTe$, CdTe and their superlattices [39] and ZnSe and ZnSe films for solar cell application [40]. More recently, Lowndes has demonstrated the growth of p-type ZnSe film by *in situ* nitrogen doping [41], opening up novel applications. Overall, the current trend suggests that PLD-grown semiconductor layers will remain in these niches and must face stiff competition from well-developed technologies such as MBE and MOCVD.

5.4.5.3 Diamondlike Carbon

Diamondlike carbon film is a composite network of an SP^2 (graphitic) and SP^3 (diamond) carbon. Although other techniques such as CVD and ion beam deposition have been used to grow this material, the qualities of thermal conductivity, microhardness and field emission efficiency are still inferior to PLD DLC films [42]. In addition, PLD has been used to grow CN_x films [15] exhibiting the highest nitrogen content.

5.4.5.4 Perovskites

Perovskite is a family of oxides with a general formula of ABO_3 where A and B represent cations. Deposition of perovskites has been the traditional stronghold of PLD and will continue to be so. PLD owes its rapid growth in the late 1980s to the successful demonstration of YBCO films. In the early 1990s, the YBCO work opened the door for ferroelectric films studies [21]. Recent activity in co-

lossal magneto resistance (CMR) materials is a further extension of this research area [43].

The structural properties of perovskite materials and the advantages of PLD match up extremely well. In fact, PLD has served as a trail blazer in the new material frontier involving perovskite thin films. However, we should re-emphasize that PLD growth of perovskite thin film is not a large-area process. Scaleup to larger substrates is underway in a number of laboratories.

5.4.5.5 Simple Oxides

Oxide films are the most logical material of choice for PLD, especially "simple oxides" referring to those containing only one type of cation. In this area, to be successful PLD must compete against E-beam deposition and sputtering, in particular RF magnetron sputtering. PLD has met such challenges and demonstrated a definitive advantage. Simple oxide film growth by PLD has flourished and is firmly established in this niche. The highest-impact work in this area is the series on heteroepitaxy of oxide/semiconductor structures by Fork [44]. In particular, the epitaxy of yttria-stablized zirconia (YSZ) on Si [45] and the epitaxy of MgO on GaAs [46], two materials of technological relevance, are important examples.

5.4.5.6 Nitrides

The main challenge in growing nitride films is to overcome the barrier of producing high concentrations of active nitrogen. PLD has proven successful in doing so. The first successful demonstration in this area is the growth of TiN [47], followed by polycrystalline AlN [19], BN [48], and the growth of GaN in high vacuum [49]. The main rival in this area is MOCVD, which has been very actively used for nitride growth in the recent year.

5.4.5.7 Organic Materials

Clearly, the growth of organic materials must involve a nonthermal process to avoid dissociation. For this reason, UV PLD is a logical candidate. The ablation mechanism of organic polymers is quite different from that of other materials. The mechanism is mainly photolytic, involving "zipping" polymer chains into small units and repolymerizing them on the substrate [50]. Examples of materials deposited include polyethylene, polycarbonate, polyimide, polymethylmethacrylate (PMMA) [51], polyphenylene sulfide [52], and even Cu-phthalocyanin films [53].

5.4.6 FUTURE OUTLOOK

PLD has made giant strides in the last 10 years and has undergone a metamorphosis from an obscure and curious laboratory innovation to a powerful thin-film deposition technique that is widely accepted by the technical community. There have been numerous articles, books, workshops, and symposia on this subject. In terms of material selection, it has stepped out from the early success with perovskites and broadened the horizon to other systems. The availability of commercial PLD systems is another sign of maturity. Along a parallel path, *in situ* diagnosis of the laser ablation process has evolved rapidly to provide insightful information relevant to process control. While UV excimer lasers used to be the weak link of this process, because they lack long-term stability, they have now been improved to a point comparable to other thin-film deposition hardware. However, despite all this progress, PLD still remains a research tool. To join the mainstream of technology, PLD must address and solve the issues concerning large-area scaleup and morphology improvement. It must also identify niches where the performance of this technique is unique and superior to other techniques.

REFERENCES

1. Jeffrey T. Cheung and Haluk Sankur, *CRC Critical Review on Solid State Material Science*, **15** (1988) 63–109.
2. *Pulsed Laser Deposition of Thin Films*, edited by Douglas B. Chrisey and Graham K. Hubler (Wiley, New York, 1994).
3. A. F. Turner and H. M. Smith, *Appl. Optics*, **4** (1965) 147–148.
4. D. Dijkamp, T. Venkatesan, X. D. Wu, S. A. Shaheen, J. Jisrawi, Y. H. Min-Lee, W. L. McLean, and M. Croft, *Appl. Phys. Lett.*, **51** (1987) 619–621.
5. V. S. Ban and D. A. Kramer, *J. Material Science*, **5** (1970) 978.
6. H. Schwartz and H. A. Tourtellotte, *J. Vac. Sci. Technol.*, **6** (1969) 3763.
7. P. Meir-Komor, *Nucl. Instrum. Methods*, **167** (1979) 73.
8. Yu-A Bykovskii, V. M. Boyakov, V. T. Galochkin, A. S. Mochanov, I. N. Nokolaev, and A. N. Orevskii, *Sov. Phys. Tech. Phys.*, **23** (1978) 578.
9. B. Holzapfel, B. Roas, L Schulz, P. Bauer, and G. Saemann-Ischenko, *Appl. Phys. Lett.*, **61** (1992) 3178.
10. R. J. Kennedy, *Thin Solid Films*, **214** (1992) 223.
11. Mamoru Iwabuchi, Kazuya Kinoshita, Hiroshige Ishibashi, and Takeshi Kobayashi, *Jpn. J. Appl. Phys.*, **33** (1994) L610.
12. K. Maruyama, S. Endo, G. Sasaki, K. Kamata, J. Nishino, and K. Kuchitsu, *J. Mat. Sci. Lett.*, **11** (1992) 1588.
13. H. S. Kwok et al., *Appl. Phys. Lett.*, **59** (1991) 3643.
14. Seth Jashree, R. Padiyath, T. H. Rasmussen, and S. V. Babu, *Appl. Phys. Lett.*, **63** (1993) 473.

15. C. Niu, Y. Z. Lu, and C. M. Lieber, *Science*, **261** (1993) 334.
16. J. T. Cheung and J. Madden, *J. Vac. Sci. Technol.*, **B5** (1987) 705.
17. T. Kawaii, *Digest, IEEE/LEOS 1996 Summer Topical Meeting, Advanced Applications of Lasers in Materials Processing*, (IEEE, Keystone, CO, 1996), pp. 3–4.
18. See David B. Geohegan, Diagnostics and characteristics of pulsed laser deposition laser plasma, in Reference 2, chap. 5.
19. G. M. Norton, P. G. Kotula, and C. B. Carter, *Appl. Phys. Lett.*, **70** (1991) 2871.
20. X. X. Xi, C. Doughty, A. Walkenhorst, et al., *Phys. Rev. Lett.*, **68** (1992) 1240.
21. R. Ramesh et al., *Appl. Phys. Lett.*, **57** (1990) 1505.
22. J. A. Greer, *J. Vac. Sci. Technol.*, **A10** (1992) 1821.
23. Roger Kelly and J. Rotheberg, *Nucl. Inst. and Methods*, **B7/8** (1985) 755.
24. J. F. Ready, *Appl. Phys. Lett.*, **3** (1963) 11.
25. J. T. Cheung, *Appl. Phys. Lett.*, **43** (1983) 255.
26. D. Lubben, S. A. Barnett, K. Suzuki, S. Gobatikin, and J. E. Greene, *J. Vac. Sci. Technol.*, **B3** (1985) 968.
27. A. B. Beavitt, *Thin Solid Films*, **1** (1967) 69.
28. H. Dupendant, *Appl. Surf. Sci.*, **43** (1989) 369.
29. Kouichi Murakami, *Laser Ablation of Electronic Materials—Basic Mechanisms and Applications*, edited by E. Fogarassy and S. Lazare (Elsevier Science, 1992) pp. 125–140.
30. S. V. Gaponov, A. A. Gudkov, and A. A. Fraerman, *Sov. Phys. Tech. Phys.*, **27** (1982) 1130.
31. M. D. Strikovsky, E. B. Klyuenkov, S. V. Gapanov, J. Schubert, and C. A. Copetti, *Appl. Phys. Lett.*, **63** (1993) 1146.
32. J. T. Cheung (unpublished results).
33. H. Sankur, W. J. Gunning, J. DeNatale, and J. Flintoff, *Appl. Phys. Lett.*, **65** (1989) 2475.
34. R. F. Xiao, *Appl. Phys. Lett.*, **69** (1995) 1022.
35. Sarath Witanachchi and Pritish Mukherhee, in *Advanced Laser Processing of Materials—Fundamentals and Applications*, MRS Symposium Proceeding Vol. 397, edited by R. Singh, D. P. Norton, L. D. Laude, J. Narajan, and J. T. Cheung (Material Research Society, Pittsburgh, 1996).
36. V. M. Boyakov, V. M. Epikhin, B. A. Kalin, M. K. Makhatov, I. N. Nikolaev, and G. N. Shiskin, *Sov. J. Quantum Electron.*, **8** (1978) 902.
37. K. Hayashi, M. Hayakawa, W. Ishikawa, Y. Ochiai, H. Matsuda, Y. Iwasake, and K. Aso, *J. Appl. Phys.*, **61** (1987) 3514.
38. F. R. DeBoer, R. Boom, W. C. M. Mattens, A. R. Miedema, and A. D. Niessen, *Cohesion in Metals* (North-Holland, Amsterdam, 1988).
39. J. T. Cheung, G. Niizawa, J. Moyle, N. P. Ong, B. Paine, and T. Vreeland, Jr., *J. Vac. Sci. Technol.*, **A4** (1986) 2086.
40. A. Aydinli, A. Compaan, G. Contreras-Puente, and A. Mason, *Solid State Comm.*, **80** (1992a) 465.
41. C. M. Rouleau, D. H. Lowndes, J. W. Macamy, J. D. Budai, D. B. Poker, D. B. Geohagen, A. A. Puretzky, and Shen Zhu, *Appl. Phys. Lett.*, **67** (1995) 2545.
42. For a review, see A. Rengan and J. Narajan, *Laser Ablation of Electronic Materials—Basic Mechanisms and Applications*, edited by E. Fogarassy and S. Lazare (Elsevier Science, Amsterdam 1992), pp. 363–376.
43. J. H. Hao, X. T. Zeng, and H. K. Wong, *J. Appl. Phys.*, **79** (1996) 1810.
44. David Fork, *Epitaxial oxides on semiconductors*, in Reference 2, chap. 6.
45. D. K. Fork, D. Fenner, G. A. N. Connell, J. M. Phillips, and T. H. Geballe, *Appl. Phys. Lett.*, **57** (1992) 1137.
46. W. Presseit et al., *Appl. Phys. Lett.*, **46** (1992) 1499.
47. J. Narajan, *Appl. Phys. Lett.*, **61** (1992) 1290.
48. G. L. Doll, J. A. Sell, C. A. Taylor II, and R. Clarke, *Phys. Rev.*, **B43** (1991) 6816.

49. R. F. Xiao, H. B. Liao, N. Cue, X. W. Sun, and H. S. Kwok, *J. Appl. Phys.*, **80** (1996) 4226L.
50. R. Srinivisan, B. Braren, and K. G. Casey, *J. Appl. Phys.*, **68** (1990) 1842.
51. S. G. Hansen and T. E. Robitaille, *J. Appl. Phys.*, **84** (1988b) 2122 and *Appl. Phys. Lett.*, **52** (1988) 81.
52. S. Kale, M. Swaminathan, and S. B. Ogale, *Thin Solid Films*, **206** (1991) 161.
53. F. Kannari, N. Matsumoto, E. Ina, and T. Fujii, *Digest IEEE/LEOS 1996 Summer Topical Meetings (Advanced Applications of Lasers in Material Processing)* (IEEE, Keystone, CO, 1996), p. 63.

CHAPTER 5.5

Plasma-Enhanced Chemical Vapor Deposition

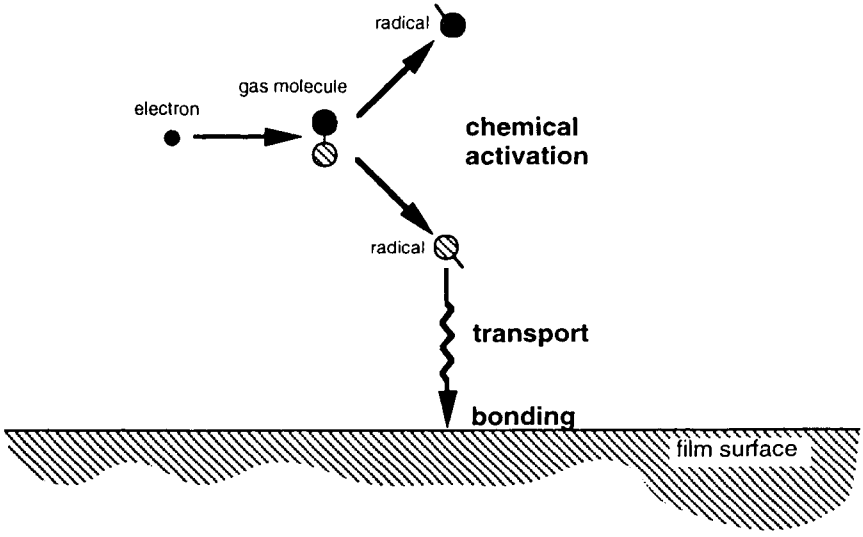
Frank Jansen
BOC Coating Technology

5.5.1 INTRODUCTION

Some thin-film deposition processes use molecular gases and vapors, rather than solid evaporants or sputter targets, as their source materials. These processes are generally referred to as "chemical vapor" deposition processes, implying that the starting gases are "chemically activated" before deposition. In plasma-enhanced chemical vapor deposition (PECVD), chemical activation is achieved by supplying electrical power to a gas at reduced pressure, typically between 10 mtorr and 10 torr. At these pressures, the application of a sufficiently high voltage creates a visible glow, called glow discharge plasma. The plasma consists of about equal concentrations of ions and electrons. The visible glow is caused by charge recombination processes and the relaxation of electronically excited atoms and molecules. The electrical power is coupled into the gas through the mediation of plasma electrons. The energetic electrons in the plasma ionize the gas, if only to a minor extent—about one part per million. A much larger fraction of the gas, about 1 percent, is chemically activated by the electrons. The increased chemical activity of the gas results primarily from dissociation of the molecules into smaller species, called radicals [1,2]. Radicals are chemically unsaturated and therefore

ISBN 0-12-325065-7
$25.00

Fig. 1.



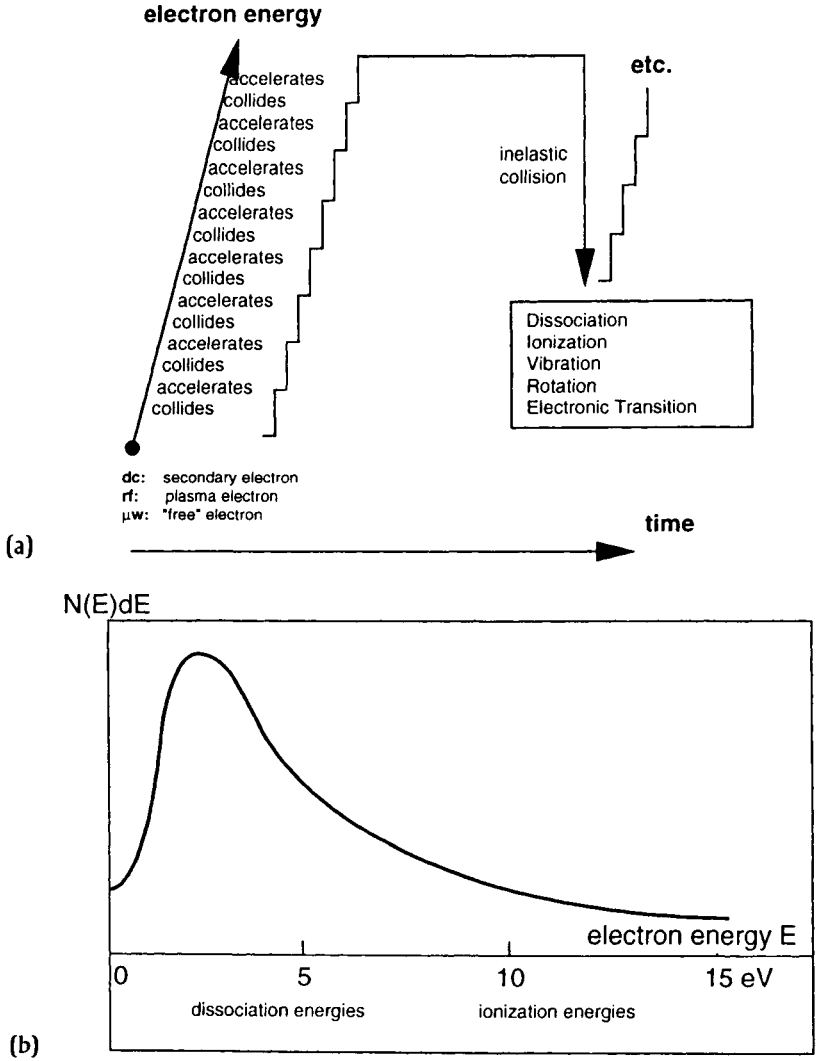
The basic elements of the PECVD process: (a) chemical activation by electron impact, e.g., molecular dissociation; (b) transport of the activated species to the film surface, e.g., by diffusion; and (c) bonding to the film surface.

capable of chemical reactions at high rates; they are the species that react at a surface and contribute to film formation.

One can picture the PECVD process as a sequence of three distinct processes: the chemical activation of a gas molecule through electron impact dissociation, the transport of the radical species to the substrate, and the chemical reaction at the film surface (Figure 1). We will discuss each of these processes in some detail.

The relatively light electrons present in the plasma are easily accelerated by the electric field. The same electrical force acts on the ions but they are much heavier and accelerate slower. Electrons don't lose much energy in most collisions with the background gas. Due to the large mass difference between the electrons and the gas molecules, low-energy electrons collide elastically with the gas molecules, like marbles bouncing off a billiard ball. A plasma electron can therefore accumulate more and more energy in the electric field. Once an electron has gained enough energy, an inelastic collision may take place. Like a bullet shot into a billiard ball, the electronic bullet loses energy and the molecular billiard ball increases its energy and often dissociates. In such an inelastic collision, the electron loses several electron volts (eV's) of energy to the gas molecule—Figure 2(a). This energy can cause a gas molecule to rotate and vibrate; it can also electronically excite the gas or cause it to dissociate or ionize. Many of these inelastic interactions increase the "chemical activity" of the gas molecules by several orders

Fig. 2.



(a) The energy vs. time for a single electron undergoing numerous elastic collisions and an inelastic collision; (b) A typical energy distribution of the plasma electrons. $N(E)dE$ is the number of electrons with energy between E and $E + dE$.

of magnitude. The ensemble of plasma electrons is characterized by a (nearly) Maxwellian electron energy distribution [3]—Figure 2(b).

The mean electron temperature of a typical glow discharge is equivalent to an energy of between 1 and 3 eV. This energy must be compared to the energy needed

to dissociate or ionize the gas, an energy range roughly between 3 and 20 eV. Therefore, only the electrons in the high-energy tail of the distribution chemically activate and ionize the molecular gas. It takes much less electron energy to dissociate a molecule and form two neutral fragments that it takes to overcome the strong coulombic forces to produce an ion and electron pair. The shape of the electron energy distribution suggests that there are far more plasma electrons with the relatively low energies at which radicals can be formed than there are electrons capable of ionizing the gas. Indeed, in a PECVD plasma, the density of neutral radicals is much higher than the density of ions but a physically correct argument should also take the various cross sections into account.

Densities typically deduced from observations on PECVD plasmas [4] are approximately

10^{16} cm^{-3} molecules in a gas at 1 torr
10^{14} cm^{-3} neutral radicals
10^{10} cm^{-3} electrons and ions

From these numbers, you can see that a glow discharge plasma is a very dilute phenomenon.

Gas flow velocities in most practical PECVD systems are so low that neutral radicals created in the plasma find their way to the surface of the film by diffusion. The concentration gradient driving the diffusion process is maintained by radical generation in the plasma (10^{14} cm^{-3}) and their disappearance at the film surface (0 cm^{-3}). This concentration gradient exists over a linear dimension approximately corresponding to width of the dark space or sheath. For reasonable estimates of the diffusion coefficients, one can calculate deposition rates to be typically of the order of 10 Å/s, in agreement with observations [5]. Much higher deposition rates can be obtained in the directional flows of convective systems. These systems require special gas inlets, flow patterns, and high pump speeds.

Once the radical arrives at the surface of the growing film it participates in a chemical surface reaction. PECVD resembles the well-known process of plasma etching and differs mainly in the outcome of the surface reaction. In etching processes, the reaction produces volatile species that leave the surface. In the case of PECVD, the reaction bonds at least part of the radical to the surface. Volatile reaction products, such as H_2, are often formed and escape. These chemical reactions take place under unusual energetic conditions: both temperature and ion bombardment affect the outcome of the reaction. PECVD takes a unique place among thin-film deposition processes because ion bombardment during film deposition is an inherent part of the process. Ion bombardment and substrate temperature, controlled by equipment and process design, are the most important parameters that determine the properties of the film.

The chemical composition of PECVD films doesn't fit the chemical nomencla-

ture as being pure silicon, simple oxides, nitrides, carbon, or carbides. Films always contain fragments of the precursor ligand groups, usually hydrogen, hydrocarbon groups, fluorine, or chlorine. Although this may be discomforting from a chemical purist's view point, materials scientists have learned that these solids have useful properties, often amazingly close to their chemically pure and stoichiometric counterparts. Physicists have realized that an incorporated ligand group such as hydrogen can make the difference between an interesting and technologically potent material such as a-Si:H, and its electronically less useful counterpart a-Si. The process engineer manipulates temperature and ion bombardment to control the incorporation of ligand groups and tailor the properties of the film. Although many different material compositions fit the single designation of, e.g., plasma oxide or plasma nitride, Table 1 lists the most important materials deposited by PECVD.

The first commercial plasma reactors were developed in the early 1970s to deposit silicon nitride films [22]. Silicon nitride passivation layers replaced the silicon oxide/metal can packaging technology, used until then to encapsulate integrated circuits, and enabled the use of plastic packaging. At about the same time, high-resistivity hydrogenated amorphous silicon was shown to be amenable to doping [23], analogous to crystalline silicon. The development of amorphous

Table 1

Examples of Materials Deposited by PECVD

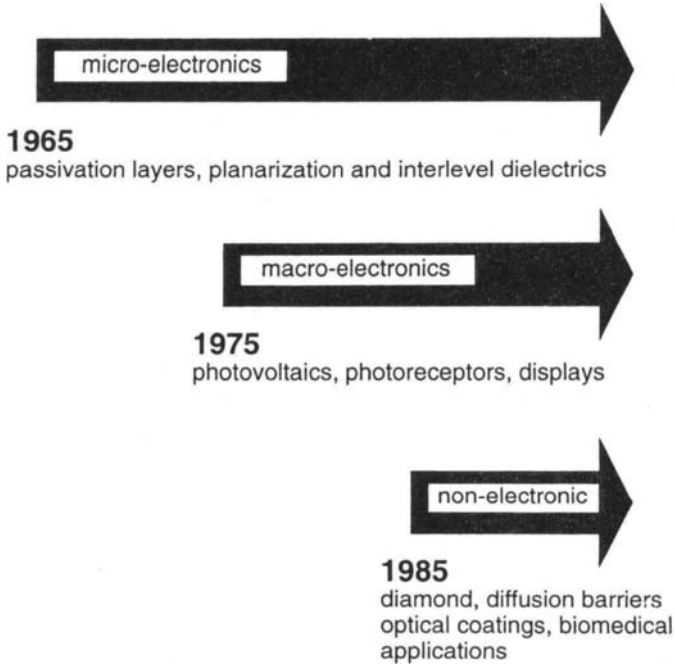
	Material	Made from
Semiconductors	Amorphous [6] and polycrystalline silicon [7]	SiH_4, Si_2H_6
	Silicon–germanium alloys [8]	SiH_4 + GeH_4, GeF_4
	p- and *n*-type a-silicon [8]	SiH_4 + B_2H_6 or PH_3
	Diamondlike carbon [9,10]	Hydrocarbon gases
	Diamond [9,10]	CH_4, H_2, O_2
Insulators	Plasma nitride [11]	SiH_4 + (NH_3, N_2)
	Plasma oxide [11,12]	SiH_4 + (N_2O, O_2) or organosilane + oxidizer
	Silicon–carbon alloys [13]	SiH_4 + hydrocarbon or organosilane + hydrocarbon
	Diamond, a- carbon [9,10]	Hydrocarbons, H_2
	Boron nitride [14,15]	B_2H_6 + NH_3, borazine
	Titanium oxide [16]	Tetraisopropyltitanate + O_2
	Aluminum oxide [17]	Aluminum acetylacetonate + O_2
	Tantalum oxide [18]	$Ta(C_2H_5O)_5$
Conductors	Copper	Copper acetylacetonate [19] Copper β-diketonates [20]
	n + silicon [21]	SiH_4 + PH_3

silicon as a dopable semiconductor material spawned several new device technologies [24] because it now became possible to deposit electronic devices with dimensions larger than a silicon wafer on low-temperature substrates ("macroelectronics").

Today, these early applications of PECVD technology are still important but with improved quality of dielectric films, applications of PECVD in microelectronics have expanded well beyond their initial use as passivation layers. PECVD films are now also used for surface planarization [25] and intermetal dielectrics [26–28]. The deposition of dielectrics at low temperature is particularly important for the processing of GaAs [29] and InP [30]. These compound semiconductors don't form good insulating oxides, and high-temperature processing causes the preferential evaporation of the group V element, resulting in the formation of defects. PECVD continues to be an important deposition technology for photovoltaics [31], xerographic photoreceptors [32], and thin-film transistor-based devices [33]. Most importantly, amorphous silicon thin-film transistors are the dominant pixel switch for large active matrix displays. Their relatively low electron mobilities of 1 cm^2/Vs (compared to polycrystalline silicon mobilities of 30–150 cm^2/Vs) is presently not limiting display performance. Rather, the high parasitic capacitance, hence transistor size, needs to be decreased from several μm's gate lengths to achieve shorter switching times, requiring innovative lithographic techniques.

PECVD is used increasingly for nonelectronic applications. Polycrystalline diamond films offer technical promise for a variety of applications, including heat sinks, X-ray windows, and tool coatings. The deposition of diamond requires an abundance of atomic hydrogen, which can be obtained from the dissociation of molecular hydrogen at microwave frequencies or in plasma jets. Another nonelectronic application of PECVD is the deposition of thin silicon oxide diffusion barriers to reduce the permeation rate for oxygen and water vapor through polymers used for packaging applications [34]. Thicker oxide layers have been developed as scratch-resistant coatings on polycarbonate and acrylic substrates [35]. This technology has been commercially applied to ophthalmic lenses [36]. Using the conformal, low-temperature characteristics of the PECVD process to advantage, abrasion-resistant nitride coatings have been deposited on machine tools to increase their useful life by an order of magnitude [37]. Capitalizing on the ease by which multilayers are deposited, PECVD has been used for the fabrication of rugate filters [38]. The availability of novel metallo-organic precursors [39] continues to expand the applications of PECVD, particularly for optical applications such as reflection modification coating. The performance of a polymer used in clinical practice can often be improved by PECVD coatings, formulated to elicit specific cellular responses. Many of the biomedical applications require the deposition of films on difficult shapes such as tubes [40] and tubing [41]. PECVD tech-

Fig. 3.



Time line of the most important commercial applications of the PECVD process.

nology allows uniform deposition on these difficult substrate shapes, which are often impossible or impractical to coat by physical vapor deposition techniques.

The time evolution of PECVD for the various applications that have been discussed, is schematically shown in Figure 3.

5.5.2 EQUIPMENT AND OTHER PRACTICAL CONSIDERATIONS

Fueled by a strong commercial incentive, PECVD technology has rapidly evolved over the past few decades. Many specialized systems have evolved from the simple diode reactors used in the early days. Multiple-wafer batch systems have been replaced by single-wafer cluster tools to increase process reproducibility and to reduce process risk. Small footprint tools are particularly important for clean-room-based manufacturing of semiconductors and displays. Also the in-line systems historically used for solar cell and display manufacturing are now

yielding to cluster tools requiring less floor space. Roll coaters to deposit photovoltaic films or barrier coatings on plastic web, and batch reactors to coat ophthalmic lenses and plastic containers, are further examples of PECVD systems designed for a specific process.

For any system, at least four process parameters can be varied and must be controlled: total gas pressure, gas flows, electrical power, and deposition temperature. These primary process parameters constitute an enormous parameter—space of operation. We will discuss the choice of operating conditions in the context of two problems often encountered with PECVD processes: powder formation and stress.

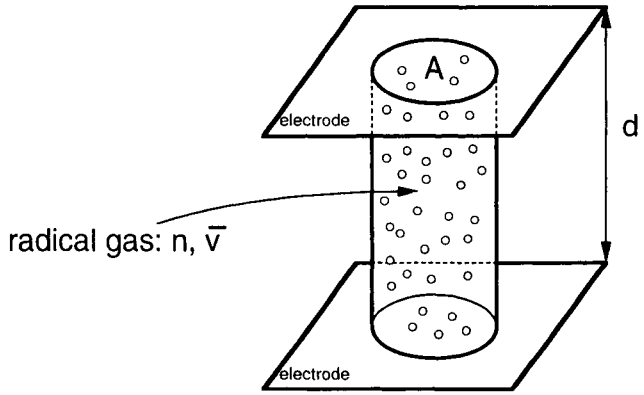
Controlling particulate contamination poses a special challenge to PECVD in general and to semiconductor processing in particular. Cylinder gases, opening and closing of gas valves, movement of substrate transport mechanisms in vacuum, heating and cooling, venting and purging all contribute to particle levels that ultimately decrease the process yield. Contamination from these sources can be minimized by engineering solutions such as point-of-use filtration, proper choice and treatment of vacuum materials, and process procedures. Technology and operating practice have progressed to the point where the process itself has become the limiting step to reduce particle counts [42]. Process-induced particulate contamination limits the device yields as geometries become smaller; the low-pressure processing environment inside a plasma reactor can have higher particle counts than the clean room that houses the system.

At typical PECVD process pressures, a radical created in the plasma undergoes many collisions before reaching the surface of the film, most likely with the source gas, because these molecules are by far the most abundant. Within a few collisions, there is an end to any chemical reaction between the radical and the source gas, very different from radical–radical interactions that keep on going [43]. The probability that a radical collides with another radical is small but finite, and the reaction product is likely to be also reactive. If the process repeats itself, the molecule continues to grow and becomes a macromolecule, then a particle. In PECVD processes, film formation at the substrate always competes with the undesirable tendency of the neutral radicals to react with each other in the gas phase: one can only control the degree to which this undesirable process happens. By properly choosing process parameters such as electrical power, pressure and the residency time, one can control the degree of gas-phase reactions. These gas-phase reactions are inherent to the PECVD process and determine to an important degree the properties of the film.

The relative probabilities of the various processes can be quite easily estimated (Figure 4). For a radical density n, the impingement flux on the electrodes is known from kinetic gas theory to be $\frac{1}{4}\,nv$ where v is the molecular mean speed. The radical–radical collision frequency is $\sqrt{2}nv\sigma$ where σ is the molecular cross

Fig. 4.

A cylindrical volume of radical gas with density n and moving at an average molecular speed v between two electrodes with distance d.



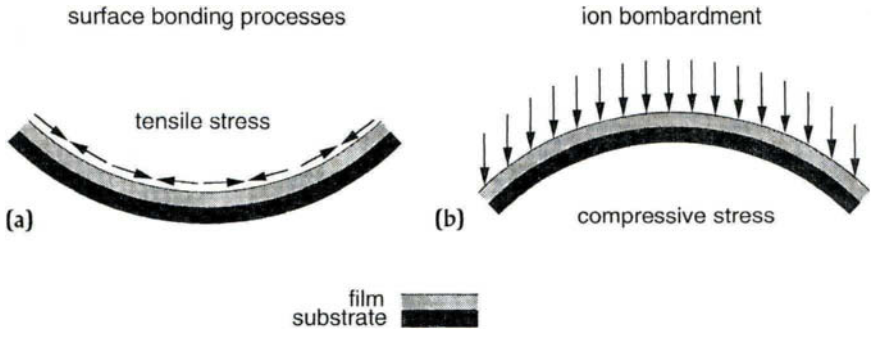
The electrode area intersected by the cylinder is A.

section. For an interelectrode distance d, the number of radical collisions in the volume between the electrodes, per unit electrode area and unit time, equals $\sqrt{2}n^2v\sigma d$. The ratio between the gas phase reaction rate and the condensation rate is therefore $4\sqrt{2}n\sigma d$. For a radical density of 10^{14} cm^{-3}, a molecular diameter of 2.10^{-8} cm and an interelectrode distance of 3 cm, the gas-phase reaction rate is approximately the same as the radical condensation rate. Under these typical conditions, some radicals will undergo multiple reactions in the gas phase to form relatively big particles observable as powder.

Low power and process pressure reduce the density of the radical gas, and a short interelectrode distance reduces the distance over which the radical must travel before it bonds to the film surface. Multiple collisions can be minimized by choosing these parameters properly, but powder formation has a finite statistical probability that can't be eliminated. Particles can be directly observed by light scattering techniques [44]. Using this method, much has been learned about their formation and behavior. In the plasma, macroparticles assume a negative charge and interact with the electric fields in the discharge zone. Under certain conditions, particles accumulate at the edge of the dark space and fall down on the wafer surfaces when the discharge power is shut off. Gas flow patterns and substrate orientation must be properly engineered to minimize this problem [45].

Film properties are strongly determined by the degree of ion bombardment during growth [46]. The PECVD process is unique in that the desired ion bombardment flux is easily engineered into the process without requiring auxiliary devices such as ion guns or additional electrodes. The amplitude of the applied voltage determines the maximum bombardment energy and the discharge current

Fig. 5.



(a) Surface-bonding processes during film growth, as well as annealing and hydrogen evolution after the film has been formed, induces tensile stress; (b) ion bombardment reduces the tensile stress and may induce compressive stress.

is a measure for the ion flux. Both frequency of the applied power [47] and the gas pressure [48] affect ion energy and flux in predictable ways. Ion bombardment during deposition at low temperatures yields film properties similar to high-temperature material deposited without bombardment. Ion bombardment, however, induces compressive stress in the film — see Figure 5(b) — because the film wants to occupy more space than the substrate allows. When for thick films the compressive force (stress × film thickness) exceeds the forces of adhesion at the interface with a rigid substrate, adhesion failure catastrophically relieves the stress. Without ion bombardment, many amorphous films exhibit tensile stress. Tensile stresses — Figure 5(a) — are often related to the formation of strained surface bonds formed during film growth. Hydride films often exhibit tensile stress after they have been heated above their temperature of deposition. The stress is caused by hydrogen evolution, causing film densification and often film cracking. Long before these unhappy endings, excessively high stresses ($>>10^9$ dynes/cm^2) cause many problems in thin-film devices and components, such as substrate deformation, cracking, void formation, enhanced electromigration, and piezoelectric charging effects. Understanding the relation of these stresses to the process parameters is key to the successful application of PECVD technology.

Symmetric reactors have the same ion bombardment on both electrodes, consequently the material harvested from both electrodes is the same. Most plasma reactors are purposely designed asymmetric to maximize or minimize the degree of ion bombardment. However, about half the material deposited in these anisotropic processing environments is lost. An asymmetric reactor design often has a large grounded electrode and a capacitively coupled powered electrode. In this configuration, the grounded electrode is subject to minimum ion bombardment and the small powered electroded receives high ion bombardment.

Dense and hard materials are deposited under conditions of high bombardment

flux and energy. These conditions are found at the cathode of DC systems and at the powered electrode of a capacitively coupled RF system. Ion bombardment can be tempered by decreasing the discharge power to the process, increasing the frequency or the process pressure.

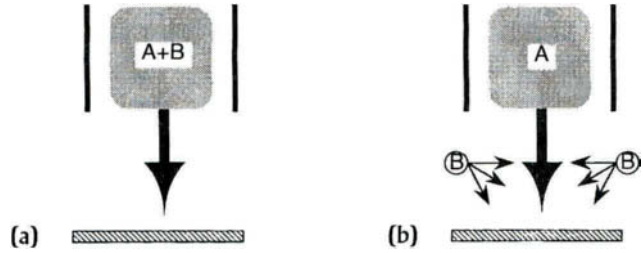
At low frequencies an ion can traverse the dark space during one half-period. Now the pressure is the main parameter to temper the ion energy, and the pressure ultimately determines if a film will be deposited or sputtered away. The maximum bombardment energy is the voltage amplitude. With increasing pressure, the average ion bombardment energy decreases due to intermolecular collisions; the ion bombardment flux, however, increases. If the frequency is higher than the corresponding time that it takes an ion to traverse the sheath, the ion bombardment energy also decreases [49]. The reason for this is that for a fraction of the cycle the ions are accelerated by the relatively low plasma potential rather than the cathode potential. Independent control over flux and energy is obtained in mixed-frequency deposition systems. In these systems, the plasma density (ion flux) is primarily determined by the high-frequency component and the bombardment (ion energy) is determined by the low-frequency bias [50].

Many device applications require minimum bombardment during growth. Interfacial charge trapping and electronic defect states in the bulk are often related to bonding defects due to ion damage. Low power, high pressure, and high frequencies are process conditions limiting the energy flux to substrates attached to a large grounded electrode of a capacitively coupled system. The bombardment energy on the grounded electrode is the average plasma potential, usually low for this system configuration [51]. Inserting a grid electrode to create a field-free region (triode system), can further decrease the bombardment energy if plasma potentials are unacceptably high.

So far we have assumed that the substrates are attached to one of the electrodes or are immersed in the plasma and electrically floating. In either case the film surface is exposed to the energetic plasma. Ion bombardment, ultraviolet light, and metastable atoms are capable of inducing structural changes that result in film densification and also in bonding defects. Defect centers act as electronic traps, degrading charge transport, photoconductivity, and interfacial charge densities. Similarly to plasma-assisted etching, ion bombardment during deposition affects the conformality of the film around steps and in trenches [52]. At high ion energies, resputtering of less dense material avoids "breadloafing" at the top of trenches. Depending on the application, anisotropy may or may not be desirable. Several plasma deposition methods, permitting some control over the process have been developed. These methods are schematically shown in Figure 6.

The simplest change is to move the substrate away from the energetic plasma — see Figure 6(a). In addition, one can change the way in which the precursor gases are chemically activated by moving the gas inlet of the condensable gas out of the plasma zone — Figure 6(b). Both these methods are referred to in the literature

Fig. 6.



(a) *Remote deposition*: The substrate is remote from the plasma. (b) *Remote reaction*: Gas B is chemically activated by gas A in a region remote from the plasma.

[53] as "remote." Here, we distinguish remote deposition — Figure 6(a) — from remote reaction — Figure 6(b). By moving the deposition zone away from the plasma, both methods permit control of the energetic processes during film growth and may yield a certain degree of control over the predominant species responsible for film formation. The details of how this selectivity is obtained are different for remote deposition and remote reaction. In either case, one pays for control by a decrease in the deposition rate.

In the remote deposition process — see Figure 6(a) — only the longest-lived condensible radicals survive the journey from the plasma to the substrate. Afterglow and downstream reactors are examples of remote deposition. Chemical specificity in remote reaction schemes, as in Figure 6(b), is achieved by a two-step gas activation scheme. Gases that don't form condensable radicals, such as He, Ar, O_2, O_3, N_2, N_2O, H_2, and NH_3, are activated in an intense plasma zone. Inductively coupled plasmas, helical resonator sources, helicon sources, hollow cathodes, microwave plasmas, and ECR methods have been used as plasma sources to activate the first gas. Molecular gases are vibrationally excited or dissociated, in which case the excited molecules or neutral radicals flow downstream. Atomic gases are likely to be electronically excited, in which case the long-lived metastable species carry the reaction energy. Again, the longest-lived radical will predominate at a downstream location where there is an auxilliary gas inlet. At this location, the specific noncondensable radical reacts with condensable precursors such as SiH_4, Si_2H_6, GeH_4, hydrocarbons, fluorocarbons, and organosilanes in supposedly predictable ways to form the radicals that build the film. In this way a measure of process control has been wrested from a process that in its one-step version contains many different radicals. Remote reaction schemes have been shown to yield unique and unusual film chemistries. Silicon dielectrics with very low hydrogen concentrations and of high electronic quality [54] have been deposited at moderate substrate temperatures, albeit at impractically low deposition rates. The rate disadvantage can be overcome in microwave-excited systems, but then the films contain hydrogen. ECR reactors are the most practical example of

a remote reaction process because ion bombardment yields films with low hydrogen concentration at practical rates.

Since their introduction [55] in the early 1980s, ECR sources have become increasingly popular for PECVD, particularly for the deposition of high-quality and conformal insulators. ECR permits the deposition of dielectric films at room temperature with properties similar to CVD films deposited at much higher temperatures. These low-deposition temperatures are enabled by the high-flux, low-energy ion stream that emerges from an ECR source. Where ion current densities for a glow discharge plasma are typically of the order of $1 mA/cm^2$ and have energies of several hundreds of eV, the ion flux from an ECR source is at least one order of magnitude higher and has ion energies between 20 to 50 eV [56]. The ion flux and energy are determined by the microwave power and the magnetic field gradient. The ion-to-neutral ratio is much higher for ECR deposition than for other PECVD methods. Another important aspect of ECR PECVD is that this deposition method is, by necessity, a remote reaction process. To keep the dielectric window clean from deposits, nondepositing gases are introduced through the resonance zone. The activated gas reacts with the condensable gases in a downstream region. As discussed previously, remote reaction can favorably affect the deposition chemistry and yield films with a low concentration of hydrogen. In ECR systems, remote deposition and ion bombardment work together to yield PECVD dielectrics at room temperature with properties similar to CVD material deposited at temperatures of 800°C.

ECR has been used for the low-temperature deposition of high-quality dielectric films for micro- and opto-electronic [57] applications. Intermetal dielectrics and planarization layers require excellent conformity. Conformal films of the requisite quality can be deposited by combining ECR with simultaneous sputter etching [58,59].

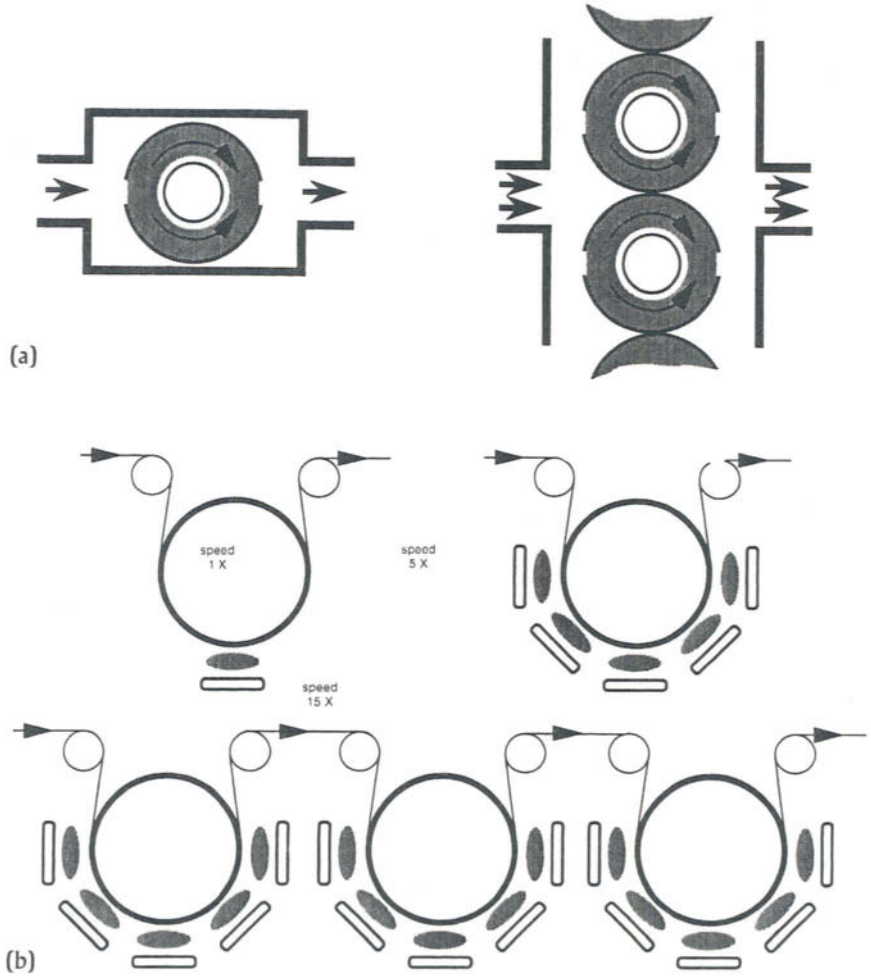
5.5.3 PROCESS SCALEUP

Transforming PECVD from a working laboratory process to a production process is generally referred to as "scaleup." Scaleup involves either increasing the substrate size or the process throughput. Both kinds of scaleup require an increase in reactor size. As large production reactors are expensive, scaleup should only be undertaken if there is a high probability of success. In this section we focus on the difficulties encountered in scaling up PECVD processes. Rather than discussing the details of particular applications, we will emphasize the aspects common to scaling up of most PECVD processes.

Plasma processes don't easily scale in size because the plasma volume has to be

increased. If all dimensions, power density, and flow rates of a small plasma reactor are multiplied by the same number, the outcome of the deposition process will be different. The low-risk path to scaleup is therefore "modular" scaleup, where one plasma environment (the plasma module) is fully developed to the right size. The required process throughput is then achieved by duplicating the module, placing them in series or parallel. Examples are drum cross-flow reactors [60] and in-line coaters [61] shown in Figure 7.

Fig. 7. Examples of modular production reactors.



(a) Drum reactor with gas inlet and exhaust plenums (parallel modules) and (b) an in-line roll coater (serial modules).

Modular scaleup minimizes the risk of meeting throughput demands, but the basic difficulties of scaling PECVD to the modular level remain. We will briefly discuss how the primary process variables of power, flow rate, pressure, and temperature change with size.

The electrical power scales nearly linearly with electrode size for frequencies <1 MHz or with the plasma volume for high frequencies. However, long before the free-space wavelength approaches the dimensions of the substrate, high-frequency processes pose a difficulty in scaleup. At 13.56 MHz, it is difficult to achieve uniformity when electrode sizes become >1 m. This problem can be solved by decreasing the frequency <1 MHz or using multiple, smaller-size synchronized electrodes. Pressure and substrate temperature usually remain unchanged in scaleup, and the greatest challenge in scaling a PECVD process is in properly engineering the gas flow pattern. Gas inlet systems have to be designed correctly to achieve the required uniformity, maximizing both deposition rates and gas utilization while minimizing particulate formation. Depending on the application, hollow electrode effects can be engineered into a distributed gas system to advantage [62] but can also be avoided by proper design. As a first rule, flow rates should be much larger than the desorption and leak rates of a system. The combined desorption and leak rate can be calculated from a pressure rate-of-rise measurement. Flow rates can be further scaled by requiring that the system doesn't operate in the gas-depleted mode; i.e., an increase in the total flow doesn't affect the deposition rate proportionally. Unless graded films are desired, the average gas residency time (process pressure $\times$ system volume/gas flow rate) must be much smaller than the total deposition time. Moreover, in a materials-efficient process, the molecular diffusion length during the residency time must be much longer than the interelectrode distance.

The hydrodynamic (= flow) part of the process is amenable to mathematical modeling. Such models have been developed for a variety of reactor configurations [63]. Some aspects of these models have been experimentally verified for cross-flow [5] and longitudinal cylindrical reactor configurations. In practice, flow rates are adjusted to ensure the uniformity of film thickness and materials properties of the film over the substrate area. Film thickness nonuniformities in the flow direction generally result from depletion of precursor gases. Film nonuniformity from gas depletion may be overcome by increasing the flow rate. At high flow rates, the gas composition of the exhaust gas will approach the composition of the precursor gas. This approach to solving the nonuniformity problem obviously reduces the process efficiency, defined as the ratio between the film weight and the weight of the condensable gas used to deposit the film. In practice it is often more cost effective to increase the process efficiency by moving the substrates with respect to the plasma to compensate for gas depletion effects. Pulsed plasmas [40] offer another way to engineer around difficult flow problems by redistributing the gas during the off cycle; fresh gas replaces the partially depleted gas. At the end

of the off cycle, the gas composition is spatially uniform along the substrate. When the on cycle is short enough that the gas doesn't significantly deplete in the flow direction, film thickness uniformity is assured, albeit at the cost of materials efficiency.

In practice, scaleup is more complicated than making a few relatively simple adjustments of the primary process variables. A change of reactor geometry potentially affects any number of important intrinsic parameters. For instance, the plasma potential is influenced by electrode geometry and details of power coupling [51]. Plasma potentials have, through ion bombardment of the film and the coater walls, a profound effect on the film properties [64]. Electric field patterns, especially at the edges of substrates and electrodes, are nonuniform, causing an effect known as "picture framing." The wall temperature of scaled-up reactors is generally different from smaller prototypes. Gas retention or "memory" effects are likely to be different, particularly if the reactor walls are covered by thick deposits.

The ease with which plasma reactors can be cleaned should receive serious consideration in the design phase of production systems. Deposits accumulating around the deposition zone have to be removed regularly to avoid flaking and excessive outgassing. Flaking increases with the thickness of the deposit and temperature cycling. The problem of cosmetic film damage is particularly relevant to applications that require thick films, and to continuous processes that produce thick films, as in roll coaters. Special constructions such as mesh-covered walls minimize flaking. If the substrate is arranged between vertical electrodes, most of the particulates generated during the process will settle on the reactor bottom rather than on the substrates.

When the complexity of reactor construction does not allow the electrodes to be easily exchanged, film buildup can be removed by plasma etching. Fluorinated gases such as CF_4 and NF_3 are commonly used for *in situ* chamber cleaning of silicon, silicon oxide, and silicon nitride deposits. These gases, however, have long retention times, and their use may well require several "sweep runs" before a reactor is restored to a controlled condition. The required purity level of a system, as for instance measured by a pressure rate-of-rise measurement, depends on the application and the nature of the contaminant. Group 3 and 5 dopants affect the electrical behavior of silicon at ppm levels [65], whereas nitrogen, oxygen, and hydrocarbons have a much smaller effect at comparable concentrations. Nonelectronic applications are generally more forgiving of impurities than electronic devices.

Chemical memory results from long-term outgassing of the inside of a coater, often exacerbated by the buildup of deposits. This carryover of small amounts of gases from previous runs results in cross-contamination. This phenomenon potentially limits the performance of multilayered devices such as thin-film transistors, photovoltaic structures, and superlattices. Some gases are worse than oth-

ers. Diborane, fluorinated and chlorinated gases, and generally most high-vapor-pressure liquids such as ammonia and metalloorganics, are especially bad actors. In addition to procedural measures such as sweep runs, the problem can be minimized by using multichamber reactors and gas lines dedicated to one gas.

But even multichamber reactors where isolated sections are dedicated to the deposition of a single material, have their limitations. At the interface between two films, the material composition is determined by startup transients that result from the system responding to pressure changes due to plasma decomposition, plasma-induced outgassing, and temperature increases. With gas flows stabilized, 10 seconds or more may pass before the process reaches a new stationary state after the discharge has first been turned on [66]. The material deposited in this time is likely to be different from the bulk material. Startup transients especially limit the performance of interfacially sensitive devices such as photosensors and thin-film transistors [67,12].

5.5.4 CONCLUSION

The key advantage of PECVD is that semiconducting and dielectric films can be deposited at relatively low temperatures. Ion bombardment can be substituted for deposition temperature to achieve the required film density. Low-temperature deposition is important for several different technologies. The continuing decrease of microelectronic device dimensions is synonymous with a decreasing tolerance for high processing temperatures. On the other side of the device spectrum, PECVD enables the fabrication of electronic devices over very large areas, often on low-temperature substrates. PECVD on plastics is rapidly emerging as an important technology to enhance the barrier properties or the scratch resistance of the plastic substrate. Significant advances have been made in enhancing the biocompatibility of polymers with films deposited by PECVD techniques.

Process performance and equipment demands will continue to be dictated by the needs of the single-wafer processing technologies for the manufacture of high-density integrated circuitry. Process engineers have learned how to deal with the inherent weaknesses and shortcomings of the process. Particularly, the formation of particles can be minimized by proper choice of process parameters, reactor geometries and materials. It is well understood how to control and manipulate the film stress. Multiple-frequency technology has become important to decouple the chemical activation process from ion bombardment, allowing process windows to be widened while maintaining tight control over the stress.

With the increasing commercial availability of novel, high-vapor-pressure organometallic precursors at affordable prices, the applications of PECVD continue

to grow. Organosilanes are increasingly used as a lower cost and safer source of silicon instead of silane. Optical multilayer devices are made possible by the commercial availability of aluminum-, titanium-, and tin-based organic compounds. Fluorinated compounds of every description continue to fascinate materials scientists, particularly for the modification of surface properties. New applications, including diamond deposition, diffusion barriers, optical filters, abrasion-resistant coatings on polymers, powder coating, fiber coating, and biomaterials have pushed PECVD technology far beyond the early electronic applications.

Major gaps continue to exist in our detailed understanding of the deposition process and the way in which the process affects the properties of materials and devices. More guided by intuition than understanding, new reactor designs sometimes cleverly exploit our limited knowledge, often building on experience in different areas of glow discharge technology such as plasma etching. Afterglow reactors, hollow cathode discharges, helical resonators, ECR, and other microwave discharges have been extensively used in plasma etching but their application to plasma deposition is relatively recent. Afterglow deposition schemes now provide additional control over, and understanding of, the plasma chemistry. As the demand for high-density, large-area electronics and high-performance plastics continues to rise, the unique advantages of PECVD as a low-temperature deposition technology guarantee its increasing role as a thin-film manufacturing technology. Industrial demand for new processing strategies will fuel fundamental research at universities and improve our understanding of the critical processes in a field that is ripe with opportunities for technological and commercial innovation.

REFERENCES

1. G. Glockler and S. C. Lind, *The Electrochemistry of Gases and Other Dielectrics* (Wiley, New York, 1939).
2. F. Kaufman, in *Chemical Reactions in Electrical Discharges*, Vol. 80, edited by R. F. Gould (American Chemical Society, Washington, DC, 1969), p. 29.
3. A. T. Bell, in *Techniques and Applications of Plasma Chemistry*, edited by J. R. Hollohan and A. T. Bell (Wiley, New York, 1974), pp. 1–56.
4. F. Jansen, in *Plasma Deposited Thin Films*, edited by J. Mort and F. Jansen (CRC Press, Boca Raton, FL, 1986), chap. 1.
5. I. Chen and F. Jansen, *J. Non-Cryst. Solids*, **59, 60** (1983) 695.
6. W. Luft and Y. S. Tsuo, *Hydrogenated Amorphous Silicon Alloy Deposition Processes* (Marcel Dekker, New York, 1993).
7. D. Pribat et al., *Rev. Tech. Thomson-CSF*, **26** (1994) 73.
8. *Properties of Amorphous Silicon*, 2nd ed. (INSPEC, The Institution of Electrical Engineers, London, 1989).
9. A. Lettington and J. W. Steeds, *Thin Film Diamond* (Chapman and Hall, London, 1994).
10. K. E. Spear and J. Dismukes, *Synthetic Diamond: Emerging CVD Science and Technology* (Wiley, New York, 1994).

11. Third symposium on silicon nitride and silicon oxide thin films, *Electrochem. Soc. Proceedings*, **94-16** (1994).
12. S. V. Nguyen, D. Dobuzinsky, D. Dopp, R. Gleason, M. Gibson, and S. Fridmann, *Thin Solid Films*, **193/194** (1990) 595.
13. P. K. Shufflebotham, H. C. Card, and A. Thanaiklakis, *J. Non-Cryst. Solids*, **92** (1987) 183.
14. M. Kuhr, S. Reinke, and W. Kulisch, *Diamond and Related Materials*, **4** (1995) 375.
15. S. V. Nguyen, T. Nguyen, H. Treichel, and O. Spindler, *J. Electrochem. Soc.*, **141** (1994) 1633.
16. H. J. Frenck, W. Kulisch, M. Kuhr, and R. Kassing, *Thin Solid Films*, **201** (1991) 327.
17. Y. Zhao and H. Suhr, *Appl. Phys.*, **A55** (1992) 176.
18. S. O. Kim, J. S. Byun, and H. J. Kim, *Thin Solid Films*, **206** (1991) 102.
19. A. Hamerich, L. Lottermoser, and J. Müller, *Surf. Coat. Technol.*, **59** (1993) 212.
20. E. T. Eisenbraun, B. Zheng, C. P. Dundon, and P. J. Ding, *Appl. Phys. Lett.*, **60** (1992) 3126.
21. R. A. Street, M. J. Thompson, and N. M. Johnson, *Phil. Mag.*, **B51** (1985) 1.
22. R. S. Rosler, *Solid State Technology*, (June 1991) 67.
23. W. E. Spear and P. G. LeComber, *Solid State Commun.*, **17** (1975) 1193.
24. D. E. Carlson and C. R. Wronski, *Appl. Phys. Lett.*, **28** (1976) 671.
25. P. H. Singer, *Semiconductor International*, (March 1992) 44.
26. A. Nagy and J. Helbert, *Solid State Technology*, (January 1991) 53.
27. A. Nagy and J. Helbert, *Solid State Technology*, (March 1991) 77.
28. G. W. Hills, A. S. Harrus, and M. J. Thoma, *Solid State Technology*, (April 1990) 127.
29. R. Williams, *Modern Gas Processing* (Artech House, Boston, 1991).
30. G. J. Valco, V. J. Kapoor, M. D. Biedenbender, and W. D. Williams, *J. Electrochem Soc.*, **136** (1989) 175.
31. D. E. Carlson, R. R. Arya, M. Bennett, and L. F. Chen, *AIP Conf. Proc.*, Vol. 353 (1995).
32. J. Mort and F. Jansen, *Plasma Deposited Thin Films* (CRC Press, Boca Raton, FL, 1986), p. 187.
33. *Proc. Electrochem. Soc.*, 94-35, (Electrochemical Society, Pennington, NY, 1995).
34. H. Chatham, *Surf. Coat. Technol.*, **78** (1996) 1.
35. S. Menichella, C. Misiano, E. Simonetti, L. D. Carlo, and M. Carrabino, *Proceedings of the 37th Conference of the Society of Vacuum Coaters*, (1994) 37.
36. C. Wohlrab and M. Hofer, *Optical Engineering*, **34** (1995) 2712.
37. L. Shizhi, Z. Cheng, X. Xiang, S. Yulong, Y. Hongshun, X. Yan, and H. Wu, *Surf. Coat. Technol.*, **43/44** (1990) 1007.
38. S. Lim, J. H. Ryu, J. F. Wager, and T. K. Plant, *Thin Solid Films*, **245** (1994) 141.
39. D. C. Bradley, *Chem. Rev.*, **89** (1989) 1317.
40. F. Jansen and S. Krommenhoek, *Thin Solid Films*, **252** (1994) 32.
41. R. Monson, master's thesis, Department of Chemical Engineering, Brigham Young University, Salt Lake City, Utah, 1991.
42. C. Steinbruchel, *Phys. Thin Films*, **18** (1994) 289.
43. M. Shiratani, H. Kawasaki, T. Fukuzawa, T. Yoshioka, Y. Ueda, S. Singh, and Y. Watanabe, *J. Appl. Phys.*, **79** (1996) 104.
44. G. S. Selwyn, *J. Vac. Sci. Technol.*, **B9** (1991) 3487.
45. G. S. Selwyn, *J. Vac. Sci. Technol.*, **A10** (1992) 1053.
46. H. P. W. Hey, B. G. Sluijk, and D. G. Hemmes, *Solid State Technology* (April 1990) 139.
47. K. Köhler, D. E. Horne, and J. W. Coburn, *J. Appl. Phys.*, **58** (1985) 3350.
48. R. J. M. M. Snijkers, M. J. M. v. Sambeek, M. B. Hoppenbrouwers, G. M. W. Kroesen, and F. J. D. Hoog, *J. Appl. Phys.*, **79** (1996) 8982.
49. R. H. Bruce, *J. Appl. Phys.*, **52** (1981) 7064.
50. A. Tsukune et al., in *Extended Abstract: Properties of silicon nitride prepared by dual frequency RF plasma deposition* (Electrochemical Society, San Diego, CA, 1986), p. 580.
51. K. Köhler, J. W. Coburn, D. E. Horne, and E. Kay, *J. Appl. Phys.*, **57** (1985) 59.
52. J. Schlote, S. Hinrich, B. Kuck, and K. W. Schroder, *Surf. Coating Technol.*, **59** (1993) 316.

53. G. Lucovsky, D. V. Tsu, R. A. Rudder, and R. J. Markunas, in *Thin Film Processes II*, edited by J. L. Vossen and W. Kern (Academic Press, New York, 1991), Section IV-2.
54. G. Lucovsky, Y. Ma, T. Yasuda, C. Silvestre, and J. R. Hauser, *Jpn. J. Appl. Phys.*, Part I, **31** (1992) 4387.
55. S. Matsuo and M. Kiuchi, *Jpn. J. Appl. Phys.*, **22** (1983) L210.
56. K. L. Seaward, J. E. Turner, K. Nauka, and A. M. E. Nel, *J. Vac. Sci. Technol.*, **B13** (1995) 118.
57. S. Dzioba, *Mat. Res. Soc. Symp. Proc.*, **165** (1990) 91.
58. C. S. Pai, *J. Electrochem. Soc.*, **139** (1992) 850.
59. B. Fowler, D. Stark, J. Xie, P. McDonald, R. Carpio, and S. Akbar, *Proc. SPIE-Int. Soc. Opt. Eng.*, **2090** (1993) 203.
60. J. Mort and F. Jansen, in *Plasma Deposited Thin Films*, edited by J. Mort and F. Jansen (CRC Press, Boca Raton, FL, 1986), chap. 7.
61. H. Chatham, A. Smith, R. Smith, J. Rietzel, and C. Woolley, *39th Annual Technical Conference Proceedings of the Society of Vacuum Coaters* (Philadelphia 1996) 64.
62. M. Chang, J. M. White, D. N. K. Wang, and D. Maydan (Applied Materials, European Patent Application EP 303508, 1989).
63. I. Chen, *Thin Solid Films*, **101** (1983) 41.
64. J. L. Vossen, *J. Electrochem. Soc.*, **126** (1979) 319.
65. F. Jansen and D. Kuhman, *J. Vac. Sci. Technol.*, **A6** (1988) 13.
66. S. V. Nguyen and P. H. Pan, *Appl. Phys. Lett.*, **45** (1984) 134.
67. Y. Nakayama, T. Ohtsuchi, M. Nakano, and T. Kawamura, *J. Non-Cryst. Solids*, **77, 78** (1985) 757.

CHAPTER 5.6

Common Analytical Methods for Surface and Thin Film

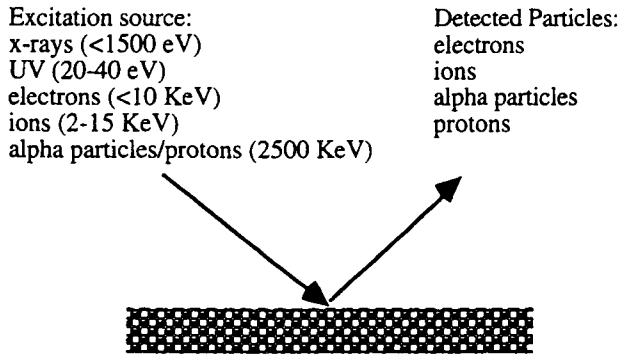
John H. Thomas, III
3M Corporate Research Laboratory

5.6.1 INTRODUCTION

The surface of a material is generally considered the topmost atomic layers or, on a crystalline surface, a monolayer (single atomic layer). To determine surface structure or chemistry, the surface under investigation must be kept in a pristine state such that the surface is not contaminated by atmospheric contaminates or residual gas components resulting from poor vacuum quality. In earlier chapters 4.3 and 1.8, it was shown that vacuum systems that were not properly precleaned could contain substantial amounts of contaminate residual gases from outgassing of components and the chamber and/or backdiffusion of gas components from the pumps; hydrogen, argon, methane, and water vapor are a few of the more common residuals. These residual vacuum gas components can readily react with almost any surface. For example, assume the surface is a single-crystal surface of a very reactive metal (Ti (100)). To produce a very clean surface, procedures such as *in situ* fracture, and sputtering followed by *in situ* annealing must be performed under very good vacuum conditions ($<10^{-10}$ torr). Measurements

ISBN 0-12-325065-7
$25.00

Fig. 1.



The analytical process of surface analysis.

must then be made before a monolayer of contamination forms on the surface, which generally takes only a short period of time, depending on the composition of the background gas in the system.

Most surface-analytical techniques are particle based; that is, they involve the detection of low- or high-energy particles (electrons and ions) that are produced by X-rays, ultraviolet irradiation, or electron or ion bombardment. Figure 1 summarizes the general methodology used to study surface structure and chemistry.

Chemical analysis of surfaces, the determination of elemental composition and some details of the chemistry existing on the material surface, is performed by two experimental procedures: the energy analysis of excited electrons, and the energy or mass analysis of ions. Instruments used in energy analysis of particles are based on electrostatics. Mass analysis is performed by one of three methods: magnetic field deflection, $\vec{E} \otimes m\vec{v}$ (RF fields) and electrostatic (time of flight) measurements. Table 1 summarizes the most commonly used analytical techniques and some important properties of the method.

The techniques just listed require the use of ultra-high vacuum. The following sections present an overview of these methods.

5.6.2

THE ELECTRON SPECTROSCOPIES

The electron spectroscopies are the most common surface analytical techniques used in industry, especially, Auger electron spectroscopy (AES). Both AES and X-ray photoelectron spectroscopy (XPS or ESCA) are used for submonolayer qualitative and quantitative surface analysis, and with the addition of ion sputtering and elemental imaging thin-film in-depth analysis. AES is most often used in

Table I

Common Analytical Techniques and Their Properties

Analytical Technique	Excitation Probe	Emitted Particles	Range in Z	Spatial Resolution	Detection Limits	Accuracy (%)
X-ray photoelectron spectroscopy XPS/ESCA)	1.5 keV photons	Slow e^- from core levels	$Z > 1$	10 μm imaging	~0.1 at. %	<5%
Auger electron spectroscopy (AES)	2–20 keV electrons	Slow e^- from core levels	$Z > 2$	500 Å imaging	~0.1 atomic %	~10%
Secondary ion mass spectrometry (dynamic SIMS)	5–20 keV ions	Slow ions	All	2–10 μm imaging	<1 ppb	Varies with Z and matrix
Secondary ion mass spectrometry (surface TOF SIMS)	5–20 keV ions	Slow ions	All	2–10 μm imaging	Surface <1 ppm	Varies with Z and matrix
Rutherford back-scattering spectrometry (RBS)	1–4 MeV α-particles	α-particles	All	2 mm	~1%	<2%

the depth-profiling mode but can be used to obtain chemical information. XPS can also be used in the depth-profiling mode, but as a result of the "X-ray beam size," depth profiling is much slower, because of the larger sputtered area required.

5.6.2.1 Introduction

Electron spectroscopy is produced experimentally by bombarding a surface with either electrons or photons (see Figure 1) and detecting the excited electrons in an ultra-high vacuum. The two most common spectroscopies are Auger electron spectroscopy (AES) and X-ray photoelectron spectroscopy (XPS) or ESCA (electron spectroscopy for chemical analysis), electron excitation, and X-ray excitation of electrons. Electrons emitted into the analysis vacuum chamber are ultimately collected and kinetic energy analyzed to determine the chemistry of a surface of unknown composition. The kinetic energy of the emitted electron ($\frac{1}{2}mv^2$) is commonly measured to an accuracy of 0.1 eV in 1000 eV (1 part in 10^4) using electrostatic energy designs [1–4]. A typical instrument consists of several components: an excitation source of either electrons or photons (X-rays), a means of analyzing the kinetic energy of the emitted particles (electrons), a detection and data acquisition system, and an ultra-high vacuum system. These components are discussed later in this section.

Electron spectroscopy provides qualitative analysis and quantitative analysis of material surfaces, thin-film elemental depth profiling, chemical surface analysis, and elemental mapping. These features and the instrumentation required to produce these effects are discussed in the following sections.

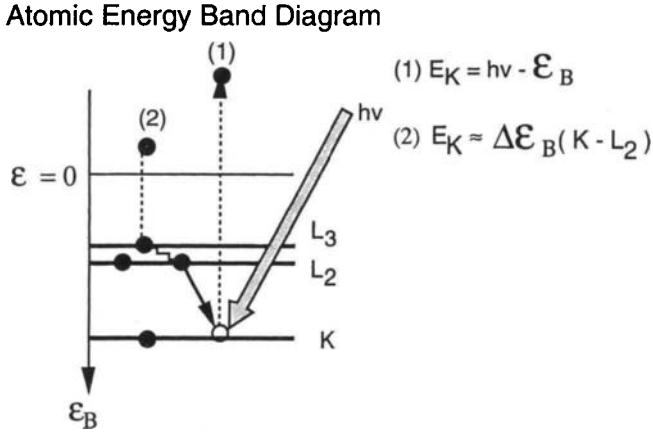
5.6.2.2 Qualitative Analysis

Qualitative analysis is performed through the analysis of the electron cloud emitted from a material surface excited by an electron beam or a low-energy X-ray source such as Mg K_α or Al K_α X-rays (1253 or 1468 eV, respectively). Materials that can be analyzed include metals (conductors), insulators (polymers, inorganic insulators, etc.), and semiconductors. Some organic materials can be damaged by electron or X-ray irradiation, but in general, most materials can be reliably qualitatively analyzed. As described next, analysis is straightforward, and spectral information can be obtained from well-documented handbooks [1–9].

The physical excitation process in Auger electron spectroscopy is the Auger transition, shown in Figure 2, where a bombarding electron inelastically scatters, for example, from a core K-level electron producing a K-level hole. The scattered and primary electron are free to leave the sample or to produce a cascade of secondary electrons through electron–electron interactions. To maintain charge equi-

Fig. 2.

(1) The photoemission process to produce a photoelectron;
(2) Auger emission process—an X-ray excitation of the atom is shown.



In an electron-excited process, primary electrons would scatter electrons from core levels and Auger transitions as shown would occur.

librium in the atom, one of two relaxation processes may occur. An electron from a shallower core level, L_2, will lose its energy by filling the deep-level hole in the K core level (generated by the scattering event) emitting an X-ray of energy $E_K - E_L$. Another process, known as the Auger process, can occur where the K hole is filled by the electron from the L_2 level by losing energy. The energy loss is synchronously coupled to an electron in a nearby level, L_3. The L_3 electron is excited to a kinetic energy approximately equal to the energy loss, $E_K - E_L$ and may leave the atom. The excited Auger electron is designated by the three levels involved in the process, KLL. The energy associated with the Auger transition is unique to the atom of origin and is used for chemical analysis. Typically, the core level is designated by standard spectroscopic notation (K, L_{12}, L_{23}, M_{34}, M_{45}, . . .) [6–10]. Note that Auger processes can also be excited by X-rays.

Since a large number of secondary and primary electrons are observed in the Auger spectrum, the signal is generally differentiated with respect to the kinetic energy. This removes a large portion of the background and produces a very sensitive method of observing the Auger transitions.

If the material is excited by nearly monochromatic X-rays, core-level electrons can be excited. The excitation process must conserve energy, and hence the kinetic energy of the excited electron $E_K \approx h\nu - E_B$ where E_B is the binding energy of the core-level electron designated by spectroscopic notation, K, L_1, L_2, M_3, M_4, etc. The core-level excited electron has a unique binding energy, depending on the atom of origin, and is used for chemical identification.

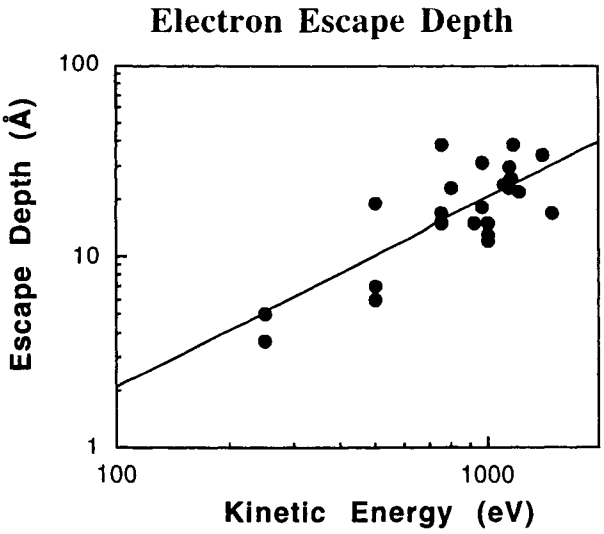
5.6.2.3 Surface Sensitivity Mechanism

When an electron leaves an excited atomic level in a material, it loses its original kinetic energy primarily through electron–electron interactions and thermal losses (phonons). The original electron kinetic energy is indicative of the atom of origin (qualitative analysis). After the electron loses energy, it is no longer readily identified with the original excitation and is called a *loss electron*. The range the electron traverses before losing its first quanta of energy through some process in the material, is known as the "electron inelastic mean free path" and is generally of the order of a few atomic layers. This range depends on material and kinetic energy. Electrons that escape from the surface can have energies equal to or less than its kinetic energy ($h\nu - E_B$). To identify surface species, the excited electrons that leave the surface must not lose kinetic energy. Only "elastically" scattered electrons leaving the surface add to the detected atomic core-level transition (either electron- or photon-excited spectra). Therefore, both Auger electron and photoelectron spectroscopy are surface sensitive for the same reason.

The electron "escape depth" is a function of electron kinetic energy, the mate-

Fig. 3.

Electron escape depths taken from literature references.



The straight line fit shows that escape depth depends on kinetic energy of the emitted electron. The usual power law dependence is $E^{0.75}$.

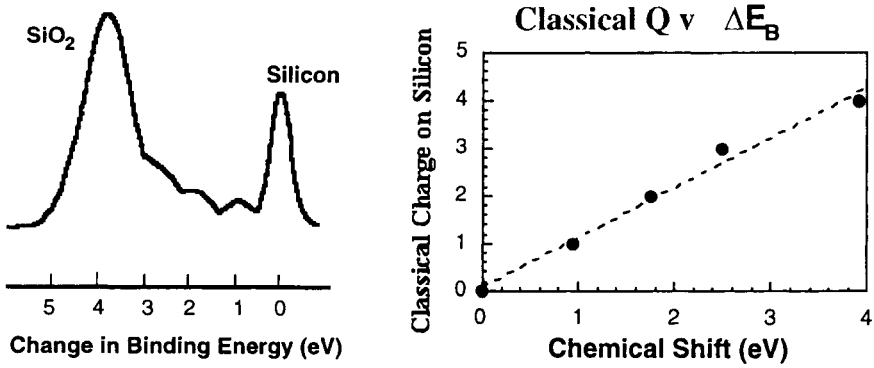
rial parameters and the physical geometry of the detection system relative to the sample surface. Figure 3 shows representative electron escape depths (at a 90-degree takeoff angle relative to the sample surface) as a function of the emitted electron kinetic energy.

5.6.2.4 Chemical Information (chemical shift)

In addition to providing near-surface region elemental analysis, electron spectroscopy can also provide limited chemical information. When an electronegative atom and an electropositive atom bond to form a compound, the electron charge distribution around the atoms redistributes to maintain electrical neutrality. The redistribution of charge produces a change in the core-level binding energies. Using photon-stimulated or electron-stimulated spectroscopy, the perturbed kinetic energy of a core level can be probed and appears as a shift in energy and/or a change in line shape. The perturbation is known as "chemical shift," and its development for use in ESCA (electron spectroscopy for chemical analysis) or X-ray photoelectron spectroscopy [1] resulted in Kai Seigbaghn receiving the Nobel prize in chemistry.

Fig. 4.

The spectrum of the Si 2*p* region was obtained using 130 eV photons from a synchrotron source.



The sample is a thin silicon dioxide layer on silicon. Interfacial states at the SiO_2/Si interface are believed to be due to silicon bonded in SiOx where x = 0, 1, 2, 3, 4. The chemical shift is plotted as a function of classical charge x and is nearly a straight line [11].

A simple example is shown in Figure 4 where the oxide states of silicon are compared. The binding energy of the Si 2*p* core level shifts ~5 eV from its elemental position. Using classical theory to compute the amount of charge present on silicon (electronegativity calculation), and plotting this as a function of the shift in binding energy (chemical shift) results in a nearly linear dependence. Therefore, the peak position along with quantitative chemicaly analysis can be used to identify functionality at the surface. This is mostly used in polymer chemistry, where the shifts and the shape of the C 1s peak are most useful in functional identification.

5.6.2.5 Quantitative Analysis

Both AES and XPS are reasonably quantitative techniques. Historically, elemental standards have been employed to calibrate an experiment. This requires the use of *in situ* cleaning of the standard to produce a clean surface for this purpose. Both methods obtain spectra in the N(E) mode, that is, direct electron counting is employed. Elemental sensitivity factors have been derived empirically and theoretically [2–6]. Over the periodic table of elements, the sensitivity factors may change by as much as 20. By the proper choice of the transitions used in the analysis, the effect of sensitivity variations can be minimized.

With elemental standards available, a good first approximation to quantitative analysis is possible by obtaining the elemental signal amplitudes from the un-

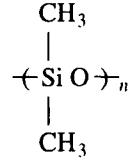
known and standard spectrum (usually by integrating the spectral region digitally). The unknown concentration is then obtained from the ratio of the unknown to the standard.

When standards are not available, quantitative analysis is still possible. By measuring all the elemental peak areas of the unknown, the normalized atomic concentration is given by

$$X_m^a = (I_m^a/\sigma^a) \cdot \left(\sum_{i=0}^{n} I_m^i/\sigma^i \right)^{-1} \tag{1}$$

where X is the normalized atomic concentration of element a in the matrix m, I is the amplitude of the i-th element in the matrix m, and σ is the elemental sensitivity factor. The sensitivity factors can be improved by including the element in a material similar to what is being analyzed. Generally, this expression provides a reasonably good normalized concentration. For further information, the reader is referred to References [2–7].

As an example of the application of this methodology, Figure 5 shows a typical XPS spectrum of an unknown surface. The elements are identified from peak binding energies. With the aid of software, the peak areas are integrated. The elemental peak areas are placed in a table and the normalization procedure, as shown in Equation (1), is performed. After the computation is performed, an elemental surface compositional table is available, as shown next for the unknown.
Compare this with what is expected from the chemical formula



where the O to Si to C ratio is expected to be 1:1:2. The experimental data are found to be in reasonable agreement. Therefore, the unknown is identified as silicone (polydimethylsiloxane).

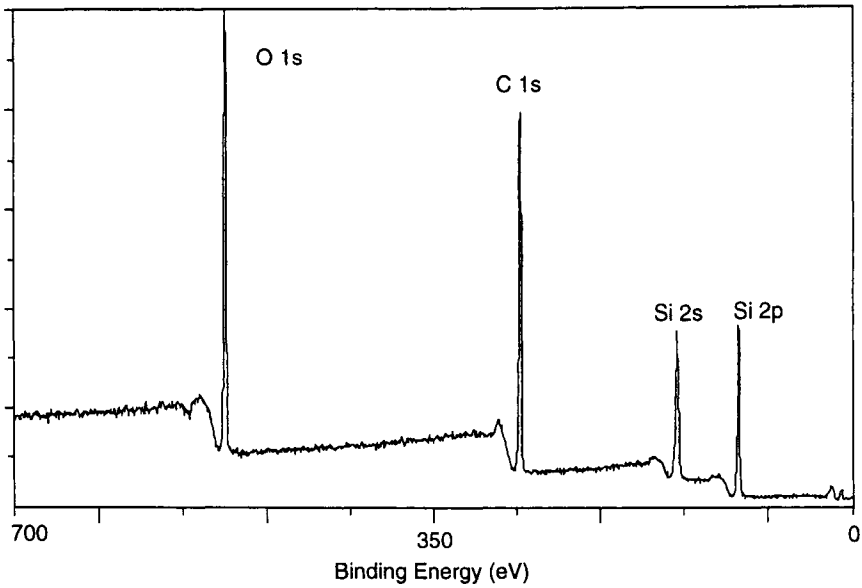
AES quantitation is very similar. The exception is when the data are obtained in the derivative mode, that is, the spectral data are $N(E) \cdot dN/dE$ versus E. Instead of measuring the peak area, the peak-to-peak signal amplitude is measured, and the appropriate elemental sensitivity factors for derivative data are applied.

5.6.2.6 Elemental Depth Profiling and Chemical Depth Profiling

Both AES and XPS are typically equipped with ion milling or sputtering using rare gases (He, Ne, Ar, Xe). Ion gun designs permit a focused beam of rare gas

Fig. 5.

Polydimethylsiloxane spectrum obtained using an aluminum K_α monochromatic X-ray source.



Elemental Composition Table

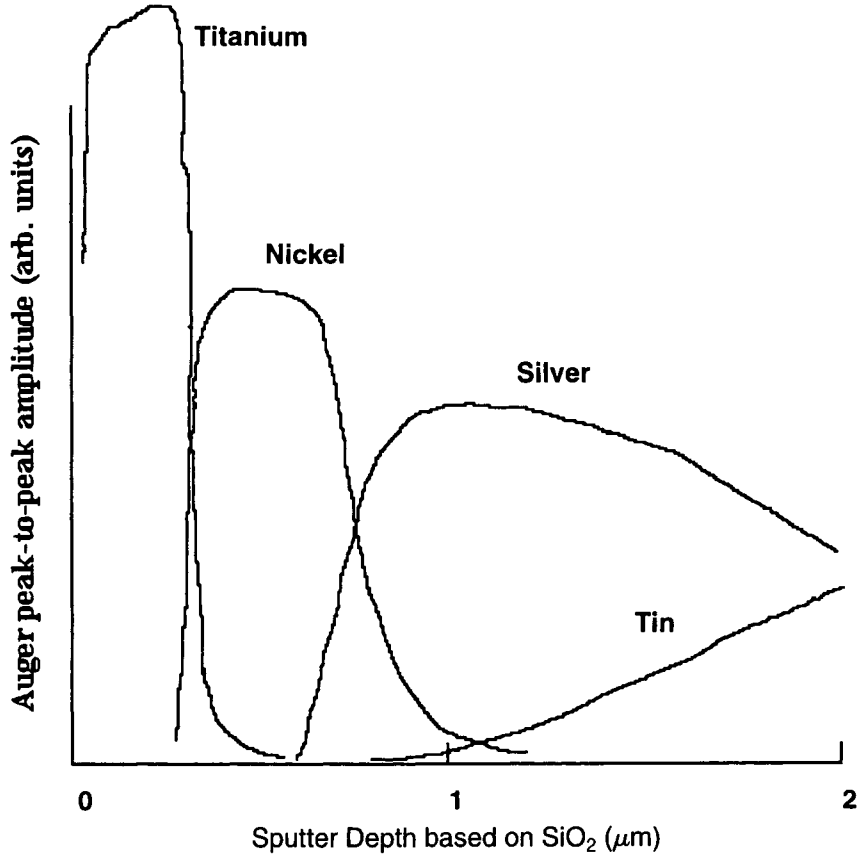
Element	Sensitivity Factor [12]	Peak Area	Atomic %
Si 2*p*	0.800	3000	27
C 1*s*	1.000	6550	50
O 1*s*	2.950	8400	23

The peak amplitudes are determined by integrating the area under the peaks.

ions to be produced with kinetic energy from 500 eV through 10 keV. The beams can be focused to spot sizes on the order of 10 μm but are generally used at larger spot sizes (depending on the ion energy). To produce an elemental depth profile that is accurate, the ion beam is rastered over an area larger than the analyzed area [3].

Auger elemental depth profiling is typically performed by obtaining the $N(E)$ spectrum, taking the derivative, and subsequently measuring the peak-to-peak amplitude of the measured peaks as a function of sputter time. Data may also be obtained from the peak amplitude minus a point near the peak to give a peak height. This information is used to obtain a normalized atomic concentration. Figure 6 shows a typical Auger elemental depth profile where the sputter time was

Fig. 6.



Elemental depth profile of a soldered connection to metalized silicon where the metalization was peeled off the silicon prior to analysis.

converted to depth in μm based on the sputter rate for SiO_2. The Auger peak-to-peak amplitude was not converted to elemental concentration.

It is possible to estimate the film thickness from the sputter time. However, this can vary widely depending on the matrix and the difference in sputter yields. The time scale can be converted to depth using thickness standards of a known material.

When X-rays are the excitation source, the photoelectron peaks are tracked as a function of depth. The peak areas are obtained and normalized using elemental sensitivity factors to provide a plot of atomic concentration as a function of sputter time. When the ion bombardment process does not degrade the surface, this technique can be used to obtain chemical information from chemical shift data as a function of sputter time.

5.6.2.7 Auger Elemental Mapping

One of the main features of using Auger electron spectroscopy as a surface analytical method is the use of the very finely focused electron beam obtainable with modern electron columns (similar to scanning electron microscopes). The electron sources are typically LaB_6 (single-crystal) or field emission sources. These sources provide a high brightness beam at very small spot sizes (<100 Å diameter beams).

Elemental mapping is accomplished by rastering the electron beam across the surface where at each point (pixel), the elemental signal amplitude is measured relative to the background. The background amplitude is used to correct for some of the topographic effects encountered in the analysis. These data are stored in a data file on a computer where it can later be manipulated using software. The picture can also be obtained in an analog fashion by photographing the Auger signal amplitude displayed as Z-axis modulation on an oscilloscope.

An Auger elemental map of silicon and carbon in silicon carbide fibers in a titanium matrix is shown in Figure 7. The fibers consist of a graphite core surrounded by SiC and a graphite coating encapsulating the SiC.

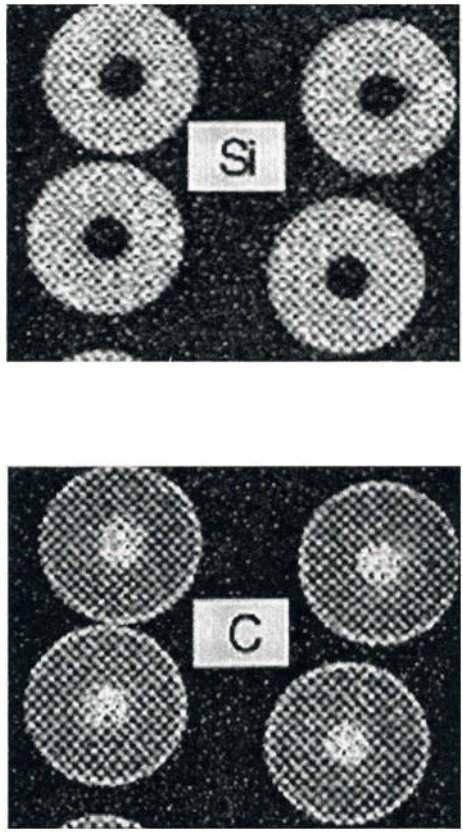
5.6.2.8 Instrumentation

Electrostatic analysis of the emitted electrons is performed using a number of common analyzers. In the following, three of the most common analyzers are briefly described.

The cylindrical mirror analyzer (CMA) is used mainly in AES systems. This instrument consists of concentric cylinders. The size of the outer cylinder is fixed depending on the inner cylinder diameter. Electrons are accepted into the structure where the outside cylinder is biased negative, and this voltage is scanned. The geometrical ratio of the cylinders fixes the energy resolution to ~0.2–0.5% of the kinetic energy.

The excitation source, generally an electron gun, is mounted coaxially in the inner cylinder of the analyzer but can be mounted externally to the CMA. Both assemblies are at ground potential. The electron gun can be a high-quality magnetically focused gun or an electrostatically focused electron gun. Both instruments are available on the commercial market. The coaxial geometry provides a small-size structure where the source is axially aligned with the analyzer. The sample for analysis must be at the proper location relative to the analyzer to angular restrictions. A typical simple AES system based on this analyzer is shown in Figure 8. Electrons are collected from a 360-degree cone at an angle of ~42 de-

Fig. 7.



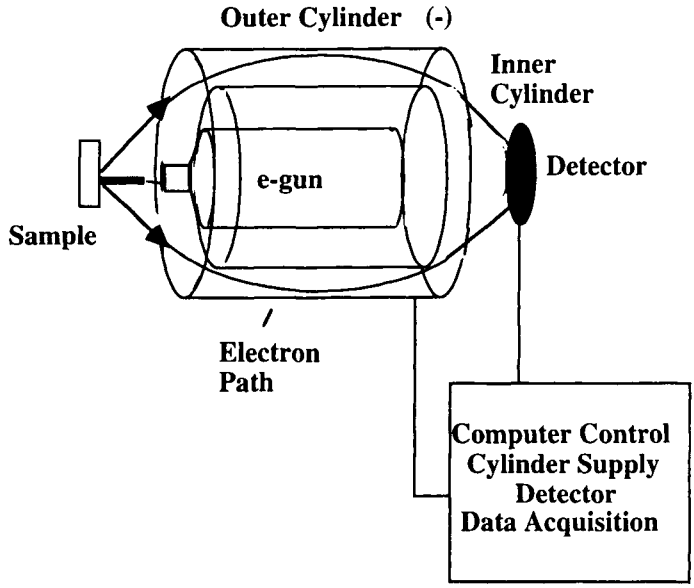
AES map of carbon and silicon in a multilayer fiber in a titanium matrix.

grees relative to the cylindrical axis. Other geometries degrade the amount of signal obtainable from the analyzer. Analyzed electrons are collected using a multichannel plate or simple multiplate electron multiplier. This information is used to produce spectra as shown.

The second most popular analyzer is the hemispherical sector analyzer (HSA). Historically, this instrument has been used with XPS systems due to its high-energy resolution. The HSA consists of two hemispherical sectors placed one within the other (Figure 9). The spacing of the sectors (the difference in the radius of the spheres) determines the basic energy resolution. As with the CMA, the outer sphere is biased negative, relative to the inner sphere. The HSA is equipped with entrance and exit slits, which may be either physical or electronic. The physical size of the slits determine the analyzed area.

Fig. 8.

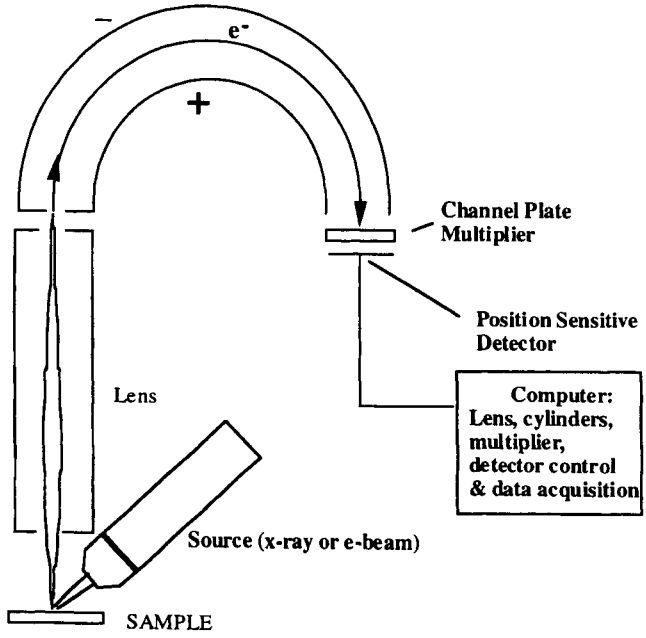
A simple AES single-pass CMA, detector, and e-gun electronics.



The detector can be an electron multiplier or multichannel plate type detector.

Fig. 9.

Hemispherical Sector Analyzer



A hemispherical-analyzer-based system used either for XPS or AES.

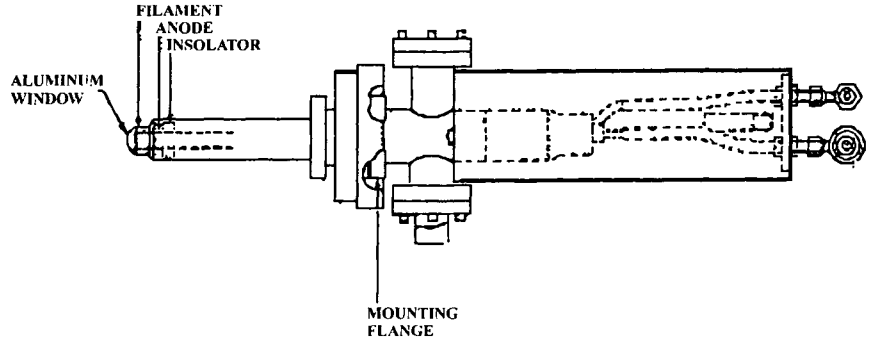
Generally, the HSA is equipped with a electron lens system that permits the sample to be placed at a remote location relative to the entrance slit of the analyzer. The use of a lens permits the electrons being analyzed to be either retarded in energy to a fixed energy (fixed analyzer transmission energy) and then analyzed by the HSA or to be analyzed by the HSA at a proportional energy to the kinetic energy (fixed retarding ratio). In addition, the angle subtended by the slits at the sample can be small (~6–8 degrees) depending on the design of the lens. This permits the sample to be tilted relative to the slits to perform grazing angle or angle resolved measurements. When the subtended angle is large, the depth of field of the analysis point is small, which can be useful in some applications. Large subtended angle designs are used to increase the collected signal amplitude.

In modern instruments, lens voltages, cylinder voltages, and detector voltages are all computer controlled by the computer used for data acquisition. This is convenient, because the instrumental functions are electronically controlled. The resolution of a CMA is for practical purposes fixed by the geometry of the instrument. This makes the HSA a more versatile instrument for high-resolution spectroscopy.

X-ray sources are simple diode structures in which the anode of the structure is typically aluminum or magnesium. The anode of the structure is generally run at high potential, while the cathode is at ground potential. Because of the high power employed in these units, the anode must be cooled. High-purity water systems are needed to minimize leakage currents to ground. The X-rays emitted from the anode are filtered with a thin foil of aluminum. This blocks secondary electrons emitted from the anode and protects the sample from high-energy electron bombardment. Figure 10 shows a typical non-monochromatic X-ray source.

Magnesium K_α and aluminum K_α X-rays are the primary sources used in XPS (1253.6 and 1486.6 eV, respectively). Both sources produce a main line with a full

Fig. 10.



A Physical Electronics Inc. X-ray source.

width at half maximum in the range of 1 eV (Mg is somewhat narrower in full width at half maximum [FWHM]). This is sufficient for XPS analyses.

Monochromatic X-ray sources are used in high-resolution studies. With a monochromator, X-ray line widths of less than 0.8 eV can be obtained. The most common monochromator employs bent quartz plates mounted on a Roland circle. X-rays produced on the circle are refocused at a point on the Roland circle. If the X-rays are produced with a finely focused electron beam, the X-ray spot reformed on the Roland circle is also focused. In this way, small-spot XPS can be performed. Sources with spot sizes of less than 100 μm diameter are presently available.

5.6.3 METHODS BASED ON ION BOMBARDMENT

In the previous sections, electron-beam-excited and X-ray-photoemitted electrons were used to analyze sample surfaces. More sensitive methods based on mass spectrometry were developed around the same time that XPS and AES were developed. In the following, secondary ion mass spectrometry used for both elemental depth profiling and chemical analysis is discussed briefly.

5.6.3.1 Secondary Ion Mass Spectrometry: Basic Mechanisms

Secondary ion mass spectrometry has become a mainstay in the semiconductor and polymer industries [13–16]. Through detection of low-energy ions emitted from the surface of materials, extreme sensitivity to the molecular composition is obtainable.

The SIMS technique uses incident medium-energy ions (1 to 20 keV) to sputter atoms or molecules from the surface of a material. The detected particle is an ion released from the near-surface region of the sample. To determine the identity of the atom or molecule released from the surface, mass spectrometry is employed. Both negative and positive ions are detected for a given incident ion beam. The positive and negative ion yields are not identical for a specific bombarding ion but are useful in unambiguous identification of the surface composition.

Two regimes have been studied, depending on the magnitude of the ion flux incident on the surface. When high-rate sputtering is used, the technique is known as *dynamic* SIMS. Dynamic SIMS can be performed using a focused ion beam to provide resolution to $\sim 1-2$ μm depending on the ion species and beam energy. By rastering the ion beam, an image can be formed of the detected ion species. Imaging is commonly performed in *static* SIMS, where the incident ion current is

minimal, so as to minimize sputtering and not disrupt the surface chemistry too severely.

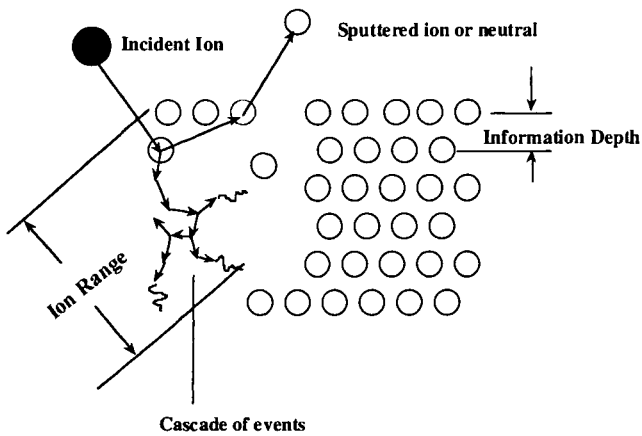
To identify surface atoms or molecules, mass spectra are obtained using a variety of instrumentation. Excitation is provided by finely focused ion beams of heavy "liquid metal" ions (cesium) or gases (argon). To ensure that the surface remains clean during the measurement, measurements must be performed under ultra-high-vacuum conditions. UHV conditions are required for the mass spectrometer and detection system. Under poor vacuum conditions, it is possible that reactions may occur during the ion drift time, producing a species unrelated to the original surface species.

Ion bombardment of a surface produces secondary ions through the loss of kinetic energy and momentum in a collision cascade of surface atoms or molecules within the primary-ion mean free path (or penetration depth) as shown in Figure 11. The secondary ion energy is at a low kinetic energy, and the distribution is not well defined. It depends on material properties in much the same way secondary electron distributions are formed by electron beams.

The penetration depth of the incident ion is a function of incident kinetic energy, the incident ion mass, and the material. Atoms or molecules that acquire momentum toward the surface of the sample can proceed into the vacuum as a sputtered species. Most particles leaving the surface are likely to be neutrals. Only a few ions leave the surface region for subsequent analysis.

Secondary ion yield is very sensitive to electropositive or negative ions at the

Fig. 11. Ion cascade produced by ion bombardment.



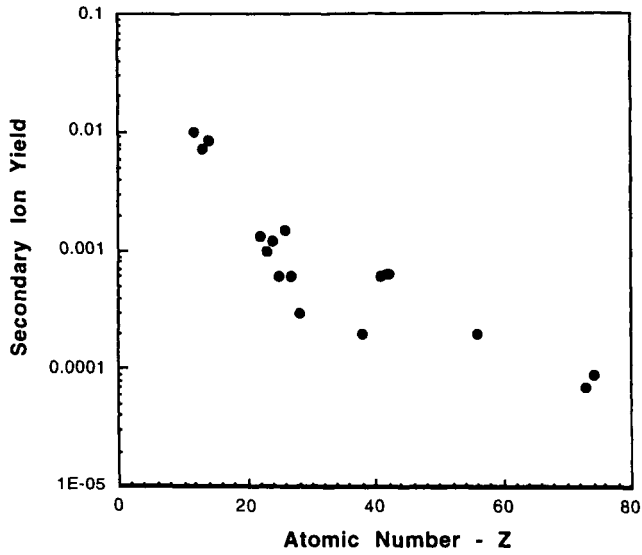
The bombarding ion will come to rest in the matrix after losing its kinetic energy. Near the surface (2–3 monolayers), neutrals and ions are ejected [13,16].

sample surface. For example, Cs^+ ions are used to lower the work function of the surface and enhance the production of negative secondary ions, whereas O_2^+ ions are be used to raise the work function of the surface, thereby increasing the yield of positive secondary ions. Common ions used are Cs^+, Ga^+, Ar^+, and O_2^+. Positive ions tend to be observed more frequently from electropositive surface species and negative ions are more frequently observed from electronegative species. In addition, the presence of electronegative species improves the positive ion yield, and the presence of electropositive species improves the negative ion yield. For a more detailed discussion of these effects, the reader is referred to References [16,17].

When the incident ion mass is very small and the ion kinetic energy is large (MeV), the sputter mechanism becomes minimal, and classical scattering of the incident particles occurs. This is known as Rutherford backscattering spectrometry (RBS) and is discussed later.

Ion bombardment is capable of sputtering all atoms from the periodic table, including hydrogen. However, the secondary ion yield is a strong function of the material (matrix effects) and varies ~5 orders of magnitude over periodic table. This is responsible for the very high sensitivity for specific atoms, where SIMS can detect some atomic species to the ppb range. A sketch of positive secondary ion yield using 3 keV Ar^+ is shown in Figure 12 (from Benninghoven's data, Reference [16]).

Fig. 12.



Some of Benninghoven's data (Reference [16]) plotted to demonstrate the variation in ion yield.

Unlike the electron spectroscopies, SIMS is not considered an "inherently" quantitative method. However, using appropriate standards in the dynamic mode, SIMS can be very quantitative. In some of the early quantitation reported, samples were implanted with the atom to be quantified. The implantation method yields a quantitative amount of implanted atoms/cm^2 (measured charge) and knowing the implantation energy (measured accelerating potential), the ion penetration depth can be obtained from tables [18]. Implants are Gaussian distributed to first order. Consequently, the peak of the implant profile as determined from the standard provides the concentration in a known matrix. This method is described in detail in Reference [19]. SIMS can also provide isotopic information. For example, hydrogen and deuterium are readily separated by most spectrometers.

5.6.3.2 Instrumentation: Dynamic SIMS

Dynamic SIMS is used primarily for depth-profiling thin-film structures. In addition to the ion source, this instrument requires an ion extraction lens close to the sample surface, a low-energy electrostatic energy analyzer, and a quadrupole mass spectrometer or other mass spectrometer. Larger ion current densities must be used to provide continuous sputtering, and in some cases very-high-rate material removal (Å per second) can be obtained using Cs, Ar, O, and Ga beams. The various sputtering ions change the sensitivity and ion yield. The ion beam is rastered over a region of interest on the sample surface, and the secondary ion signal is raster gated in a smaller region to eliminate or minimize crater edge effects, as shown earlier in the case of elemental depth profiling with AES. In this mode of operation, very good depth resolution can be obtained over large film thicknesses (μm).

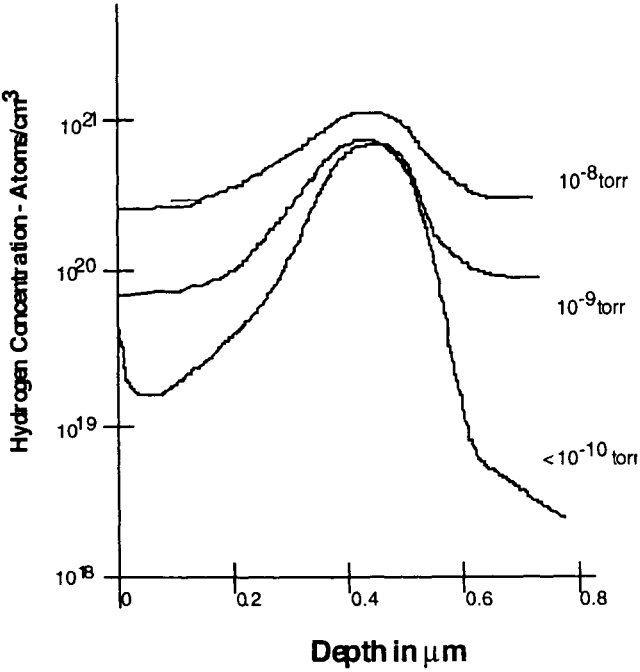
Ion signals are analyzed in a multiplex mode. Instruments are computer controlled, and the computer sets the amount of time spent per mass channel. The resulting profile is plotted as the number of counts in a channel as a function of sputter time. As noted earlier, the counts scale can be changed into a quantitative scale. The depth scale can be measured directly, using a surface stylus instrument after the sample has been sputtered. Conversion to concentration versus depth is therefore quantitative for simple materials systems.

Elemental depth profiling is generally performed over many thousand angstroms of material. Depth resolution may be very material dependent. For example, if the sample consists of palladium on silicon, it is likely that atomic mixing will occur as a result of the cascade energy transfer during profiling. In this case the depth resolution would be poor. Resolutions of less than 200Å over a depth of several tens of micrometers have been routinely achieved in other material systems.

Dynamic systems have been used to study ion implantation profiles with sensitivities over five orders of magnitude in crystalline materials such as silicon or

Fig. 13.

Sketch of a depth profile of hydrogen implantation in silicon was obtained at three partial pressures of H_2O using a Cs^+ 5 keV ion beam.



As the partial pressure of water is decreased, the sensitivity to hydrogen increases as a result of less interference (after Reference [19], figure 1).

GaAs. An early application of this method was to depth-profile hydrogen in silicon solar cells. It was found that the background hydrogen and water vapor in the analysis chamber could easily swamp the hydrogen signal from the sample. To prevent this, a vacuum of 10^{-11} torr is required [19]. To demonstrate this, Figure 13 shows a SIMS profile of a hydrogen implant in silicon (incident ion was Cs^+ 5 keV). The background pressure of water vapor was changed with remarkable results. Only the tip of the hydrogen profile was observed at 10^{-8} torr pressure. When the background pressure of water vapor was decreased, hydrogen was detected to below 10^{19} atoms/cm^3.

5.6.3.3 Static SIMS (time-of-flight SIMS)

Static SIMS is performed using the same type of apparatus as described for dynamic SIMS where the bombarding ion current is decreased to a level low enough to ensure that the sputtered ion (atom or molecule) has not been previously bom-

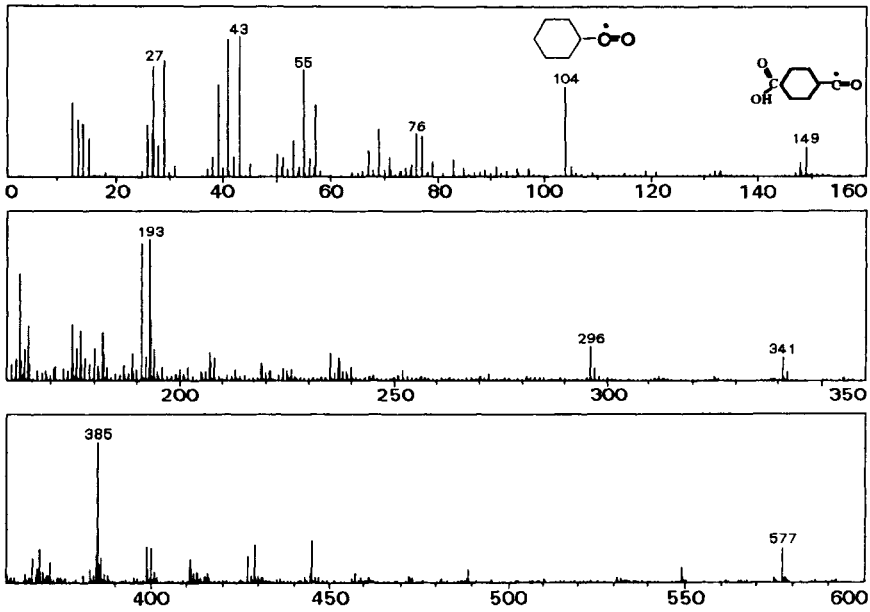
barded by a sputtering ion. This makes the static SIMS very surface sensitive. With the low current densities, the ion beam can be focused to ~1 μm diameter and rastered to obtain an ion or molecular image.

Static SIMS is performed in much the same fashion as standard mass spectrometry, where mass spectra are obtained and interpreted. Polymeric and organic materials can be probed directly with considerable mass resolution and sensitivity, permitting identification of the molecular ionized fragments and hence molecular species on surfaces. Static SIMS spectra are typically difficult to interpret unless the structure of the surface is rather simple. However, it has had amazing success in solving material problems. One approach is to use what has been termed "fingerprint" spectra method to identify material surface species. A typical fingerprint spectrum of polyethylene terephthalate is shown in Figure 14. Two of the positive ion fragments are identified [17]. By comparison of the major features of the spectrum with features of an unknown material spectrum, the unknown may be identified. Sophisticated computer programs have been and are being developed to perform this type of analysis.

In more recent instruments, time-of-flight mass spectrometers have been employed. These spectrometers offer the greatest sensitivity due to their large throughput and a very large mass range with very good resolution as compared

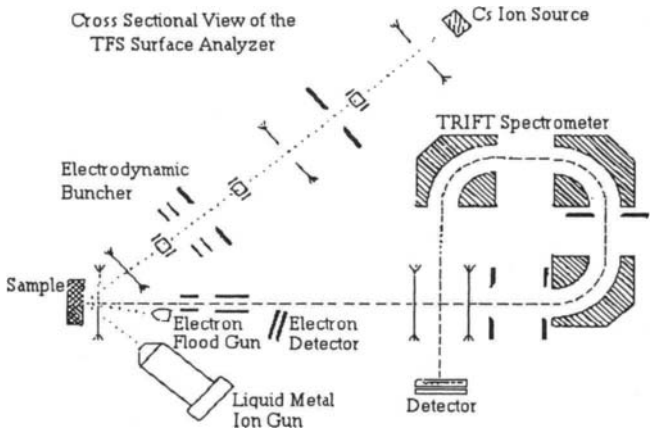
Fig. 14.

PET positive ion fingerprint spectrum.



The main mass/z peaks are noted on the spectrum. (Data courtesy of S. J. Pachuta, 3M CRL, 1996, obtained on a TOF-SIMS instrument.)

Fig. 15.
Schematic diagram of a TOF SIMS instrument (Charles A. Evans & Associates).



Both ion guns operate in a pulsed mode.

with instruments based on the quadrupole mass analyzer. Secondary ions are excited by a pulsed ion beam. Figure 15 shows a diagram of a typical TOF-SIMS (Charles Evans & Associates). Two other designs have been made and commercialized: the ion microscope and magnetic sector ion microprobe [16,17,20]. These instrumental designs have been available for some years. Compared to other instrumental designs, the TOF-SIMS instrument is simple and easy to operate. Computerization consists of data collection, which includes the mass spectra and data to provide an ion image, with higher current, elemental depth profiling can be performed.

Imaging using static SIMS has become a quality means of studying material surfaces. It is generally fast because it simultaneously collects the total mass spectrum. Imaging is limited to qualitative information, because it is difficult to obtain a "standard" image for quantification purposes.

5.6.3.4 Rutherford Backscattering Spectrometry (RBS)

Rutherford backscattering spectrometry (RBS) differs from the techniques described so far in that it uses high-energy, nuclear-physics-type apparatus. It is based on the use of high-energy ions as a probe and analyzes the same ions as they are scattered from the material being analyzed. References [13,18,21,22] present a good overview of the details.

From the physics of high-energy ion scattering, it is known that some particles such as α-particles (He^{++} or He^{+}) and protons (H^{+}) can scatter from a surface where classical physics applies. Generally α-particles are used, and scatter classi-

cally in the range of 1–3 MeV (10^6 eV). This means that when an α-particle scatters from an atom at the surface of a material, both the energy and momentum are conserved (Σ E and Σ mV = 0).

Classical scattering implies that α-particles scatter in a "billiard ball" fashion. To simplify analysis, a scattering event of 180 degrees relative to the incident beam is employed to identify the surface scattering atom(s).

At the surface of the sample, the energy of the α-particle is given by

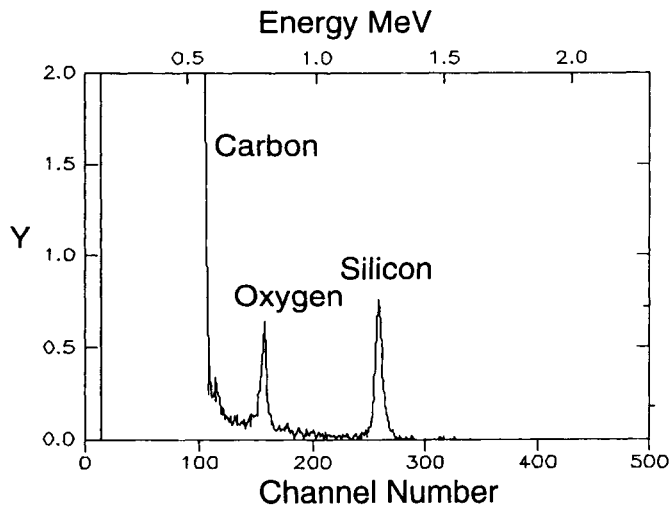
$$E_1 = E_0\left[1 - \left[\frac{M_\alpha M_x}{(M_\alpha + M_x)^2}\right](1 - \cos\theta)\right] \text{ where } \theta \approx 180° \qquad (2)$$

E_1 is the scattered energy, E_0 is the primary energy, and M_x is the unknown mass. Equation 2 reduces to $[(M_x - 4)/(M_x + 4)]^2$ at 180°. By comparing the energy of the backscattered α-particle to the primary beam energy, the scattering atom can be identified. The ratio of E_1/E_0 is known as the kinetic factor and is tabulated [18,21]. Elemental sensitivity depends on the Rutherford scattering cross section, which, for 180-degree scattering, is simply

$$\frac{d\sigma}{d\Omega} = 4(Z_\alpha Z_x e^2/4E_0)^2 \qquad (3)$$

where e is the electron charge (1.6×10^{-19} coulombs), E_0 is the beam energy, and Z is the atomic number. This expression becomes somewhat more complex when $\Theta \neq 180°$. An example is shown in Figure 16, where a very thin layer of sil-

Fig. 16.



The spectrum was obtained from a very thin layer of silica on a carbon substrate.

ica was deposited on carbon. The Si and O peaks correspond to E_1 as computed from Equation (3). The silicon and oxygen peaks are well separated out but are not very intense, as explained by Equation (4). The relative sensitivity becomes greater as the atomic number is increased and ranges two orders of magnitude over the entire periodic table. However, the ability to resolve the unknown energy becomes less as Z increases or sensitivity increases.

Thin-film analysis is possible because the α-particles scatter from atoms beneath the surface of a film. The α-particles lose energy, dE/dx, where x is the depth into the film, as they travel through the thin film (the energy loss occurs primarily through electron interactions). The energy loss is a function of the mass of the target and is known as the stopping cross section or power. dE/dx is presented in a computed table as a function on particle energy and the atomic number of the scattering atom. The units are presented in eV $\cdot$ cm^2/10^{15} atoms, and hence the depth scale for thin-film data is a function of the atomic density of the material. Film thickness determination is reasonably accurate and given sufficiently powerful software, the raw data can be directly converted to an elemental depth profile.

By performing a forward scattering experiment where the α-particles are permitted to strike the sample surface at a shallow angle, the scattered α-particles can be filtered out using a thin aluminum window and a particle detector placed in line with trajectory of the α-particles. This system will collect protons knocked out of the solid by the incident α-particles. The α-particles do not hit the particle detector but the protons can easily traverse the aluminum window and be collected. This is known as forward recoil scattering (FReS) and is used in obtaining hydrogen depth profiles.

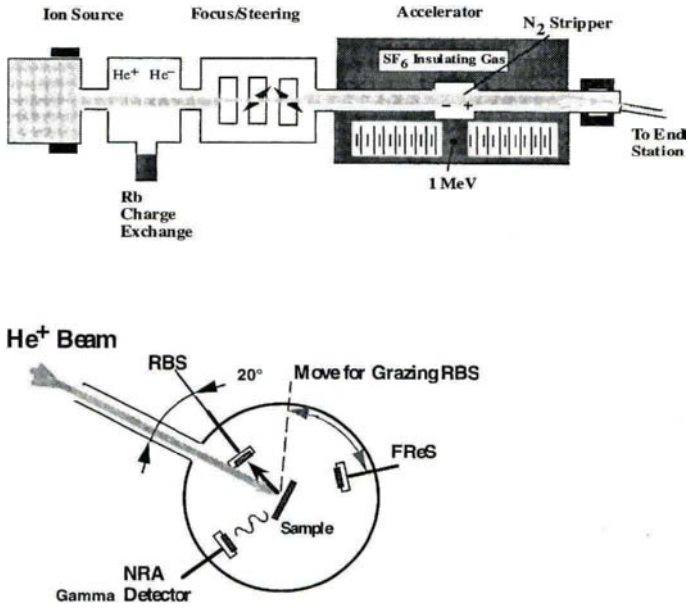
As the beam energy is increased, γ-rays are generated. They are the result of the α-particles interacting with the nuclei of the atoms. The γ-ray is elemental specific and is produced at a very specific incident kinetic energy. As a result, since the α-particles lose energy to the material system, a pseudo depth profile can be obtained by varying the incident beam energy and plotting the γ-ray yield as a function of energy loss that is converted to depth. The yield can be converted to a quantitative elemental concentration.

INSTRUMENTATION

The apparatus required to perform RBS and FReS (forward recoil scattering), is shown in Figure 17. The main portion of the system is the ion accelerator. This system electrostatically accelerates ions to a potential of < 4 MeV. The ion energy is determined by the accelerating voltage applied to the accelerator. He ions are accelerated up to 1 MeV and then stripped of electrons to produce either He^+ or He^{++}. The He^+ ions are accelerated to ground potential, dropping 1 MeV. In the case of He^{++}, the alpha particles obtain a potential of 2 MeV accelerating to

Fig. 17.

Typical RBS system using a high-energy accelerator and an end station equipped with the standard RBS particle detector, a nuclear reaction photon detector, and a forward recoil detector (used for hydrogen detection).



The end station is as supplied by Charles A. Evans & Associates.

ground. The total energy of the alpha particles is either 2 or 3 MeV for an applied 1 MeV accelerating potential.

The alpha particles are steered using a bending magnet. This acts as a selector to ensure a pure alpha particle beam. After beam alignment on the sample in the end station, most instruments are equipped with a collimator to restrict the beam size to 1–2 mm diameter at the sample.

The end station consists of a moderate high-vacuum chamber (10^{-8} torr). Better end stations work in UHV. The end station is equipped with two particle detectors and a γ-ray detector. The particle detectors are used to obtain the RBS and FReS or grazing-angle RBS data, while the γ-ray detector is only used for γ-ray detection. For these analyses, the sample is positioned using a computer-controlled goniometer stage. Some end stations are equipped with surface analytical apparatus such as AES to further increase the utility and flexibility of the instrument.

5.6.4

UHV GENERATION AND SYSTEM CONSIDERATIONS FOR SURFACE ANALYSIS

In the previous sections, it is clear that surface analysis must be performed in an ultra-high-vacuum ambient. Among the four techniques discussed, there is considerable parallelism in the analytical chamber construction, vacuum requirements, etc.

Over the past 30 years, vacuum technology has undergone great changes. This book is devoted to making us aware of some of these changes. Vacuum systems used for commercial analysis of materials depend heavily on the production of reliable ultra-high vacuum at a reasonable cost. As noted in Chapters 2.4 and 4.2 of this book, most ultra-high-vacuum systems are produced using stainless steel construction and copper metal seals in the form of, e.g., Conflat™ seals. The metal-to-metal seals and stainless steel construction can maintain vacuums in the 10^{-11} torr region. Many of the systems are either ion pumped or cryopumped, and in some cases where gas loads are large titanium sublimation pumps and cold panels are employed.

5.6.4.1 Sample Introduction System and Manipulation

To maintain a good ultra-high-vacuum ambient for surface analysis in an analysis vacuum chamber, manufacturers equip their apparatus with rapid-turnaround air lock systems. The detailed design varies, depending on the nature of the technique and the sample manipulation system in the analysis chamber. Most systems generally consist of some type of sample manipulation and transfer device, a fast-pumping system that will take an ancillary chamber from air pressure to near UHV conditions and a small volume ancillary chamber to reduce the pumpdown time.

Stainless steel construction is used throughout both the air lock and the analysis chamber. The ancillary chamber is typically metal sealed or may have a single O-ring-sealed port for quick sample introduction. Modern systems are universally equipped with a small-size turbomolecular pumping system, although some systems used in the academic areas still employ ion pumps and sorption pumps to evacuate the air lock chamber. Ion pumps were found to be slow and in many cases not able to handle the gas load from the samples. Turbomolecular pumps in the 50 to 200 L/s range are typically employed, depending on the size of plumbing used to rough-pump the system and where the pump is mechanically located. The mechanical pumps used to back the turbomolecular pumps use an in-line oil filter that should be changed on a periodic basis to keep the pump oil out of the

turbomolecular pump chamber (see Chapter 2.3.3). The chambers are equipped with a dry nitrogen purge system while at atmospheric pressure. This reduces the amount of water vapor that may adsorb on the stainless steel chamber walls.

Pumpdown times depend on the time the chamber is exposed to air/dry nitrogen purge and the sample(s) being introduced into the analysis chamber. Fast pumpdown times are encountered with materials that do not degas; for example, metals and inorganic semiconductor and inorganic insulating materials. An introduction system could be made to cycle from air to a good vacuum of 10^{-7} torr in 5 to 10 minutes. However, many polymeric materials tend to degas and hence require a long pumpdown before introduction into the analysis chamber. Most degassing must be done in the introduction chamber to keep the main analysis chamber clean. This is done with reproducible results in all the modern analytical manufacturers where the base pressure is 10^{-9} torr or less.

Sample introduction is typically done using a horizontal introduction system. The details of the systems vary depending on the type of sample manipulation being employed in the main analysis chamber. Some manufacturers introduce one sample per pumpdown. The sample is docked on a stage in the high-vacuum system. If more than one sample is to be introduced into the chamber, samples must be handled sequentially. Other systems handle many samples for a single pumpdown. These systems are, again, based on a horizontal introduction system that is computer controlled or hand operated.

In the analysis chamber, the sample manipulation system varies, again, with manufacturer but in all cases consists of a minimum of x-y-z-θ manipulation. Sample position can be computer controlled or manually set. For multisample analysis, it is necessary to automate the sample manipulation. This is done using computer-controlled motor drives on the manipulator where the positions of the samples can be manually read into the system. In some more sophisticated systems, sample position can be read into the computer through video selection. This type of sample manipulation is usable where positioning need not be too accurate. Repositioning a sample within a few μm is well within the capability of the equipment available.

When performing submicrometer analysis, such as in the case of small-spot AES or small-spot SIMS, manual manipulation of the sample is preferred. It is useful to be able to see the point of analysis through video imaging of the sample (through secondary electron imaging). This is obtained in both AES and SIMS instrumentation (both TOF SIMS and dynamic SIMS). x-y-z-θ manipulators are used to perform the necessary spatial maneuvers required to set up the sample for analysis. Many goniometers are capable of performing an eucentric alignment. With this alignment, the sample can be rotated without losing the analysis position.

Fixturing is available for either heating or cooling the sample during the analysis process. In some AES and XPS instruments, cooling and heating are an integral part of the sample manipulator. In addition to routine measurements, such

stages can be useful in performing elegant analysis *in situ*. Other instruments are equipped with ancillary chambers that provide sample preparation operations such as cleaning. This equipment is all UHV compatible.

5.6.4.2 In-Line Use of Analytical Systems

The semiconductor industry has led the use of highly technical instrumentation aimed at surface analysis. A large portion of the work that goes into making integrated circuits and semiconductor devices requires the preparation of clean, highly oriented, submicrometer thin-film structures on very clean surfaces that may themselves be thin-film structures. Cleanliness and chemical purity are critical. In well-equipped fabrication sites, instrumentation is included in the clean room facility to perform secondary electron microscopy and in some cases AES and other structure-related methods. Incorporating analytical instrumentation into the production of high-technology devices, materials, and structures is essential.

REFERENCES

1. K. Siegbahn et al. *ESCA, Atomic, Molecular and Solid State Structure Studied by Means of Electron Spectroscopy* (Almqvist and Wiksells, Uppsala, 1967).
2. *Handbook of X-ray and Ultraviolet Photoelectron Spectroscopy*, edited by D. Briggs (Heyden, London, 1977).
3. *Practical Surface Analysis*, Vol. 1 edited by D. Briggs and M. P. Seah (Wiley, Chichester, England, 1990).
4. J. F. Moulder, W. F. Stickle, P. E. Sobol, and K. D. Bomben, *Handbook of X-ray Photoelectron Spectroscopy* (Perkin-Elmer, Eden Prairie, MN, 1992).
5. G. Beamson and D. Briggs, *High Resolution XPS of Organic Polymers* (Wiley, Chichester, England, 1992).
6. L. E. Davis, N. C. MacDonald, P. W. Palmberg, G. E. Riach, and R. E. Weber, *Handbook of Auger Electron Spectroscopy*, 2nd ed. (Physical Electronics, Eden Prairie, MN 1976).
7. *Handbook of Auger Electron Spectroscopy* (JEOL Ltd., Tokyo, 1982).
8. *Auger Electron Spectroscopy*, Michael Thompson, et al. (Wiley-Interscience, New York, 1985).
9. *Auger Electron Spectroscopy*, edited by C. L. Briant and R. P. Messmer (Academic Press, Boston, 1988).
10. For example, see G. Hertzberg, *Atomic Spectra and Atomic Structure* (Prentice Hall, Englewood Cliffs, NJ, 1937).
11. F. J. Himpsel, F. R. McFeely, A. Taleb-Ibrahimi, and J. A. Yarmoff, in *The Physics and Chemistry of SiO_2 and the Si-SiO_2 Interface*, edited by C. R. Helms and B. E. Deal (Plenum Press, New York, 1988).
12. J. H. Scofield, *J. Electron Spectrosc. Rel. Phenom.*, **8** (1976) 129.
13. G. A. Somorjai and H. H. Farrell, *Advan. Chem. Phys.*, **20** (1970) 215.
14. J. J. Lander, *Progress in Solid State Chemistry*, Vol. 2 (Macmillan, New York, 1965).
15. G. A. Somorjai, *Principles of Surface Chemistry* (Prentice-Hall, Englewood Cliffs, NJ, 1972).

16. *Introduction to Surface Chemistry and Catalysis*, Gabor A. Somorjai (Wiley, New York, 1994).
17. *Fundamentals of Surface and Thin Film Analysis*, Leonard C. Feldman and James W. Mayer (North-Holland, New York, 1986).
18. C. W. Magee, W. L. Harrington, and R. E. Honig, *Rev. Sci. Instrum.*, **49** (1978) 477.
19. J. A. McHugh, Secondary Ion Mass Spectrometry, in *Methods of Surface Analysis*, edited by A. W. Czanderna (Elsevier, New York, 1975).
20. A. Benninghoven, F. G. Rudenauer, and H. W. Werner, *Secondary Ion Mass Spectrometry* (Wiley, New York, 1987).
21. A. E. Morgan, Secondary Ion Mass Spectrometry, in *Characterization of Semiconductor Materials*, Vol. 1, edited by G. E. McGuire (Noyes, Park Ridge, NJ, 1989).
22. *Handbook of Modern Ion Beam Materials Analysis*, edited by J. R. Tesmer, M. Nastasi, J. C. Barbour, C. J. Maggiore, and J. W. Mayer (MRS, Pittsburgh, 1995).
23. C. W. Magee and E. M. Botnick, *J. Vac. Sci. Technol.*, **19** (1981) 47.
24. J. H. Thomas, III, and B. L. Bentz, Surface analysis in *Electronic Materials Handbook, Vol. 1: Packaging* (ASM International, Metals Park, OH, 1989).
25. *Backscattering Spectrometry*, W-K. Chu, J. W. Mayer, and N.-A. Nicolet (Academic Press, San Diego, 1978).
26. W. D. Mackintosh, Rutherford scattering in *Characterization of Solid Surfaces*, edited by P. F. Kane and G. B. Larrabee (Plenum Press, New York, 1974).

PART 6

Large-Scale Vacuum-Based Processes

6.1 Roll-to-Roll Vacuum Coating 761
- 6.1.1 Overview of Roll-to-Roll Coating 761
- 6.1.2 Typical Products 764
- 6.1.3 Materials and Deposition Processes Commonly Used in Roll-to-Roll Coating 765
- 6.1.4 Vacuum Systems for Roll-to-Roll Coating Applications 775
- 6.1.5 Substrates (Webs) 779
- 6.1.6 Process Control 783
- 6.1.7 Specific Problems Exhibited by Coatings 784

References 787

6.2 The Development of Ultra-High Vacuum Technology for Particle Accelerators and Magnetic Fusion Devices 789
- 6.2.1 Introduction 789
- 6.2.2 Storage Rings and the Need for UHV 790
- 6.2.3 UHV Developments for the First Generation of Electron and Proton Storage Rings 793
- 6.2.4 Storage Ring Vacuum Vessel and Pumping System Developments 796
- 6.2.5 Cold-Bore Machines 798
- 6.2.6 Superconducting RF Accelerators 800
- 6.2.7 The Next-Generation Big Accelerator? 801
- 6.2.8 The Magnetic Fusion Road Map 801
- 6.2.9 The Early History of Magnetic Fusion 803
- 6.2.10 Model C: The First UHV Fusion Device 804
- 6.2.11 The Russian Revolution in Fusion: Tokamaks 805
- 6.2.12 Plasma Impurities and Vacuum Technologies 806
- 6.2.13 Toward the Breakeven Demonstrations 808
- 6.2.14 The Next Step in Fusion 810

Acknowledgments 810

References 812

This Page Intentionally Left Blank

CHAPTER 6.1

Roll-to-Roll Vacuum Coating

William B. Robbins
3M Corporate Research Laboratories
Thin Film Technology Resources

6.1.1 OVERVIEW OF ROLL-TO-ROLL VACUUM COATING

6.1.1.1 History

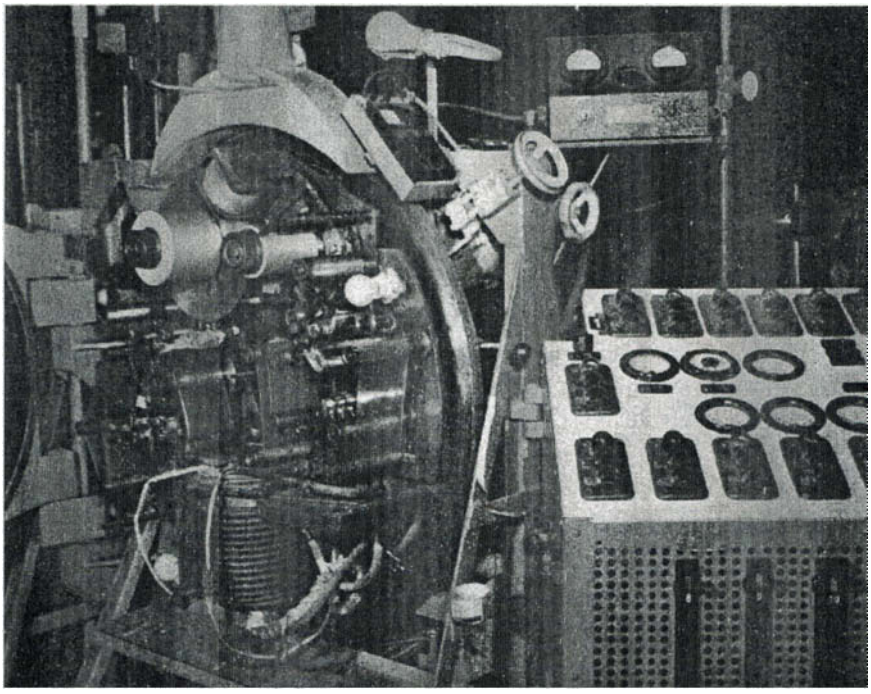
In 1917, sputtered thin-film resistors deposited on long glass rods were produced on a commercial basis. In the early 1930s, capacitor materials were being produced, made of vacuum-metallized paper tissues, using such advanced techniques as a cellulose nitrate smoothing lacquer and silver nucleation of a zinc coating. This material was produced using a roll-to-roll semicontinuous vacuum coating process; such vacuum coating is the primary focus of this chapter. The vacuum process and equipment for this capacitor film were developed in Germany. Figure 1 is a photo of this early roll-to-roll coater.

In the late 1930s Penning observed that magnetically confined plasmas enhanced sputtering rates from cathodes. This is the beginning of magnetically assisted sputtering. Also, large telescope mirrors, eventually 200 inches in diameter, were aluminized by John Strong, and aluminized automotive headlamps were in production.

From the mid-1930s to mid-1950s, many vacuum deposition processes evolved

ISBN 0-12-325065-7
$25.00

Fig. 1.



1930s roll-to-roll vacuum coater. Source: Courtesy of Collin H. Alexander, Private Communication.

into commercially viable technology in roll-to-roll systems. In the early 1950s, Q-meters were first used to measure electrical conductance of deposited films inside a vacuum system, on-line, with no contact. Monitoring having noncontact capability is important to vacuum coating technology, and sophisticated equipment has been developed to monitor properties of deposited materials and process parameters. By 1953, 3M Retroreflective Sign Material, a material used for traffic control products, was in production using vacuum deposition as a key step.

6.I.1.2 Overview

With the advent of both plastic film and paper technology, large area, roll-to-roll vacuum deposition processes became feasible. These strong, thin, flexible plastic and paper substrates are typically referred to as "webs." Roll-to-roll vacuum coating systems range from very small laboratory coaters, having webs 4 inches wide, to large units having a web width of 120 inches or 3 m. Typical commercial units coat webs that are 40 to 60 inches (1 to 1.5 m) wide; wider webs may provide

low-cost coatings if equipment is optimally operated. The process is versatile, relatively low cost, and nonpolluting.

In large modern roll-to-roll vacuum systems, a roll of web is loaded and threaded automatically; the chamber is evacuated or pumped down, and deposition is initiated. One roll may weigh 450–3500 kg and have an area of 10 thousand to 90 thousand square meters. The material is coated during one evacuation cycle, in reality a semicontinuous process but often termed continuous coating. Evacuation or pumpdown time may be as short as 5 to 10 minutes, typically 15–30 minutes, before initiating deposition, even though pressure required for much vapor coating is on the order of 3×10^{-4} to 3×10^{-6} torr. Even for sputtering processes, where pressures are on the order of 3–6 mtorr (mt), background gases must, in general, be less than a few percent of the pressure of sputtering gas, and even less for deposition of materials particularly sensitive to reaction with background gas. Occasionally, a truly continuous air-to-vacuum-to-air process is used, but this approach is rarely economical for roll-to-roll coating using flexible substrates, and will not be considered in this chapter. Such an air-vacuum-air process is generally used, and appropriately so, for coating glass.

Typical web speeds for the most commonly deposited material, evaporated aluminum, may be as high as 15 m/s (2950 ft/min). For many other materials and processes, a speed of approximately 2 m/s (400 ft/min) is common. Films deposited in various vacuum processes range in thickness from about 1 nm to about 600 nm; thicknesses of 30 to 180 nm are most common. To promote adhesion, modify nucleation and growth morphology of films, a layer is often deposited prior to the primary thin-film deposit, a practice similar to that used in batch coating. Typically, such a layer has a thickness ranging from 0.1 to 10 nm. Coatings are generally limited to a few layers of films, both for economic reasons and because of limitations imposed by properties of flexible substrates.

Although webs in roll form make these materials useful and convenient as substrates for vacuum coating, they bring with them numerous problems intrinsic to their properties. These substrate problems include outgassing, low modulus, temperature limitations, and moisture permeability. Substrate properties also contribute to poor adhesion and high corrosion rates of deposited films. These problems are significantly reduced when coating on inorganic materials, such as glass.

Two very common applications of this technology include production of packaging materials and capacitor films. Vacuum coating systems using wide webs have been optimized for producing packaging products, typically polyester coated with aluminum. Wide coating systems also provide materials serving architectural markets where wide web widths are often essential. Roll-to-roll vacuum systems for capacitor metallization have been optimized at narrower widths, 0.6 to 1 m width, as very thin substrate materials are used in this application. These capacitor coating systems require a technology to create longitudinal stripes or bands of uncoated substrate material, typically using oil masking or, in the past, continuous metal bands running on the substrate as contact shadow masks.

6.1.2

TYPICAL PRODUCTS

These are produced by a roll-to-roll vacuum coating in a "semicontinuous" process. Packaging products dwarf all other product applications combined.

Packaging products and vapor barriers
- Snack (potato chip) bags
- Permeation reduction films
- Packaging for medical products
- Transfer films

Decorative products
- "Polyester" balloons
- Wrapping papers and films

Electrical products
- Capacitor films
- Flex circuitry materials
- Microinterconnect materials
- Microwave-absorbing materials
  - Food heating
  - Stealth products
- Transparent conductors
- Specialty antennas

Optical products
- Retroreflective signs
- Mirrored films
- Solar control products
  - Reflective window treatments
  - Neutral density window films
  - Spectrally selective window films
- Reflective thermal insulation
  - Evacuated thermal insulation
- Protective fabrics
  - Reflective fire suits
- Glare reduction coatings
- Color filters
- Optical recording materials
- Infrared absorbers

Printing products
  Offset printing plates
  Security inks

Electro-optical products
  Photoreceptors
  Electrochromics
  X-ray detectors

Hard coats, friction reduction materials

6.1.3

MATERIALS AND DEPOSITION PROCESSES COMMONLY USED IN ROLL-TO-ROLL COATING

6.1.3.1 Process Concepts

Adhesion to substrate, corrosion resistance, and also reproducible optical and electrical properties are typically demanded of thin-film coatings. These properties contribute to the concept of "functional" films. The following six basic process steps help achieve functional films during vacuum coating:

1. Outgas web to avoid contamination downstream and, on occasion, to modify polymer properties comprising the web.
2. Prime the web surface, changing chemistry and mechanical properties to improve adhesion, for example, by exposing webs to plasmas of appropriate composition.
3. Deposit a thin film to nucleate the subsequent primary deposit; this first thin film, a nucleating film, helps control growth and morphology of primary film and may also act as a barrier and adhesion promoter.
4. Deposit primary film, minimizing voids and impurities.
5. Deposit protective top coatings.
6. Monitor relevant properties on-line, in a noncontact fashion.

All these steps are generally performed sequentially within a vacuum coater. This means that the web passes only once from unwind roll to windup roll, all within a vacuum coater, making for a "one-pass" process. The steps are certainly optimized to fit each product, and each step is not always necessary nor appropriate. For materials needing minimal functionality, some steps may often be rudimentary, or even optional.

6.I.3.2 Materials for Deposition

Using at least one of the vacuum deposition methods described at length later in the chapter, a list follows of a few of many materials that can be deposited on appropriate webs to achieve a wide variety of functional properties and products.

Metals: Al, Cu, Ag, Au, Zn, Cd, Pt, Pd, Rh, Sn, In, Ti, Zr, Hf, V, Cr, Fe, Co, Ni, Nb, Li, Mg, Ca, Se, Te, Bi, etc.

Alloys: Stainless steels, NiCrFe (Inconel™), NiCu (Monel™), FeCo, FeNi, NiCr (Nichrome™, Nimonic™), magnetic materials

Compounds:
Oxides: SiOx, In_2O_3, TiO_2
Nitrides: TiN
Carbides
Sulfides
Fluorides: MgF

Intermetallic compounds and *ceramic/metal mixtures*

Organics: Monomers, oligomers, pigments, dyes

The vacuum processes by which these materials can be deposited are both physical and chemical in nature. These deposition processes are thus termed physical vapor deposition and chemical vapor deposition. A few key requirements of deposition sources of either type are important if sources are to be used for roll-to-roll coating. These include the need for deposition sources to be reliable and cost effective, to have long life, and high and stable deposition rates. Physical vapor deposition is described in [1], pp. 1–27 to 1–97; the following treatment is oriented toward the specific requirements of roll-to-roll coating.

6.I.3.3 Evaporation

High-enough temperature provides adequate vapor pressure to evaporate most materials. Evaporation does not work for every material. Instead of evaporating, however, some metals react with crucibles; various inorganic compounds dissociate; constituents of alloys may selectively evaporate; some organic materials decompose, char, or cure.

GENERAL CONSIDERATIONS FOR EVAPORATION

In all evaporation processes, the following issues are dominant factors in determining coating performance and appearance: speed of deposition, materials to be deposited, substrate (web) properties, residual gas composition, and residual gas availability.

Pressure in the chamber must be low enough to allow materials having acceptable properties to be deposited. Typically, the reason given to reduce chamber pressure is to increase mean free path of evaporant, but reaction between gas and evaporant is generally a more important reason to reduce residual pressure. At very high rates of deposition, residual gas is not in equilibrium within the beam of evaporant. Both rate of evaporation and "gettering" by freshly deposited reactive metals determine effective "pressure" during deposition. Pressure is well defined only when a gas is in hydrostatic equilibrium, so pressure of residual gas within the evaporant stream is not a particularly useful concept. Measurement of pressure at a remote location in chamber is also not a reliable indicator of those film properties that result from reaction with residual gas. A more realistic determination of those properties affected by residual gas is provided by noncontact monitoring of the properties of material actually deposited on the web. Electrical and optical properties are affected by oxygen content in deposited material; residual gas also affects nucleation and growth morphology of the primary deposit.

Infrared radiation and heat of condensation each deliver significant amounts of heat to web material, heat that often must be minimized. As power to an evaporation source is increased, flux of evaporant emitted by the source increases faster than radiant emission from the evaporant source. This is due to the fact that vapor pressure increases exponentially with temperature, whereas radiant emission increases with only the fourth power of temperature. Thus, for a fixed coating thickness, more power to evaporating sources typically results in reduced web heating, wrinkling, and distortion. To control web temperature during deposition, webs may contact a cooled or chilled roll; this temperature control is essential for thin webs having low heat capacity. In addition, because substrate heating is reduced when depositing an infrared reflective material, such materials are easier to produce than infrared-absorbing materials.

Radiant heat from evaporation sources is significant; water-cooled shields and radiation shields are thus used for control of heating caused by thermal radiation. Of particular importance is the need to shield rolls of feed wire, typically supplied on plastic spools or reels. To allow heatup and cooldown of evaporation sources, water-cooled shutters are essential in thermal evaporation processes; significant quantities of web are otherwise wasted in moving the web fast enough to prevent heat distortion.

The shutters mentioned earlier provide another function, as they can remain

partially closed, providing a surface of freshly deposited metal on which to react residual gas. Although the shutters are typically closed only enough to reduce deposition rate by about 10%, deposits of higher purity and higher density typically result. Most reactions between evaporant and residual background gases are heterogeneous or surface reactions; homogeneous reactions or gas-phase reactions are rare.

Various means of heating are used for evaporation sources, including resistive, inductive, and electron beam heating, each described in sequence, with attention to its performance when used as a source in roll-to-roll processing.

1. Resistive Heating

Resistive heating using composite intermetallic bars as the evaporation source is used for approximately 90% of roll-to-roll coating. Most of this is evaporation of aluminum in the form of wire, 1199 alloy, 99.99% pure, used in wire feed systems having independent speed control for each wire. Heater bars must be electrically conductive; bars are typically composites of TiB_2, BN, and AlN tailored to minimize erosion. Heater bars are spaced approximately 0.15 m (6 inches) apart, across the width of the web, with outermost bars centered under the edge of the web. For very-high-speed coating of aluminum, two rows of bars are used, the second row positioned down-web of the first row, and typically staggered with respect to the first.

2. Inductive Heating

Inductive heating is used for about 7% of roll-to-roll coating, used for depositing aluminum or similar metals such as copper or silver. Semiconductor switching power supplies, or in the past, motor-generator sets, provide power in the frequency range of 10–100 KHz, inductively heating susceptors, crucible material, or evaporant metal directly.

Spitting of small metal balls from boats or crucibles is observed frequently during inductive heating. Cross web nonuniformity of coating is also a concern when using such heating methods. Spitting and nonuniformity across web can be reduced significantly by careful boat design, selection of boat material, and inductor design. Depending on the evaporant material, crucible life may be a problem that cannot adequately be resolved.

3. Electron Beam Heating

Electron beam heating accounts for 2% of roll-to-roll coating; typical beam voltages range from 10 to 40 kV. Typically, arcs cause instability in deposition rates;

arcs result when pressure is too high or when feedthroughs and insulators become contaminated, especially by magnetic materials. Arcs may be minimized by using electron beam guns that are differentially pumped, reducing pressure inside guns and contamination of guns. An indirectly heated electron emitter improves filament life. Power supply reliability is also a recurrent difficulty for many electron beam systems.

Rods fed through the bottom of the crucible, or wire fed into the melted pool of material permit much longer deposition runs than would be permitted by inventory of materials in boat or crucible at the beginning of run. Such techniques are particularly important when trying to control composition of evaporated alloy, by helping to stabilize level of melt.

Control of alloy composition is easily realized when an alloy evaporates congruently. Congruent evaporation is analogous to an azeotrope, providing a fixed composition on evaporation; NiFe (15%) (Permalloy) is a good example, as is CrGe. Materials that closely approximate the composition of a congruent evaporant can be deposited by focusing an e-beam to a small spot on a stable level of melt. This small spot is highly heated, and diffusion in the melt limits the availability of the higher-vapor-pressure component. Use of rod feed or wire feed as described may provide requisite melt level control. An example of material deposited using this procedure is NiCr (17%) Nichrome™.

CRUCIBLE AND BOAT MATERIALS

The materials from which crucibles or liners are made often significantly limit the evaporation process, as some molten metals attack almost any material. Molten aluminum exhibits this problem, dissolving most materials that it contacts. Materials have been developed for crucibles and bars for wire feed heating that are quite adequate for evaporation of aluminum. Given the time and expense to replace wire feed bars or crucibles, replacement becomes largely an economic issue; operational performance isn't severely affected. A few evaporant materials have no appropriate crucible or bar material to meet demands for long life, reliability, low cost, and minimal contamination typically demanded by roll-to-roll coating. Evaporation in small laboratory operations is not an adequate indicator of bar or crucible performance in typical industrial roll-to-roll operations.

Some of the limitations just discussed that are produced by available crucible materials are relaxed in the e-beam process for the following three reasons:

1. Typically, crucibles sit in a copper hearth that is water cooled. The temperature at the interface between melt and crucible is significantly cooler than the temperature of that same interface in a resistively heated system. In a resistively heated system, all necessary power is provided through that interface, whereas

in an e-beam system, power is delivered through the top molten surface. Lowering the interface temperature by only 100–200 K may provide more thermodynamic stability, reducing diffusion rates, reaction rates, thermal stresses, spalling, and erosion, consequently reducing reaction compounds formed at the interface. Spalling and erosion are significant failure mechanisms for resistance heater bars. Crucible and boat materials exhibit better stability when used in e-beam and inductively heated applications and less stability when used in resistively heated systems.

2. A significant failure mode of crucibles is related to thermal expansion and contraction differences between crucibles and evaporant on heating to the melting point or cooling below the melting point of the evaporant. Volume changes at the melting point also produce stresses. Crucibles and boats occasionally crack near this temperature. If evaporant source material does not wet boat or crucible, these stresses are reduced, but in resistive evaporation, heat is poorly transferred at such a nonwetting interface. This forces a selection of boat or bar materials that react and wet, at least marginally. This is not a requirement for e-beam crucibles.

3. Crucibles for e-beam use need be only slightly electrically conductive at operating temperatures; whereas in resistively heated applications, crucibles, boats, and bars need to be quite conductive at temperature of use, restricting choice of material.

One method to reduce contamination of melts by crucible material is through the use of unmelted shells of evaporant material surrounding a pool of molten evaporant, called a "skull." To maintain the severe temperature gradient usually implicit in such a structure requires very high fluence of e-beam power. An extreme case of this is to use no crucible, directly melting evaporant in hearth. At the very-high-power densities needed for these approaches, hydrodynamic instabilities become evident in the formation of surface waves and melt stirring, both of which may become intolerable.

Thermal insulation of a flexible felt or woven fabric placed between crucible and hearth, for example, zirconia felt, will reduce heat loss from crucible to hearth. This permits higher evaporation rates or lower e-beam power, quite the opposite of skull formation, but demanding greater stability of crucible material. Used with thermal insulation, the crucible material may then be selected with little need to minimize thermal conductivity.

6.1.3.4 Sputtering

A basic sputtering process requires an electrically conductive target maintained at a negative electrical potential, surrounded by gas from which ions can be drawn

to the target, dislodging target atoms. See Chapter 5.1 for additional detail. Rate of deposition is inordinately slow if diode sputtering is used. The primary technique to enhance sputtering rate is to magnetically confine or trap energetic secondary electrons emitted by ion collisions with target. These electrons are magnetically trapped near the target, generating a copious flux of ions that sputter the target; thus extremely high sputtering rates are achieved. Electron traps for energetic electrons are created with magnetic fields of 300 to 3000 gauss (0.03 to 0.3 tesla) crossed with the electric fields generated by target. The target is quite cool during sputtering; evaporation plays no role. Deposition rate achievable using this process becomes primarily limited by effectiveness of target cooling.

Alternatively, to increase sputtering rates either of the following procedures can be used: (1) an enhanced supply of electrons such as created in triode sputtering, or (2) copiously supplied ions, attempted with limited success in ion beam deposition, or realized with microwave-enhanced sputtering. Microwave-enhanced sputtering is very useful for sputtering magnetic materials.

Power supplies for sputtering must be carefully selected if sputtering is to be successful in roll-to-roll applications. Reliability and ability to suppress arcs are both essential in power supplies useful for roll-to-roll sputtering, and both become harder to achieve as system power approaches one megawatt (MW). Response time of arc suppression is a crucial factor in high-speed roll-to-roll coating, as a power supply shutdown of 100 milliseconds may produce an uncoated section of web on the order of 0.5 ft.

In sputtered films, stress of deposited film can be minimized by adjusting sputtering pressure to an optimum value, typically in the range of 3 to 6 mT depending on atomic weight of sputtered species [2].

1. HIGH-RATE SPUTTERING

Sputtering is often viewed as a slow process. It is fast if it can be operated at high total power. It is also compact in the downweb direction if it can be operated at high power density; this requires low target and plasma impedance and adequate target cooling. For many product applications, deposition of a wide variety of materials is possible at relatively high web speed, 200 to 400 ft/min. Material is removed by sputtering from targets at a rate of 0.5 to 3 μg/J or 1.8–11 g/kWhr, depending primarily on target material and sputtering gas composition. For web coating, high rate sputtering requires power of 0.1–1.0 MW, with target design providing effective water cooling, especially at high power densities. A 60-inch-wide system, having sputtering targets operating at a total power of 1 MW, removes, for example, 1000 g/hour of copper from targets, producing up to 20 nm coating thickness at 200 ft/min. Many target materials may be operated at power densities of 200 W/cm^2, given effective target cooling. At such high power den-

sities, the reservoir of material stored in the target is, however, generally inadequate, so a power density of 50 W/cm^2 is often a useful compromise between target life and a compact sputtering installation. Certain target designs and magnetic field configurations permit more effective target cooling and lower target impedance than other designs. These characteristics have been achieved with tubular sputtering [2], pp 48–50 and 107–109; [3].

2. REACTIVE SPUTTERING

The high deposition rate and very-high-power densities achievable with this sputtering equipment makes possible controlled reactive sputtering, the final topic on sputtering.

Advantages of reactive sputtering are that many compounds can be formed using relatively easy-to-fabricate low-cost metallic targets; deposition rates of these materials may be high; low-cost targets and the high deposition rates result in low-cost products. The difficulty in reactive sputtering is the perceived complexity of process control that accompanies this versatility. Equipment design and operating technology eliminate most problems that reactive sputtering is often perceived to have.

Reaction films on cathode (target) surface may decrease sputtering rate. One must balance flux of reactive gas species with physical vapor flux to achieve desired composition of deposited film. This is not readily achieved through measurement and control of partial pressure. If the sputtering process is well designed for stability, control of gas flow is best accomplished by mass flow controllers; bistability is avoided by proper equipment design and operation, not by control of partial gas pressures.

First, though, a description of the common wisdom suggests that there are perceived complexity and control difficulties. When sputtering with mixtures of reactive-gas and argon, the relationship between film properties and the reactive gas injection rate is generally observed to be very nonlinear. As the reactive gas injection rate approaches that required to produce a stoichiometric coating, an increase in the partial pressure of reactive gas in the sputtering system results from a reduced getter pumping rate of the deposited coating. The result is that the sputtering discharge undergoes a sharp transition into a mode in which the metal sputtering rate, and therefore the reactive compound deposition rate is dramatically reduced.

A large decrease in discharge voltage and relative deposition rate during a typical transition is commonly observed. The voltage decrease at high oxygen injection rates is a transition from a deposition of metal to deposition of a compound. The voltage increase at low oxygen injection rates is the compound-to-metal transition. The reduction in sputtering rate is believed to result primarily from compound formation on the target surface. For discharges driven at a constant current,

the reduction in discharge voltage is the result of higher emissions of secondary electrons typical of compounds.

This hysteresis or bistable behavior exists for both discharge voltage and for deposition rate. The process will remain on the lower curve until the working gas is made sufficiently lean in the reactive species so that removal by sputtering of the compound layer on the target can occur. A target on which such a layer has formed is often referred to as "poisoned." Very pronounced poisoning effects may occur during reactive sputtering of materials with oxygen that form tenacious, nonconductive oxides such as Al, Cr, Ti, and Ta. No poisoning occurs for Au, where the sputtering rate for pure oxygen is not much different from that for argon.

The poisoning effect introduces practical problems: The sudden decrease in deposition rate means thickness control is lost. Precise control of stoichiometry is nonexistent, a very important problem for transparent conductive indium tin oxide. To recover and return to the proper operating point, flow of reactant gas must be decreased to the lower flow in the hysteretic curve. Time passes until operation is again stabilized at this desired operating point; meanwhile, only scrap web has been generated. Limited success has occurred in operating right at the transition point [4], pp. 231–235.

Properly designed equipment and process technology does not behave this way. Most discussions of these hysteretic transitions focus on sputtering processes at the cathode, ignoring the rest of the system. The total system includes target design, system geometry, deposition and gettering on surrounding walls, positions of gas injectors, and total gas flow, including nonreactive argon flow, which is very dependent on pumping capability of the vacuum system.

Most sputtering models are rate independent. They assume that pressure of reactant gas is homogeneous throughout the system, so they predict target poisoning. In reality, high sputtering rates generate spatial gradients in concentration of reactant gas; concentrations of reactant gases are not in hydrostatic equilibrium. High-power densities keep targets quite clean and free of reactants. High argon gas flows are used so that the concentration of reactant gas changes little as gettering rate changes on approach to saturation; the concentration depends primarily on system gas flows as set by the operator.

The geometry of tubular sputter targets produces a radial spread of metal flux; reactant pressure needed at the substrate for full reaction is thus significantly less than reactant pressure required at the target to produce poisoning. Magnetic fields are high, 0.1 T, and configured to produce efficient traps, so the system is not rate-limited by secondary electrons as are typical planar targets. The voltage is low even with no reactant gas, so it never exhibits a drop with additional reactant gas. The target operates stably over a wide range of reactant flows, exhibiting no transitions and certainly no bistable behavior. This makes for a very stable and controllable process.

Gas mass flow meters for sputtering should be selected that have no zero drift and good gain stability. Response speed, permitting rapid turn-on, important for batch semiconductor processing, is not important for roll-to-roll deposition. A slow response (30 sec) is acceptable, and this helps minimize overshoot on turn-on.

Comments follow on selection of materials and alloys for sputtering:

- Materials used for roll-to-roll coating need not, in general, be as pure as those used for semiconductor applications. Often only a few specific impurities are of concern and need to be reduced. Generally, gas content of targets needs to be minimized, especially oxygen, nitrogen, or hydrogen; oxygen in copper is an example. Oxide inclusions and voids in metal and alloy targets contribute to arcs, and thus need to be minimized.
- Generally, a commercially available alloy is used, options of composition thus become limited.
- Material must have low permeability as received, typically in an annealed state, otherwise the magnetic field used to enhance sputtering will be shorted by the target. "Ferromagnetism" [5] is useful for determining such magnetic properties.
- After sputtering, the material should also be nonmagnetic. If it becomes magnetic, for example, through a change of crystalline structure, then chips and shards of sputtered material cling to targets, causing arcing. This is a problem observed on sputtering some types of stainless steel. Target fabrication can represent a significant technical challenge for certain materials and target designs, especially for targets used at high-power densities.

6.I.3.5 Cathodic Arc Deposition

Cathodic arc sources convert solid cathode material to plasma, vapor, and molten droplets. The plasma has a higher concentration of ions than is typical in other physical vapor deposition processes; these ions are often in multiple charge states and have high kinetic energy. These ions permit external control of film morphology and coating on non-line-of-sight surfaces. Easily achieved with this technology are deposition of refractory metals and reactive deposition of refractory compounds, producing materials that are hard and wear resistant. Molten droplets and microdroplets emitted from these sources can be a serious limitation for many applications; much effort has been expended in reducing this problem. Useful cathode materials must be electrically conducting and fully dense; operating with a melted target reduces microdroplet emission. This technology has been applied primarily in batch coating of tools and equipment, but has some utility in roll-to-roll coating.

6.1.3.6 Chemical Vapor Deposition

Low-pressure CVD in various forms has occasionally been used for roll-to-roll coating. This process tends to require high substrate temperatures, which can be reduced by enhancing reaction conditions using plasma, laser, or microwave energy. In addition, to improve adhesion and to change physical and chemical properties of surfaces, plasma modification may use materials supplied as gas, which are excited and react or interact with web.

6.1.3.7 Deposition of Organic Materials

Huge numbers of organic materials may be thermally deposited and used in roll-to-roll products. Monomers and oligomers often evaporate easily, and polymeric source material may decompose, depositing short-molecular-weight material. After deposition, curing of materials with functional groups is appropriate, using exposure to low-energy beams of electrons or ions. Diparaxylelene from Union Carbide (vaporized, cracked, polymerized) deposits an organic film with no explicit curing step. Thin-film deposits of organic materials may provide protective coatings, lubricant layers, hydrophobic films, and other functional properties in addition to the properties provided by the primary deposition.

6.1.4

VACUUM SYSTEMS FOR ROLL-TO-ROLL COATING APPLICATIONS

6.1.4.1 Vacuum Pumps

Roll-to-roll vacuum systems are not very clean, due to the introduction of copious quantities of water, air, and organic materials emanating from the large rolls of webs coated in these vacuum systems. The surface area of web being coated may be 100 to 1000 times the internal surface area of the vacuum system itself, thus the web may dominate outgassing during coating. Typically, rolls of polyester web contain 0.5 to 1% by weight of gas and volatiles. These various factors largely determine pump selection, operation, and performance. They certainly contrast with practices and concerns in other areas of vacuum technology. Pumping concerns also relate to chamber "pumpdown time," outgassing of substrate and chamber during coating, gas loads from evaporating, sputtering and reactive sputtering.

Refer to Section 4.7.4 for a discussion of outgassing from polymeric materials,

with the additional consideration that little outgassing occurs from the roll of web until it is unwound during coating. In addition, small shards of polymeric material generated by previous operations, especially slitting operations and cinching of roll on windup, may contaminate system and pumps, unless substrate cleaning procedures have been used to remove dust and particles prior to coating.

The pumps commonly used in roll-to-roll systems are listed here, and are selected to meet the specific needs of these systems:

Backing, roughing
- Mechanical Compressor Pumps
  - Rotary piston
  - Rotary vane
  - Axial screw

Medium vacuum
- Roots blowers (Chapter 2.3)
- Vapor booster pumps

High vacuum
- Diffusion pumps (Chapter 2.4)
- Turbomolecular pumps (often with molecular drag sections) (Chapter 2.6)
- Capture "pumps" or traps
  - Cryogenic (helium refrigerator) (Chapter 2.5)
  - Meissner traps (liquid nitrogen chilled)
  - Refrigerated traps (CFC refrigerator)

In addition, reaction between residual gas and reactive metal vapors such as aluminum or titanium effectively pumps residual gas. This pumping or "gettering" happens as a by-product of the deposition process, often affecting both product and process quite dramatically. Commercially manufactured getter and sputter pumps are not used in roll-to-roll applications.

Diffusion pumps have been the high-vacuum pumps of choice for roll-to-roll coating. Originally, this choice was due to availability; later, low cost and reliability became primary reasons. Reliability of diffusion pumps is somewhat overrated, as they need more maintenance, cleaning, and care than they generally receive. They also need water-cooled baffles to minimize backstreaming; further, they do not provide stable pumping performance if operated at gas loads near their rated throughput.

Cryotraps in various forms are very useful pumps for roll-to-roll processes. Cryopumps using helium refrigeration are best used alone or with turbo pumps; cryopumps are generally not used in combination with diffusion pumps. Meissner

traps and CFC refrigerated traps are used especially for water trapping, particularly in conjunction with diffusion pumps.

Reliable pump performance is essential, even when pumps are assaulted by high gas loads and condensable gas, especially water vapor. Oil-recirculating systems, gas purges, and ballast valves help keep oil cool and dry in these systems, extending both mechanical pump and oil life. Large pumps, with relatively slow rotational speed, typically operate at lower temperature, wear less, and have longer seal life than smaller units with similar capacity achieved by using high rotational speed. Tolerance to water vapor and ability to pump water vapor are key specifications for mechanical pumps used in web coating operations, especially those used during regeneration of traps or those which pump chambers in which rolls of web are unwound.

Throughput or mass flow of gas is related to the power input and efficiency of a pump. Throughput is often more important than pump speed for much roll-to-roll coating, as pumps are often pushed to mass flow limits, especially by sputtering processes. All these pumps are pressure and power limited. If driven beyond rated capacity, mechanical pumps lose oil, may overheat, and seize; jets of diffusion pumps stall, and chamber pressure rises sharply; capture traps may warm, expelling previously captured gas.

The supply roll of web should, in general, be isolated from coating chambers by conductance-limiting slots or rolls. These design features help decrease the gas from substrate outgassing that flows to coating processes. Isolation between various sections of a vacuum chamber with various functions are provided by conductance limits between sections, with separate pumps for each section [6]. This design approach reduces cross-contamination between sequential processes, often operating at various gas compositions and pressures. Commonly, supply roll chambers are pumped by large Roots blowers backed by small Roots blowers, which in turn are backed by rotary piston pumps. This is a reliable package if properly operated to minimize water contamination, corrosion, and oil degradation.

Pressures less than 20 torr are tolerated by vapor booster pumps. Roots blowers are best selected with bypass valves permitting high input pressures. Roots blowers without bypass valves may be started at pressures between 10 and 20 torr, whereas Roots blowers with bypass valves may be operated at atmospheric pressure and contribute to pumpdown performance at pressures significantly above 20 torr. Under certain conditions of pressure and throughput, Roots blowers operate with small compression ratios, limiting performance.

Systems are roughed to between 0.1 and 1 torr, depending on types of pumps used, system volume, and valve-opening speeds. At lower roughing pressures, the potential for backstreaming of mechanical pump oil must be reduced by using filters, traps, and Roots blowers. To minimize backstreaming, diffusion pumps must not be exposed to high pressures during crossover.

6.1.4.2 Pressure Gauging (Chapter 3.1)

Pressure gauges useful for roll-to-roll coating processes are tightly constrained by the following considerations:

- Stability of zero on longer runs, especially as chamber temperature increases
- Ability to set zero accurately, even if residual background pressure is high, a common situation
- Response to contaminating materials
- Reproducibility on gauge tube replacement
- Reliability of gauge tube and electronics
- Response to gas mixtures, and whether or not gas factor can be ascertained

THERMAL CONDUCTION (CONVECTION) GAUGES

The wide range and reliability of these gauges make them useful for roughing control, backing pressure limits and operating interlocks, including safety interlocks on power supplies and interlocks controlling ion gauge tubes. Stability of zero is quite adequate, and can be set using sealed evacuated tubes, useful when gauge is used in locations where pressures do not get very low. Contamination of gauge tubes with pump oil may be a difficulty, but coarse sintered metal filters interposed between system and gauge tube alleviate this problem.

ION GAUGES

At the higher pressures, 10^{-4} torr, typical in roll-to-roll coating and especially in sputtering systems, triode tubes (type 504) or Granville-Phillips Stabil-Ion™ gauges are particularly useful. A standard Bayard-Alpert (BA) type of ionization gauge is ineffective above pressures of 2×10^{-4} torr, though wide range BA gauge tubes may be of some value. Some gauge controllers achieve linear performance even with BA tubes at higher pressures, by using short excitation times. Electrometer zero is automatically set in ion gauge systems even at high pressures, and zero stability is excellent. Reliability of electronics for driving ionization tubes is improving.

CAPACITANCE GAUGES

Zero stability is inadequate on older designs, but is becoming better on certain new equipment. If the system does not achieve pressures that are low enough, gauge zero can not be accurately set. Certainly, a valuable feature is the measurement of pressures, which are independent of gas composition and gas properties.

6.1.5

SUBSTRATES (WEBS)

6.1.5.1 Web Materials

Polymers (see expanded discussion at end of list)
- Polyester (Polyethylenceterepthalate) (PET) is the workhorse of this industry and is readily available in thicknesses from 1.5 to 175 μm, 0.00006–0.007 inch.
- Polyimide (Kapton™) is used for flexible circuits. This material tolerates very high temperature, so products made using it are solderable.
- Polyamide (Nylon™)
- Polymethylmethacrylate (Plexiglas™, Lucite™)
- Polycarbonate (Lexan™)
- Polyethylenenapthalate
- Polyethylene
- Polypropylene
- Polytetrafluoroethylene (Teflon™)
- Polystyrene
- Cellulose acetate
- Polyvinylidene Chloride (Saran™)

Metal foils

Paper products
- Kraft and standard papers

Textiles
- Woven and felts or nonwovens of natural fibers and synthetic fibers, including glass fibers

In addition to the wide variety of polymers just listed, there are copolymers, polymer blends, block copolymers, coextrusions, layered, and laminated materials. Many processing variables determine internal structure and therefore properties and performance of each polymer system.

As an example, polyester (PET) can have a range of thermal expansion, hygroscopic expansion, shrinkage on heating, modulus, and level of trimers. The value of each of these properties depends on the amount of stretch during orientation, on temperature and time of heat setting, all of which result in different amounts of crystallinity. Stretch is a factor that is typically different in the longitudinal direction of the web (machine direction) from the cross-web direction (transverse direction), creating properties that are anisotropic.

Various limitations of polymer substrates are notable. Low elastic (Young's) modulus is typical, especially in comparison with the modulus of glass and metals. A modulus of 0.5 to 0.8 $\times$ 10^6 psi is typical for PET, one tenth the value of glass and metal. In addition, polymers degrade with heat, they hydrolyze and are permeable to water vapor, often supplying acids or bases to interfaces between polymer and thin films. Plasticizers in some polymers migrate; vinyl (polyvinyl chloride) is a good example. Some of these plasticizers easily diffuse to the surface of the web, causing two problems: (1) increased pressure in the vacuum chamber during deposition as plasticizers evaporate and decompose, and (2) plasticizers form greasy layers on the polymer surface, resulting in bonding difficulties for vacuum deposited films.

All these limitations of polymers tend to speed corrosion and decrease adhesion, resulting in possible product failure. Many polymers also exhibit birefringence, scattering and surface roughness, degrading optical performance, and compromising polarization.

6.I.5.2 Web Handling

The primary goal is to unroll the web to be coated, move it past deposition sources, including, if necessary, such process steps as outgassing, plasma exposure, nucleation, and monitoring of properties; then roll it up again. Difficulties may occur, with wrinkles and creases forming, rolls slipping sideways (telescoping), or wound too tightly, edges building up faster than the center of the roll, webs bagging and sagging in the center of the web, scratches generated, and point defects created.

Before addressing cures for these problems, consider a typical web drive system. Such a system has at least three motors, each motor having four quadrant control. Two of these motors provide web tension, one each for the wind and unwind rolls, the third motor drives a roll at user-selected speeds. Motors deliver torque, which is adjusted by feedback to control web tension. Web tension is determined using either direct measurement, or a system estimating roll diameters and friction to calculate web tension from motor current. Mechanically, many web drive problems can be addressed by using appropriate motors and drive systems that are independently adjustable along web path. Various methods are used for isolating sections to permit separate wind and unwind tensions. These include either a pull roll or "pacer" roll, with a "nip" roll sometimes used to hold web against pacer roll to prevent web slippage. If web that hasn't been outgassed is used against a high-temperature pacer roll, nip rolls are essential to avoid web slipping caused by gas lubrication. To avoid edge buildup, windup rolls are often moved axially back and forth a few millimeters. This motion is slow, hydraulically actuated, and referred to as "side lay."

Drive motors and tension motors are usually mounted outside the vacuum chamber. Rotary feed-throughs, sealed with either O-rings or ferrofluids, permit transmitting mechanical power from motors into vacuum chambers. O-ring seals are often quite acceptable in roll-to-roll coating applications, as pressures are not extremely low, and constant rotary motion quickly generates a differential pressure drop across a multiplicity of O-rings, resulting primarily in startup "transients," as differential pressures are established. Drive motors are coupled to pacer rolls through high gear ratios and tension motors are coupled through low gear ratios, matching the mechanical impedance of tension motor to web roll. Flexible couplings, for example, Lovejoy Uniflex™ brand couplings, are essential on each side of ferrofluidic feedthroughs to decrease radial loading on such feedthoughs.

For polymer webs, tension typically used is such that web stress is in the range of 80–600 lb/in^2, depending on web modulus. Thin polymer webs, 4–20 μm and thinner, require controlled acceleration and tendency drives to reduce wrinkling. Tendency drives provide torque to idle rolls, through bearings, by rotating the shafts on which idle rolls are mounted, requiring four bearings for a shaft and roller set. For very thin or low modulus web materials, limited or controlled acceleration rates and "tendency drives" are both useful to reduce the effects of bearing friction and roll inertia. In comparison to polymeric materials, paper webs need higher tensile stress, and metal foils may require much higher stress, especially if contact needs to be maintained between web and a chill roll.

Alignment of rolls, or parallelism between the axis of successive rollers, often described as "tramming," is critical to minimize wrinkling and creasing of webs. In addition, web wrinkles are a function of web caliper (thickness), web width, web tension, web modulus, wrap angles, velocity, span lengths, roller alignment or tram as already mentioned, roller diameters, roller runout, and friction between web and rollers.

Arrangement or layout of rollers within the vacuum system plays a significant role in the sensitivity of web wrinkling to roller alignment. This arrangement generates the degree of wrap of web around each roller, and the free-standing span between each pair of rollers, both of which determine the precision of required roller alignment. Defining a span or length (L) of web between any pair of rollers, and normalizing for a given web width (W), generates an L/W ratio. If the L/W ratio is increased, roller alignment need not be so precise, but increasing this ratio is difficult within the confined space available inside typical vacuum coaters. Increasing web caliper increases web stiffness, and has the same effect as increasing L/W ratio, requiring less precise alignment. The web material or coating may also cling to rollers, demanding more precise alignment. A system may be aligned or trammed for web moving in one direction, and not be trammed for web moving in the reverse direction. This is due to the need for web to approach each roller perpendicular to the rotational axis of each roller. Wrinkles arise not only from roller misalignment and rollers that are not precisely cylindrical, but also from poor

web quality, poor stock roll formation, and from nonuniform shrinkage and expansion of webs, both caused by web heating during film deposition.

A slight collapsing of any vacuum chamber occurs on pumpdown, even in the stiffest of chambers; this is an elastic flexing of chamber walls. This flexure or elastic distortion of vacuum chamber walls can easily destroy roller alignment if not prevented by proper design. Web-handling systems inside a vacuum system are often mounted on plates quite well isolated from walls of the vacuum system, and flexible couplings are used on all mechanical feedthroughs. These roll-mounting plates are best made of steel or hard aluminum, as stainless steel is not dimensionally stable enough. A web drive system mechanically coupled too rigidly to a chamber that is too flexible cannot be aligned or trammed simultaneously at both atmospheric pressure and at vacuum, a very bothersome problem, and totally inexcusable.

Adhesion of freshly deposited films to rollers handling the substrate, especially if film is oxide free, can cause wrinkles, particularly with web-handling equipment that is not well aligned or trammed. Therefore, the front surface of a freshly coated film should not be contacted by rollers, if possible. This is particularly important with evaporated metals, as there may be inadequate film-to-substrate adhesion. Lifting of small specks of vacuum-deposited metal from the coated web, and transferring this metal to rollers is a problem occasionally observed. This problem, termed "picking," is solved by (1) improving adhesion between metal film and web through use of plasma priming and barrier layers, (2) rethreading web in coater to avoid front surface contact, or (3) oxidizing the surface of the metal film before it contacts a roller.

6.1.5.3 Web Temperature Control and Web Defects

WEB TEMPERATURE CONTROL

Control of web temperature during deposition is an important aspect of controlling product quality, and is typically achieved using chilled or heated rolls to stabilize web temperature and provide web dimensional stability. Films may be deposited onto webs that are either in free span or held against a temperature-controlled roll. A free span is particularly useful if web cannot be held uniformly against such a roll, as is often the case with metal foil webs. Infrared lamps and ultraviolet lamps aid outgassing from polymer webs.

Heat transfer coefficient in vacuum is typically 140–280 W/m^2K (25–50 Btu/hr-°F-ft^2) between web and chill roll for PET and polyimide. This is very sensitive to the amount of outgassing from web, the pressure of trapped gas from this outgassing, the surface finish of roll and polymer surface finish and polymer compliance [7–12].

A convenient formula relating web length to roll diameters and web thickness, using the same dimensional units throughout, is

$$L = \frac{\pi(Do^2 - Di^2)}{4t} \tag{1}$$

WEB DEFECTS

Various defects are often observed in polymer webs and rolls of web. A major problem of rolls is air that is entrapped into stock rolls prior to vacuum coating, caused by inadequate web tension, high web speed and method of winding. This entrapped air may cause telescoping because it expands in vacuum and lubricates the roll, permitting the roll to slip sideways or telescope in a characteristic stepped fashion. This air also affects deposition if too much of it flows into the region of deposition. Web defects include hard bands, which represent thicker material; baggy web, which is longer in some regions than others; wavy edge-slitting; point defects, which exist in webs, and are also generated or enhanced during coating by heat, dirt on rolls, slip pimples, and dirt on webs.

6.1.6 PROCESS CONTROL

6.1.6.1 Monitoring and Control During Deposition

To prevent scratching either the web or the deposited film, when measuring properties of thin films on moving webs, noncontact measurements are essential, especially in vacuum. On-line, noncontact monitoring of electrical conductance, optical transmission and reflection, magnetic properties, and substrate temperature are readily accomplished, even in vacuum. Significant effort may be required, however, to achieve these capabilities. Delcom Instruments, Inc. produces conductance monitors that achieve requisite performance. Monitoring equipment must have excellent gain and zero stability to be useful for roll-to-roll coating, as zero cannot be reset in the middle of runs. By combining data from the preceding techniques, along with off-line calibrations using XPS, ICP, ellipsometry and other analytical methods, additional material properties may be measured on-line, in a noncontact fashion.

A generally useful concept that is not often applied is the elimination of some unknowns shared by different measurements to determine additional parameters. As an example, measurements of both electrical conductance and optical transmission allow determination of volume resistivity and film thickness on-line

during deposition; for example, in an Al/AlOx cermet used for microwave heating films. Again using measurement of both electrical conductance and optical density of a copper film, the oxygen content can also be derived from these measurements, on-line, during deposition. Too much oxygen in copper makes it corrode more easily in some environments. As previously mentioned, both examples require off-line calibration of properly designed experimental runs followed by appropriate data reduction. Once appropriate data have been collected, optical models and effective media approximations permit understanding of deposited products.

6.1.6.2 Measurement of Thin Films after Deposition

Measurement of thin films on flexible plastic substrates presents a few problems not encountered with flat, rigid glass substrates. Surface flatness and back surface reflections can be overcome, permitting optical and ellipsometric measurements to be made; electrical properties can be measured using a variety of standard and nonstandard techniques.

Continuous coating processes provide a few advantages: One advantage, for example, is that the first few feet of material may be discarded, to eliminate material produced during "transients" occurring at startup of coating processes. Quite a different situation exists for deposition onto semiconductors, where such transient conditions are too often easily incorporated at the beginning of the deposit. Startup transients are caused by a variety of events, both in batch and continuous coating processes. Some of these include target cleanup at the beginning of sputtering processes, O-ring leakage through rotary feedthroughs upon starting rotation, adjustment to specifications and source heatup.

Another advantage of continuous coating is that, for some parameters, monitoring properties is often easier on moving substrates than on fixed ones; the moving substrate proceeds out of the deposition zone. At this point measurement of optical, electrical and magnetic properties of the deposited film are more easily made, as are measurements of substrate temperatures, all using noncontact methods. Good process monitoring and control allow successful production of functional films that withstand all types of environmental attacks and assaults.

6.1.7 SPECIFIC PROBLEMS EXHIBITED BY COATINGS

Functional vacuum-deposited coatings must adhere to the substrate and not corrode nor easily abrade. Consider some environmental causes and mechanisms that delaminate and corrode films.

Here is a list of environmental assaults to which functional vacuum-deposited films may well be exposed during use or subsequent manufacturing steps. Equally, it is also a list of tests to which films might be deliberately exposed. Some of these environmental assaults include exposure to higher temperatures, acids, bases, water, electrical fields and also mechanical wear and flexure. Specifically, temperatures of 325°C are demanded of polyimide-based materials; acid attack must be resisted by many conductors; bases, either KOH or bleach, must be tolerated; and boiling water is often used as a quick test for water immersion. Electrical fields and current flow may cause ion and vacancy migration, cause thin films to heat, and encourage various electrochemical reactions.

These accelerated tests, or what might be called *hypertests*, can provide useful information regarding the nature of mechanisms that may degrade adhesion. These attacks on a polymer–metal interface often have an electrochemical component. Although peel and adhesion tests are difficult to interpret and often provide limited and contradictory information, they become easier to interpret if performed after exposing products to such hypertests, especially if the products fall apart. Before considering approaches to improving adhesion, consider corrosion and abrasion resistance.

6.1.7.1 Corrosion Resistance

The problem with corrosion of thin films is simply stated: A corrosion rate averaging 25 μm/year (0.001 inch/year) is generally acceptable for most materials applications as in chemical engineering, automobile bodies, and even bridges. At this rate, however, a film with a thickness of 100 nm has totally corroded in a day or two. Often, functional thin films must maintain properties for five or more years.

Here are some techniques to decrease corrosion rates:

- Reduce impurities in the primary deposit, especially oxides, by depositing films at high rate, and at low background pressure, with partially closed shutters.
- Reduce voids in deposited material both by nucleating to avoid initial colloidal deposits, and by minimizing oxygen content, as noted earlier, with high deposition rates and low pressures.
- Select those alloys that may form adherent, dense, impermeable oxide films on their surface. Stainless steel, Inconel™, and Monel™ are examples.
- Add top coats: Deposit a very thin, dense, adherent and insoluble oxide or a tightly bound organic film. Such an oxide film functions as do the oxides formed on alloys noted previously, but is not self-generating, so may be applied to various materials that otherwise would not produce a self-limiting oxide. A typical example of an organic system is the use of 1% benzotriazole in Vitel™ PET to protect copper.

- Try amorphous materials, as these are amazingly low in corrosion rates.
- Improve adhesion. Adequate adhesion is essential to prevent formation of aqueous films of water at interfaces, an important step in reducing electrochemical effects at interfaces.

6.1.7.2 Abrasion Resistance

Abrasion resistance is also a valuable property exhibited by functional thin films. This can be improved by lubricating the thin film and again, by improving adhesion to avoid shearing of thin film from substrate material during abrasion.

6.1.7.3 Adhesion

These are some mechanisms causing loss of adhesion:

1. Anodic undermining is associated with metal oxidation, providing an electron and counter-ion to a polymer, which is then reduced.
2. Cathodic blistering and gas blistering by corrosion are both likely the same mechanism. An adjacent metal, such as chromium, may oxidize, permitting protons from an available acid to be reduced. This reaction generates hydrogen gas bubbles, which then appear at interfaces or in the weakest material. This problem may be prevented by adequately oxidizing but not overoxidizing barrier layers.
3. Cathodic delamination is an oxidative process, with polymers typically suffering oxidation. Provided there is a small amount of UV light, then photo-oxidative degradation will cause similar oxidation reactions. As an example, polyimide turns a characteristic red color on oxidation. This ease of oxidation makes it a poor choice to use with some metals.
4. A commonly observed effect is loss of adhesion or delamination of metal film when wet, an effect which is reversible on drying. Hot water often produces this effect, suggesting that electrostatic and van der Waals are the only operative bonding forces, although hydrolysis of polymer may play a role. Reactive silane and titanate coupling agents help solve this problem.
5. Polymer swelling
6. Thermal cycling stress

I have just given two key reasons to improve adhesion: Decreasing corrosion and increasing abrasion resistance. Adhesion itself is also an essential requirement for most functional thin films. Consider the following ways to improve adhesion.

Polymers degrade with heat, hydrolyze, and are permeable to water vapor, of-

ten supplying acids or bases to interfaces between polymers and thin films. These limitations of polymers tend to speed corrosion and decrease adhesion, resulting in possible product failure.

An amorphous layer generated by exposure of some polymers to flashlamps may be used to toughen the polymer surface. On deposition of a thin film, the amorphous layer improves cohesive strength in the polymer near the interface.

Plasma priming at low pressure can improve adhesion and reduce corrosion rates of metal films deposited on polymer substrates. Prior to thin-film deposition, a plasma of an appropriate gas may be generated adjacent to a polymer surface. This process may create coordination or covalent bonds between polymer substrate and vacuum-deposited thin films. Oxidizing plasmas also remove surface layers and weakly bound contaminants, typically a thickness of 10 to 20 nm; plasmas often smooth polymer surfaces. The effects of plasma treatment are very stable, because plasma exposure produces a significant chemical change at the surface. Surface modification and adhesion enhancement by plasma chemistry is a process confined to surface layers. Vacuum ultraviolet light generated by such plasma affects deeper layers of polymers.

Reduced oxidation and improved stability may be achieved at interfaces between polymer surfaces and metal films by forming an appropriate barrier film at such interfaces. The barrier should coordinate well with the plasma-treated substrate. Barrier films should be thin, insoluble, and not electrically conducting; the last item helps decrease electrochemical interactions. Once the substrate is in vacuum, one wants to provide as many coatings as possible in one pass, such as plasma priming, barrier and nucleation films, functional films, overcoats, and lubricants.

REFERENCES

1. L. I. Maissel and R. Glang, *Handbook of Thin Film Technology* (McGraw-Hill, New York, 1983) originally published 1970.
2. J. L. Vossen and W. Kern, *Thin Film Technology* (Academic Press, New York, 1978).
3. Don Griffin, "The new C-MAG™ dual rotatable sputtering cathode," *Proc. Third Int. Conf. on Vacuum Web Coating*, San Antonio, Texas (1989), pp. 62–74.
4. R. F. Bunshah, *Deposition Technologies for Films and Coatings—Developments and Applications* (Noyes Publications, Westwood, NJ, 1982), pp. 231–235.
5. R. M. Bozorth, *Ferromagnetism* (Van Nostrand, New York, 1951).
6. J. Kieser, W. Schwarz, and W. Wagner, "On the Design of Vacuum Web Coaters," *Thin Solid Films*, **119** (1984) pp. 217–222.
7. W. M. Rohsenow and J. P. Hartnett, eds., *Handbook of Heat Transfer* (McGraw-Hill, New York, 1973), pp. 3–14 to 3–18.
8. H. Fenech and W. M. Rohsenow, "Prediction of thermal conductance of metallic surfaces in contact," *J. Heat Transfer, Trans. The American Society of Mechanical Engineers*, **85** (1963) pp. 15–24.

9. T. N. Veziroglu, Correlation of thermal contact conductance experimental results, *Prog. Astron. Aero.*, **20** (Academic Press, New York, 1967).
10. A. M. Clausing and B. T. Chao, *J. Heat Transfer*, **87** (May 1965) pp. 243–251.
11. M. M. Yovanovich and H. Fenech, *American Institute of Aeronautics and Astronautics 3rd Aerospace Sciences Meeting*, New York, January 24–26, 1966.
12. H. L. Atkins and E. Fried, *1st American Institute of Aeronautics and Astronautics Annual Meeting*, paper 64–253, Washington, DC, June 29–July 2, 1964.

CHAPTER 6.2

The Development of Ultra-High-Vacuum Technology for Particle Accelerators and Magnetic Fusion Devices

H. F. Dylla
Jefferson Laboratory and Department of Physics, College of William and Mary

6.2.1 INTRODUCTION

The development of the science and technology of ultra-high vacuum has been strongly coupled to the technologies of particle accelerators and magnetic fusion devices for the last 35 years. This coupling has been bidirectional—sometimes innovations in vacuum technology have driven machine development, and sometimes the technology transfer has been reversed. In the 1990s when we discuss accelerators and fusion, the immediate images we conjure are those of multibillion-

This paper was adapted from a paper given at the 40th National Symposium of the American Vacuum Society and was published in *J. Vac. Sci. Technol.*, **A12** (1994) 862. (Reprinted with permission of the American Institute of Physics.)

ISBN 0-12-325065-7
$25.00

dollar national or international megaprojects, such as the recently canceled Superconducting Super Collider (SSC) in the United States, the Large Hadron Collider (LHC), currently under design at CERN, the European Laboratory for Particle Physics in Geneva, and the International Tokamak Engineering Reactor (ITER), which is just entering an engineering design phase at three different international design centers. These are the current offspring in an evolution that began 70 years ago for particle accelerators and more than 40 years ago for magnetic fusion.

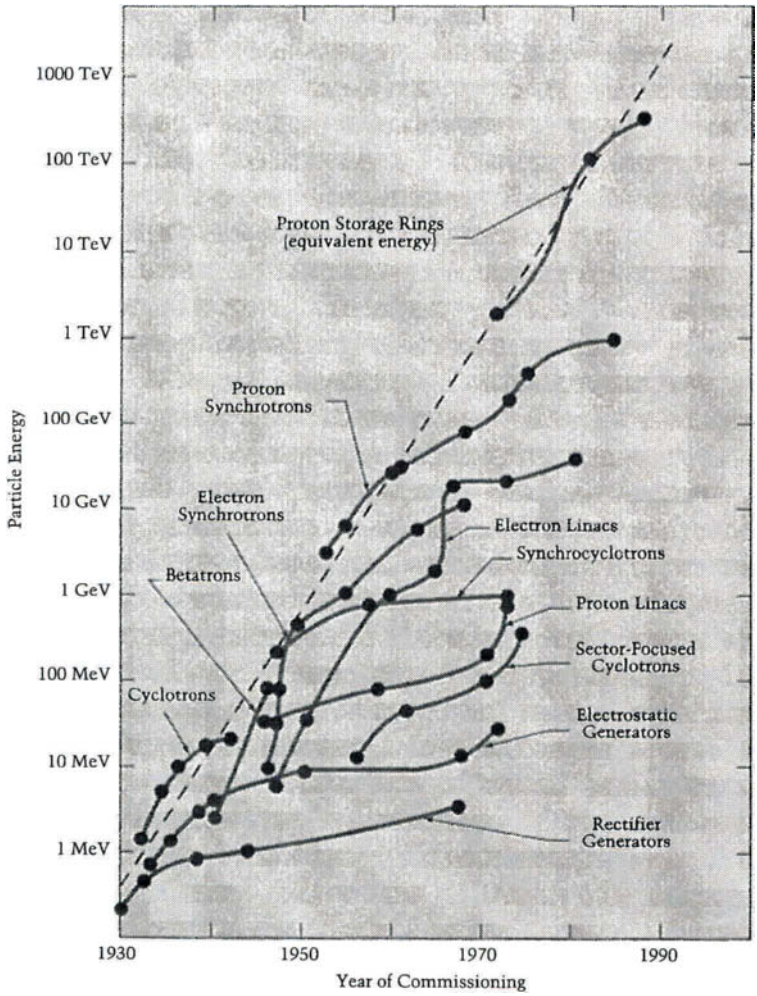
During this evolution a wealth of ultra-high-vacuum technology has been developed for or incorporated into these machines, including unique vacuum vessel designs, vacuum sealing techniques, high-performance vacuum materials, UHV cleaning and conditioning techniques, vacuum pumps, vacuum instrumentation, and vacuum controls. Often the technology development was associated with advances in the science underlying the technology, such as the surface physics necessitating the development of vacuum vessel-conditioning techniques. In this chapter I will highlight some of the key scientific and technical advances driven by the necessity for a vacuum environment for accelerating particle beams or for heating and confining high-temperature plasmas.

A well-known plot of the performance limits of various types of particle accelerators (Figure 1) was first drawn by Stanley Livingston in his 1954 book *High Energy Accelerators* [1]. The original plot showed acceleration limits increasing by an order of magnitude every six years. The predicted trend remained surprisingly accurate through 1985, largely because of the incorporation of new acceleration schemes roughly once a decade and the innovation of colliding beams versus fixed-target experiments.

6.2.2 STORAGE RINGS AND THE NEED FOR UHV

The family of accelerators that existed when Livingston published his now famous evolutionary plot (i.e., electrostatic, cyclotrons, betatrons, and synchrotrons) did not place significant demands on vacuum technology. As the machines grew larger, their design gave engineers the opportunity to incorporate cost-effective vacuum vessel fabrication techniques and vacuum pumping systems. The first incentive to incorporate UHV technology into accelerator designs came with the invention [2] of the storage ring based on a concept of Gerard K. O'Neill in 1956. O'Neill proposed a means of storing accelerated particles in a ring of strongly focused guiding magnets. Particles would be accelerated in a conventional synchrotron or linac and then injected into the storage ring (Figure 2). The full momentum transfer advantage of colliding beams could be obtained by intersecting counterrotating beams from two identical storage rings. The advantage of inter-

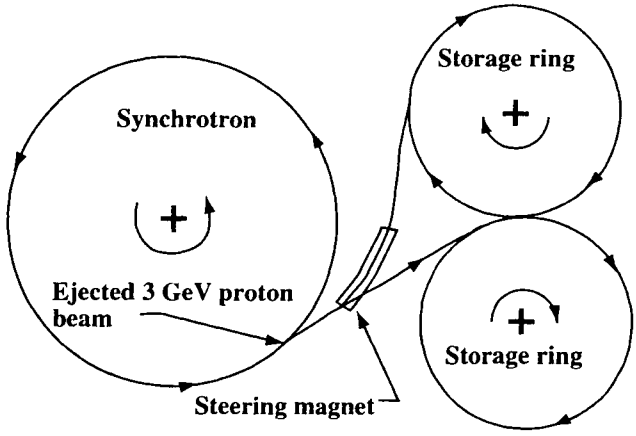
Fig. 1.



The Livingston curve of the evolution of accelerator performance first appeared in 1954 in Reference [1]. The version shown here (courtesy of the Stanford Linear Accelerator Center *Beam Line*) has been updated to include machines that came on line in later decades. The filled circles indicate new or upgraded accelerators of each type.

secting beams for achieving high center-of-mass energies was suggested by Kerst and associates [3] at about the same time as O'Neill's description of the storage ring; however, potentially higher beam currents, and thus higher interaction rates, were possible with O'Neill's scheme than with the maximum beam current allowed by Kerst's "beam stacking" method applied to conventional synchrotrons.

Fig. 2.



G. K. O'Neill's concept for the storage ring (1956), Reference [2].

In O'Neill's first paper [2] on the storage ring concept, he estimated that storage times would be a few seconds in the typical ($\sim 10^{-7}$ torr) high-vacuum environment. A follow-up paper published three years later by O'Neill and Woods [4] noted that if "vacuum technology already developed in thermonuclear power research were employed, storage times would be hours." They were referring to a neighboring development at O'Neill's institution, the C-Stellarator fusion device [5] under construction at Princeton in 1958 — which had a vacuum design specification of 10^{-9} torr. O'Neill began a collaboration with physicists at Stanford to construct two 500 MeV electron storage rings. Construction began in 1958, and electron scattering experiments began in 1961. The machine never produced results of great importance for particle physics; it was used to confirm quantum electrodynamics at 1 GeV collision energies and to measure the expected small cross section for producing heavy particles (muons). The Stanford-Princeton machine, along with a contemporary machine — the 250 MeV electron storage ring AdA, built at the Frascati laboratories in Italy [6] — were milestones in accelerator machine physics. These were the first in a long succession of storage rings and colliders to be built over the next two decades. At least two important and related lessons were learned from the operation of the Stanford-Princeton collider that would affect the subsequent development of these machines. UHV conditions, and hence useful electron storage times, could not be maintained in the stainless steel vacuum chamber because of contamination from the appendage oil diffusion pumps [7]. Second, a new phenomenon was recognized: A large dynamic outgassing effect was observed in the presence of stored electron beam that

E. Garwin [7] attributed to synchrotron-radiation-induced gas desorption. The effect has subsequently been well studied by numerous groups [8] because it is the dominant gas load to be dealt with in the design of electron or positron storage rings.

6.2.3 UHV FOR EARLY STORAGE RINGS

The next storage ring built at Stanford, SPEAR, the Stanford Positron-Electron Asymmetric Ring [9], and the first proton storage ring, the Intersecting Storage Rings (ISR) [10] built at CERN, are significant milestones in both particle physics and UHV history. Both machines incorporated specific design innovations to ensure that UHV conditions could be obtained with stored beam.

In general, storage ring vacuum chambers, because of their lengths and narrow cross sections, are highly conductance limited. Dynamic beam-induced vacuum problems could not be solved by simply increasing the number of appendage vacuum pumping systems. Innovations were necessary to meet both performance and realistic construction cost goals. These innovations involved the following:

- New vacuum vessel design and fabrication techniques
- Vacuum vessel-conditioning techniques
- Distributed vacuum pumps
- UHV-compatible synchrotron-radiation absorbers

The required dynamic pressures ($<10^{-9}$ torr) are approximately the same for achieving stable high current ($\geq$0.1 A) stored beams with reasonable storage times (~hrs) for both proton and electron rings—even though the offending gas-phase and surface interactions are different in the two cases. For proton rings, low gas pressures are required to minimize scattering of the beam and to minimize an instability that was first observed in ISR. With insufficient vacuum and insufficiently cleaned vessel surfaces, ionized residual gas species repelled by the beam impact the surface, releasing more gas by ion-induced desorption. At pressures greater than 10^{-9} torr, the effect can avalanche, leading to quenching of the beam current [11]. Careful measurements of ion-induced desorption coefficients by Mathewson and Achard [12] show desorption coefficients greater than unity for unconditioned metal surfaces. With various combinations of vessel bakeout and discharge cleaning, ion-induced desorption coefficients can be lowered by two to three orders of magnitude.

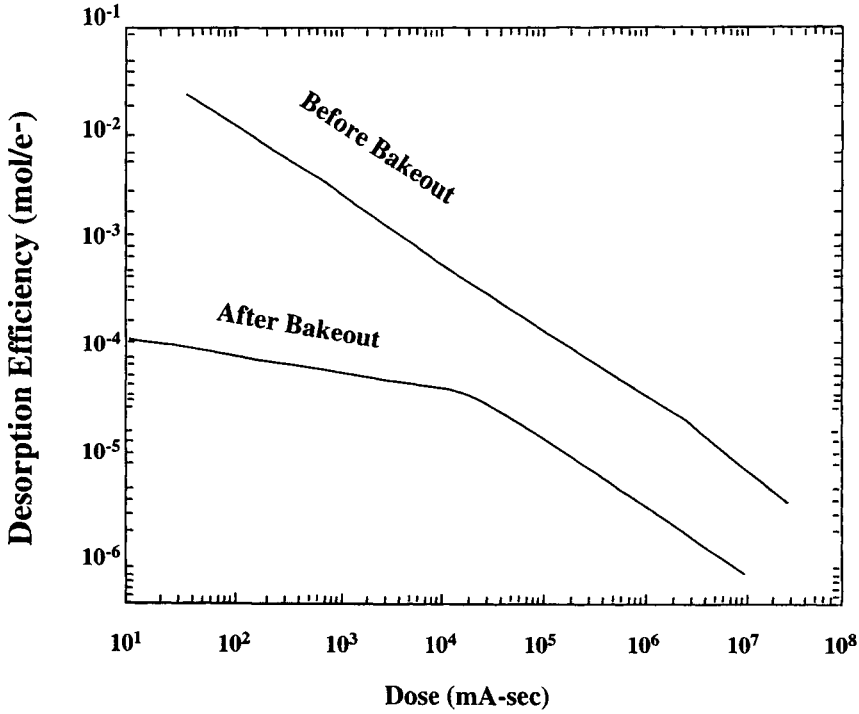
Let us examine the specific design attributes of the ISR device that affected its

vacuum performance. The ISR consisted of a pair of 1-km-long vacuum vessels that intersected at crossing points for colliding beam experiments. Proton beams with currents up to 20 A at energies up to 28 GeV were stored in each ring. The long, successful operation of this device from 1971 until 1984 owes much to the successful design and implementation of the vacuum system by the ISR vacuum group headed by E. Fischer [13,14]. Some fraction of the effort might be ascribed to brute force, in terms of incorporating the best of the existing UHV technology known at the time into the design of the system. For example, the 2-km-long, all stainless steel vacuum vessel was bakeable *in situ* to 300°C and was pumped by 300 triode ion pumps and 70 turbomolecular pumps. A significant effort was invested in properly characterizing the vacuum environment: 500 modulated Bayard-Alpert gauges and 36 residual gas analyzers instrumented the system. The ISR vacuum group developed several vessel-material-conditioning techniques that were used on many subsequent machines at CERN and elsewhere to optimize UHV performance. As a result of study of the diffusitivity of hydrogen in stainless steel [15], sheet stock that was used for vessel fabrication was prebaked in vacuum for 2 hours at 900°C, lowering the hydrogenic content by an order of magnitude. Significant efforts were devoted to developing and qualifying the effectiveness of chemical recipes for cleaning vacuum materials [16] and to developing discharge cleaning methods using argon and argon-oxygen mixtures [17]. These surface-conditioning studies were extended to aluminum surfaces [18] in the 1980s as CERN focused on the design of the 27-km Large Electron Positron (LEP) Collider, which became operational in 1989 [19].

At about the time the ISR device was being commissioned, the SPEAR device [20] at Stanford became operational (1972). Many of the design innovations developed for this collider were incorporated in subsequent electron-positron colliders and storage rings. Dynamic pressure requirements of $\leq 10^{-9}$ torr are also necessary for electron or positron storage rings, but in this case the primary gas-scattering mechanism is bremsstrahlung losses in electron-molecule collisions [21]. The beam lifetime determined by this scattering mechanism scales inversely with molecular mass, thus putting a penalty on the higher-molecular-weight components of the residual gas. As mentioned earlier, the source of dynamic gas load is the synchrotron-radiation-induced desorption. This phenomenon is now well understood in terms of gas desorption by photoelectrons generated by synchrotron radiation. Measurements to study the phenomenon were first done with electron sources to simulate the photoelectron emission, including the original measurement (Figure 3) by the SPEAR vacuum design team [7]. Later, actual measurements of photo-induced desorption coefficients were obtained using beam lines on storage ring light sources by the vacuum groups at CERN [8], Brookhaven [21], and KEK in Japan [22].

Measured desorption coefficients start out in the range of 10^{-2} to 10^{-1} de-

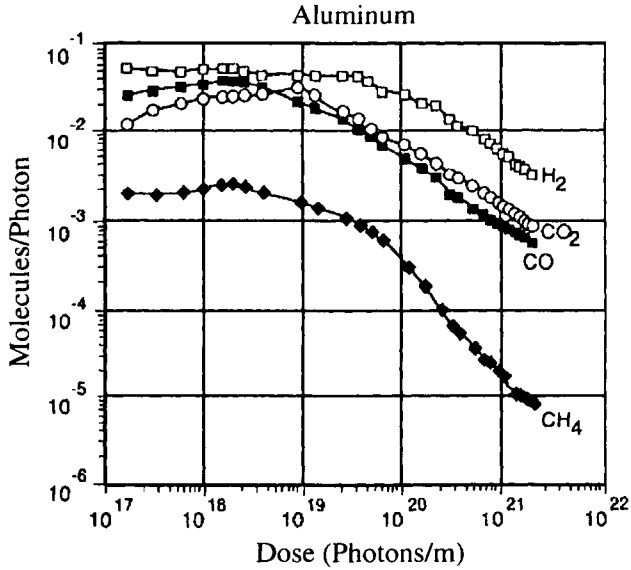
Fig. 3.



Electron desorption studies at Stanford (Garwin, 1963), Reference [7].

pending on whether the substrate is stainless steel, aluminum, or copper, and generally decrease inversely with the photon dose to 10^{-5} or 10^{-6} after a beam exposure that is typically 1 to 10 ampere-hours, corresponding to a photon dose of 10^{22} to 10^{23} photons per meter of vacuum vessel (Figure 4). This "beam cleaning" effect has turned out to be the most efficient *in situ* surface-conditioning technique for electron storage rings. The synchrotron emission that generates the photoemitted electrons, and the resulting photoemission rate, can be calculated for a given set of electron beam parameters and knowledge of the vacuum vessel geometry. Combining the results of these calculations with the measured photo-induced desorption coefficients allows the designers of new machines to design the vacuum chambers and associated pumping systems. Some effort has been made to model [23] the dose dependence of the photo-induced desorption, which is clearly rate limited by the diffusion of H, C, and O species from the near surface of the bombarded metal.

Fig. 4.



Photodesorption studies at CERN (Gröbner, 1983), Reference [8]. (Reprinted with permission from Elsevier Science Ltd., The Boulevard, Langford Lane, Kidlington OX5 1GB, U.K.)

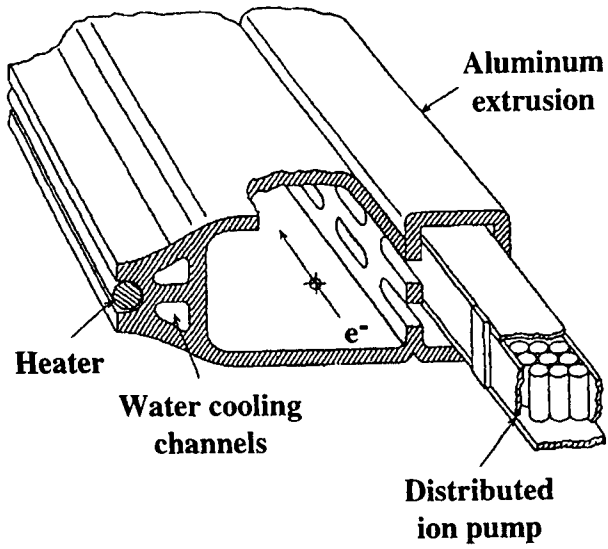
6.2.4

STORAGE RING VACUUM VESSEL AND PUMPING SYSTEM DEVELOPMENTS

The large dynamic gas loads that are present in electron storage rings have driven the development of unique vacuum chambers and associated pumping systems in order to meet the UHV operating requirements. Solutions adopted for the SPEAR storage ring [7] were innovative and became prototypical for the next two decades of storage ring design. The use of a simple pipelike geometry for accelerator vacuum chambers with appendage pumping ports spaced at convenient intervals between magnet gaps is sufficient for the modest vacuum requirements of most circular accelerators. Typically, pumping speeds per unit length of chamber of the order of 1–10 L/s-m can be obtained with appendage pumping. To cope with the synchrotron-radiation-induced gas load in an electron storage ring, so-called distributed pumps were placed within the beam chambers based either on ion pumps placed within the bending magnet field or nonevaporable getters (NEGs).

Figure 5 shows the beam chamber design [7] used in the SPEAR storage ring, which accommodated the use of distributed ion pumps. The chamber was fabricated using an aluminum (type 6061) extrusion technique. This was a cost-

Fig. 5.



SPEAR vacuum vessel design (1971), Reference [7].

effective solution for producing a long (200 m), multichambered, nonconcentric vessel. A cooling channel and ridged surface were extruded in the outer face of the chamber to deal with the localized heat load from incident synchrotron radiation. Many storage rings built after SPEAR, particularly those specifically designed as synchrotron light sources, have taken advantage of extruded aluminum chambers, because of the ease of fabrication and the ability to deal with the synchrotron radiation heat loads with coextruded cooling channels (and the intrinsic high thermal conductivity of aluminum). Extruded chamber designs made after SPEAR often included a multichamber arrangement that allowed the beam chamber to be isolated from a chamber that contained the distributed pumping elements, in order to minimize beam impedance effects.

The use of aluminum for storage ring vacuum chambers has generated a wealth of related development for the use of aluminum alloys in vacuum, such as effective surface-cleaning methods [18], aluminum–stainless steel bonding methods [24], vacuum-certified aluminum welding, and aluminum vacuum components [25].

As the synchrotron radiation power was increased for storage rings designed as dedicated light sources, specific synchrotron-radiation-absorbing structures had to be designed for the dipole magnet chambers to handle the localized heat loads. These structures have been as simple as copper bars brazed to the radiation plane of stainless steel chambers and as sophisticated as water-cooled absorbers with Cu, C, or Be as the radiation-absorbing material.[26]

The development of two classes of distributed pumps for storage rings has had

a significant impact on the further development of both ion pumps and nonevaporable getters. The innovative use in SPEAR of the dipole magnet fields as the confining field for ion pump Penning cells was an extension of the studies by Schurrman [27] and originally Penning [28] of the magnetic field dependence of cross-field gas discharges. Maler and Trachtenberg of Novosibirsk [29], and later Hartwig and Kouptsidis of DESY [30], developed specific formulas applicable to the design and performance of distributed ion pumps in both low- and high-field situations.

For the 27-km Large Electron Positron (LEP) storage ring at CERN, a more cost-effective distributed pumping scheme than *in situ* ion pumps was required. LEP was the first to incorporate nonevaporable getter (NEG) pumps within the vacuum chamber as the primary pumping element. An impressive quantity of design analysis and prototype testing of the selected NEG system (ZrAl alloy) was performed by CERN under the direction of C. Benvenuti [31]. As a result, the system has performed well since the startup of LEP in 1989. Numerous other storage rings [32,33] had proceeded with incorporating NEGs as the distributed pumping element even before LEP startup.

6.2.5 COLD-BORE MACHINES

What could very well be the last members of the family of colliders for high-energy particle physics research are represented by the Superconducting Super Collider (SSC) in the United States (which was canceled by U.S. Congress in September 1993), and the Large Hadron Collider (LHC) at CERN (which is under design and awaits construction approval by the CERN member nations). The enormous cost for these machines (> \$11B for the SSC and > \$4B for the LHC), which is driven by their size and need for thousands of state-of-the-art superconducting magnets, often outweighs discussions of the scientific benefits.

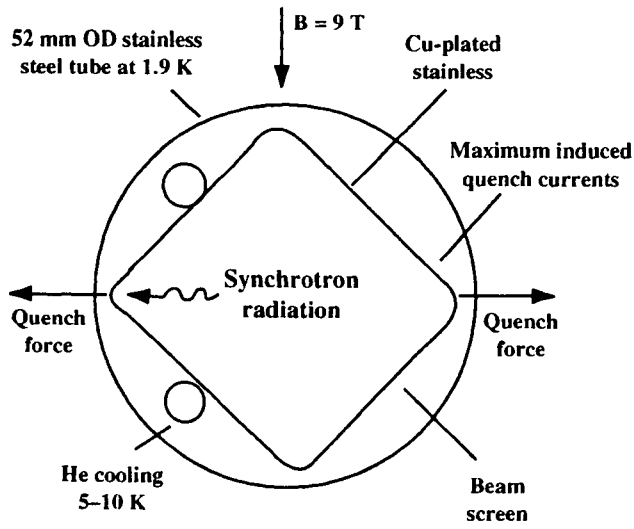
For the purposes of the present historical discussion, the design work that has been accomplished for the SSC and LHC has already contributed to the UHV knowledge base because of the particular problems associated with storage ring vacuum chambers that must operate at liquid helium temperatures [34,35]. To minimize magnet costs, both the SSC and LHC beam chambers are relatively small in diameter ($\approx$5 cm) and are thermally shorted to the magnet temperature (4.2 K for the SSC and 1.8 K for the LHC).

Normally, the synchrotron radiation emission from the SSC or LHC proton beam would not be problematic in terms of producing a dynamic gas load if the vacuum vessel was at ambient temperature. For the case of SSC, the collider would have stored 70 mA proton beams at 20 TeV. The synchrotron radiation power

onto the dipole magnet vacuum chambers is 0.14 W/m. The gas load desorbed from such a radiation flux could be easily handled by the addition of modest distributed pumping as demonstrated in the electron rings. With ambient temperature surfaces, the photodesorbed gases will not be readsorbed on nearby surfaces and would have a high probability of being removed from the gas phase by the distributed pumps.

The case of photo-induced desorption and readsorption from surfaces near 4 K is quite different. All the desorbed species of interest (CO, CO_2, CH_4 and H_2O) except for H_2 are pumped well and have low equilibrium vapor pressures on 4 K surfaces. Isotherm measurements on clean stainless steel at low temperatures show that H_2 is pumped and the H_2 vapor pressure remains low ($<10^{-11}$ torr) only if the H_2 surface coverage remains below a monolayer [36,37]. Dynamic pressures for the operation of both the SSC and LHC must remain below the 10^{-10} torr range in order to prevent excessive beam scattering and excessive heat load on the magnet cryostat as a result of the beam scattering. A proposed design feature (Figure 6) within the cold-bore vacuum chambers to prevent the dynamic pressure from exceeding the 10^{-10} torr limit is the incorporation of an intermediate temperature liner (@ ~20 K) between the beam aperture and cold vessel wall. This liner would intercept the synchrotron radiation and be partially slotted or perforated to allow desorbed H_2 molecules into the interspace. The 1.9 K operating temperature of the LHC cold bore is sufficiently low to pump H_2 above monolayer quantities without exceeding the dynamic vacuum limits. However, for the

Fig. 6.



LHC vacuum vessel design (1993), Reference [35]. (© 1993 IEEE)

SSC, some type of cryosorbing material would have to be engineered to pump multilayer quantities of H_2. The most likely candidate appears to be charcoal, based on previous studies of the pumping characteristics of this material for the fusion program [38] and recent work instigated by the SSC.

With the demise of the SSC, these studies will most likely proceed entirely focused on the application to LHC. The primary issues to be resolved are an understanding of photo-induced desorption phenomena at low temperatures (1.8–20 K) on candidate vessel materials, and the design of a cost-effective cold-bore liner. Information gained on the first topic adds to our basic understanding of this important process and has application to other important scientific and technical problems such as the formation of molecules on interstellar dust grain surfaces [39], and the dynamic pressure response of the superconducting synchrotron light sources used for X-ray lithography [40].

6.2.6 SUPERCONDUCTING RF ACCELERATORS

In addition to the use of superconducting magnets for saving power in accelerator operations, the other significant use of superconductivity in accelerator technology is RF acceleration cavities [41]. Linear accelerators consist primarily of an array of RF cavities, separated by occasional magnetic elements for correcting beam optics. In synchrotrons, the energy that the circulating beam loses due to synchrotron radiation and other processes is restored by RF cavities interspersed among the guiding magnets. For both applications, superconducting RF cavities save operating costs, offer lower impedance to the accelerating beam, and can be operated CW (continuous wave) at high gradients. The disadvantages of superconducting RF cavities are the cost and complication of cryogenic cooling and the smaller operating experience base. The latter concern is being alleviated. During 1993 the installed acceleration capacity provided by superconducting RF cavities exceeded 1 GeV, largely due to the completion of the superconducting linac at CEBAF, the Continuous Electron Beam Accelerator Facility of Jefferson Laboratory in Newport News, Virginia [42]. Large installations of operating superconducting RF systems are in the storage rings at KEK [43], DESY in Hamburg, Germany [44], and LEP [45]; the latter machine is in the process of upgrading all its RF cavities to superconducting versions in order to double the achieved final energy (to 100 GeV) in a cost-effective manner.

Over 20 years of development has been invested in superconducting RF technology, resulting in the current state of the art, which is capable of fabricating production-scale cavity assemblies with accelerating gradients greater than 20 MV/m [46]. These high-performance cavities will also find use in compact CW linacs for driving free-electron lasers, which will be useful for industrial

applications such as materials processing where high-average-power, energy-efficient UV and IR light sources are needed [47]. The successful development of superconducting RF cavities is strongly tied to UHV technology. High-performance and reliable cavities require extremely careful cleaning and surface treatments, assembly in an environment free of particulate contamination, and vacuum-sealing techniques that are UHV compatible [48].

6.2.7 THE NEXT-GENERATION BIG ACCELERATOR?

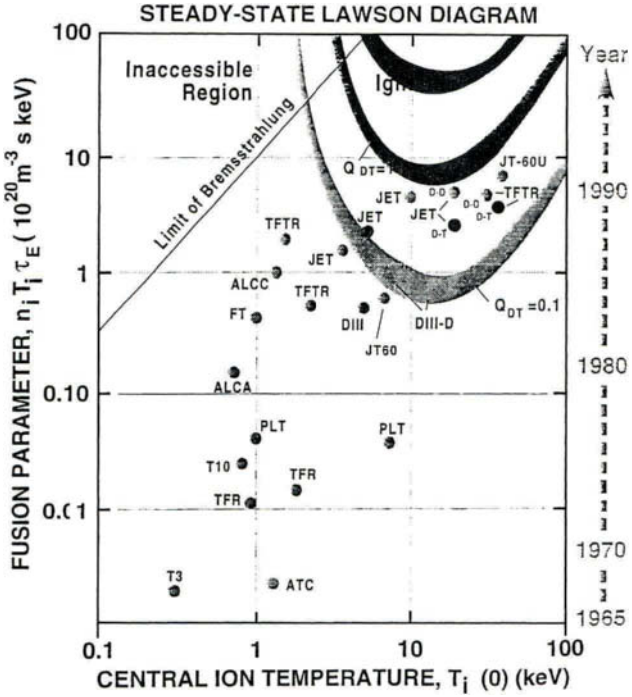
As I noted in the discussion of the Livingston curve, the evolution of accelerators was punctuated by new technologies that allowed performance to grow at affordable costs. Have the circular colliders reached this limit with the SSC and LHC? To provide alternative technologies for the next generation of particle accelerators, several international collaborations are exploring designs for linear colliders [49]. This Next Linear Collider project has teams designing various options for a pair of 250 GeV linacs for e^+e^- collisions. Several teams are investigating warm accelerating structures from 3 to 30 GHz, and one team is investigating a superconducting RF structure at 1.3 GHz. The primary problems with the warm structures, which are largely based on the success of the Stanford Linear Collider (which became operational in 1980 [50] shortly before LEP), are the extremely tight alignment specifications and the availability and cost of the required RF power. The primary problem with the superconducting option is the cost of the cavity/cryostat structure. As in most endeavors, competition is healthy, and the design exercises have already generated valuable results in materials, RF, and cryogenic engineering.

The field of accelerator physics has always managed to attract a cadre of bright scientists and engineers willing to gamble on dark horse candidates for novel accelerator schemes that have both high risk and high potential payoff, including the present exploratory work on laser-plasma beat-wave accelerators and wakefield accelerators. Who knows which dark horse candidate will add another family of machines to the Livingston curve over the next 20 years?

6.2.8 THE MAGNETIC FUSION ROAD MAP

In the history of particle accelerators, the Livingston curve has provided a useful road map for the evolution of the various families of accelerators. The field of development of devices to harness thermonuclear fusion for power generation has

Fig. 7.



The Lawson criterion for energy gain in D-T plasmas, Princeton Plasma Physics Laboratory.

an analogous road map. It is called the Lawson curve, or Lawson criterion, and is based on John Lawson's calculations [51] of the minimum conditions of plasma density, plasma temperature, and energy confinement necessary for energy breakeven in a D-T (deuterium-tritium) plasma (see Figure 7). For the first two decades of fusion research, devices were designed to magnetically confine and heat low- to moderate-density hydrogenic plasmas. During the last two decades, "inertial confinement" schemes, based on laser or ion beam compression of D-T pellets, have provided an alternative development path for fusion energy. In such magnetic confinement schemes, plasmas would have to obtain densities of $\geq 10^{19}$ m^{-3}, temperatures ≥ 10 keV and energy confinement times >1 s to reach the Lawson condition. To the physicists and engineers who started this field of research over 40 years ago this seemed to be a rather straightforward problem; however, it has turned out to be one of the most difficult scientific and technical challenges of our times. A sustained effort over the last four decades from practitioners in Europe, Japan, Russia, and the United States has seen the plasma triple-product rise over 10 orders of magnitude from the primitive, small devices of the 1950s to within 50–80% of the Lawson criterion using the gigantic tokamaks at Princeton [52] and Abingdon, England [53]. UHV technology has played a sig-

nificant role in the development of magnetic fusion devices. Before I cite specific examples that are largely based on the developments on the last two decades, it is informative to review the early history in the 1950–1960s, when the key problems of magnetic fusion and high-temperature plasma research were being defined.

6.2.9 THE EARLY HISTORY OF MAGNETIC FUSION

Fusion research has always had strong political overtones from its initiation in the 1950s through several large boom-or-bust cycles of government funding to its present uncharted future. The possibilities of fusion power or "controlled thermonuclear reactors" were clearly in the minds of physicists who witnessed the first man-made thermonuclear reactions with the explosion of the first D-T weapons. The first serious effort in the United States to launch a fusion research program was initiated by an announcement in March 1951 by the Argentinean dictator Juan Peron that his nation had harnessed fusion power with the help of an exiled German scientist, Ronald Richter. The claim proved unfounded, but it helped launch a secret program of fusion research that was code-named Project Sherwood [54]. This program had proponents who championed three different schemes for magnetically confining plasmas. Lyman Spitzer from Princeton University, who is generally given credit for having first elucidated the basic physics of magnetic confinement, conceived of a figure-eight confinement geometry he called a stellarator. A solenoidal field would confine the plasma, an inductively driven current would heat the plasma, and the twist in the field geometry would compensate for the tendency for charged particles to drift out of a simple torus. James Tuck, a British physicist on loan to the Manhattan project during World War II, had returned to Los Alamos with a scheme he developed with his colleagues at Oxford to confine and heat plasmas in a sample focus by "pinching" the plasma to a small radius stream by rapidly increasing the toroidal field. Tuck had a more sanguine view of the probable vicissitudes of confined plasmas and named his device the Perhapsatron. A third team, led by Richard F. Post and Herbert York, at the Livermore branch of the University of California's Radiation Lab (the ancestor of the Lawrence Livermore Laboratory), proposed an open solenoidal confinement scheme they termed "mirror machines."

All the early schemes were fraught with difficulties: the plasma confinement appeared poor, plasma temperatures were low, and the primitive vacuum technology that was used guaranteed that the plasma species of interest was overwhelmed by impurities from the confining chamber walls. Adding to the difficulties were the poor state of developed instrumentation to measure plasma properties and a primitive level of theoretical understanding of this complicated multicomponent

fluid system. The situation took on a decidedly different perspective and pace when the secret Project Sherwood was declassified in 1958 on the eve of the famous Atoms for Peace Conference in Geneva. At this time it became clear that the ambitious Russian program had not progressed significantly ahead of the Western efforts. The claims of near success had once again been heralded, this time by the British with their Zeta model of a plasma pinch in 1957, but by the time of the Geneva Conference the neutrons observed in the Zeta experiments were shown to be not thermonuclear in origin [55]. A growing realization of the difficulty of the confinement problem, the embarrassment of the false claims of success, and the visibility engendered by the program declassification forced the fusion research to a more balanced mix of theory, machine, and measurement.

6.2.10 MODEL C: THE FIRST UHV FUSION DEVICE

When Spitzer had conceived the stellarator concept, he developed a plan for moving the technology along an orderly route from the pioneering Model A, which was a table-top demonstration, to a larger Model B, designed to push the plasma parameters, to an engineering prototype (Model C), which would be a scale model reactor, to finally a full-scale prototype reactor (Model D) [56]. A remarkable document, completed by Spitzer's team in 1954, actually spells out many of the important subsystems that a fusion reactor would require: an auxiliary heating system to reach plasma ignition temperatures, a magnetic "diverter" to remove plasma impurities, and a lithium blanket for absorbing the fusion product neutrons and extracting energy from the reactor. The disappointing results coming in from the experiments on the B-series of stellarators, and similar results from the plasma pinch and mirrors programs, considerably stretched out Spitzer's development plan for a reactor. Model C would not be a scale model reactor—it would be the next step in elucidating the plasma physics. However, Model C would represent the largest single investment ($35M) in the program to date, and for the first time industrial contractors would participate in the design and construction of the machine. The Model C stellarator came to life in 1961. Its stainless steel vacuum vessel was the largest UHV system built to date [5]. Special double-joint, gold wire flange seals allowed the system to be baked to 450°C. After bakeout, base pressures for the system were in the 10^{-10} torr range. Because of the concerns for hydrocarbon contamination, the Model C was pumped with two large mercury diffusion pumps that were isolated from the torus by lead-sealed valves and freon-cooled traps. Mercury was chosen as the pumping fluid because its accidental presence in the torus vacuum could be detected very sensitively by plasma spectroscopic techniques, and the high-temperature vessel bakeout would easily remove the contamination. The performance of the Model C vacuum sys-

tem was a huge success—the plasma performance, on the other hand, was disappointing.

For most of the 1960s none of three confinement schemes being pursued in the United States or United Kingdom were making any significant progress toward the Holy Grail represented by the Lawson criterion. The knowledge base on fully ionized plasmas was expanding at a healthy rate fed by careful measurements with new plasma "diagnostic" instruments and increasingly sophisticated plasma theory. Despite these advances, by the end of the 1960s neither the measurements nor the theory could tell the experimenters why their plasmas remained relatively poorly confined and cold [57].

6.2.11 THE RUSSIAN REVOLUTION IN FUSION: TOKAMAKS

As the 1960s drew to a close, a pair of closely coupled political events caused a dramatic change in the pace of fusion development. The cold war prevented any significant exchange of information between the Western and Soviet programs except for brief interactions at plasma physics conferences organized by the International Atomic Energy Agency (IAEA). At the 1965 [58] and 1969 [59] conferences, the Russians claimed that their toroidal magnetic chamber or "tokamak" geometry was producing plasmas with confinement times of milliseconds and plasma temperatures near a kilovolt. These results were met with skepticism, because they represented at least two orders of magnitude improvement over the best results from the competing Western devices. The skepticism faded in 1969 when a thaw in cold war relations allowed the leader of the fusion effort at the Kurchatov Institute in Moscow, Lev Arsimovitch, to come to MIT for a lecture on disarmament issues. An impromptu series of lectures on the results from the Kurchatov T-3 tokamak began to convince a steady stream of American plasma physicists that traveled to Cambridge to hear and question Arsimovitch. The last doubts disappeared when a team of British physicists visiting the Kurchatov Institute cabled to Washington that they had confirmed the T-3 results with laser scattering measurements of plasma temperature and density made with equipment they had imported from their laboratory at Culham [59].

The tokamak race was on. Tokamaks were under construction at Oak Ridge [60], MIT [61], and the General Atomics Co. [62] in San Diego. The Model C stellarator at Princeton was converted to a tokamak in four months, and in the first two months of operation reconfirmed the confinement properties and stability limits of the tokamak geometry. Moreover, the confinement properties appeared to possess a favorable scaling with size: The energy confinement time would improve with the square of plasma radius. Plans were quickly drawn up for a second generation of tokamaks by scaling up a factor of 2 in size, followed by another

factor of 2 increase in size, which would yield a device capable of reaching the Lawson condition. These ambitious plans languished for two years until the Arab oil embargo of 1973 pumped significant funding into any energy research. The Princeton Large Torus (PLT) [63], a tokamak with a 6,000 L vacuum vessel and 0.5 m plasma radius came on line late in 1975. With the use of high-current, hydrogen-neutral beams developed at Oak Ridge National Laboratory, PLT became the first fusion device to exceed the Lawson criterion minimum temperature of 10 keV in the summer of 1978 [64].

Scientific and technical progress came at an exhilarating pace as the fusion budgets and resulting activity grew at 30–40% a year for most of the decade following the threshold year of 1973.

6.2.12 PLASMA IMPURITIES AND VACUUM TECHNOLOGY

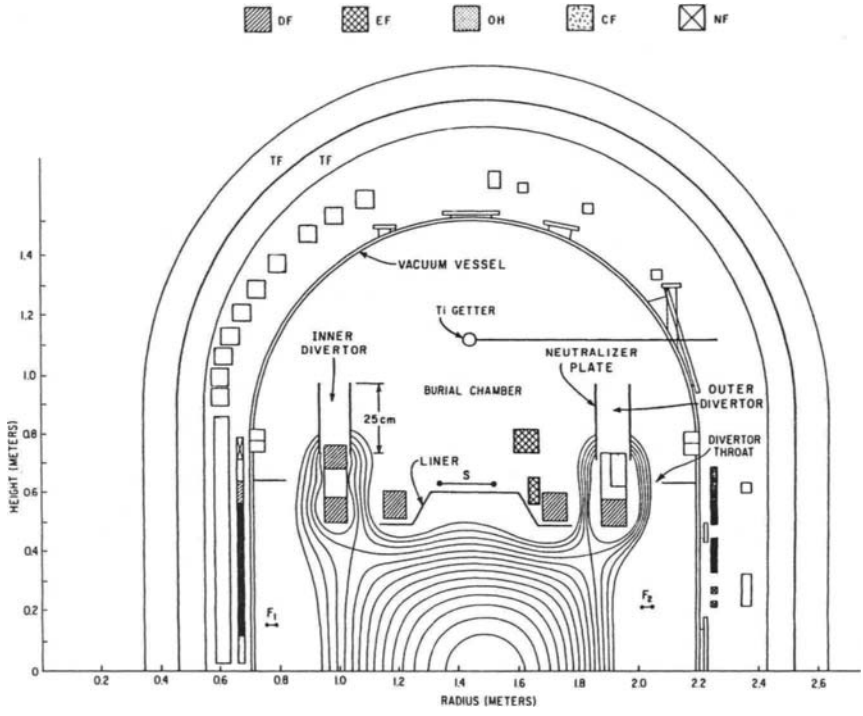
The achievement of the high plasma temperatures in the PLT tokamak did not come without several key developments in controlling how the plasma interacts with the vacuum vessel wall. Spectroscopic and plasma resistivity measurements on the first generation of Western tokamaks (ST and ATC at Princeton, Ormak at Oak Ridge, Pulsator in Germany, and TFR in France) were indicating that plasma impurities were a persistent problem. The ATC tokamak, which was built at Princeton shortly after the Model C conversion to test a magnetic compression scheme of heating plasmas [65], achieved in 1974 the first pure hydrogenic plasma that was undiluted with carbon and plasma impurities desorbed from the vacuum vessel wall. This result was achieved by evaporating titanium onto a large fraction of the vessel wall prior to a plasma discharge to suppress the impurities, combined with hydrogen gas injection during the discharge [66].

In 1975, the first in a series of compact, high-field tokamaks built at MIT, Alcator A, also achieved pure, "$Z^{eff} = 1$" plasmas [67], by a combination of UHV vessel design, programmed fueling of the discharge with hydrogen gas injection, and pioneering use of a pulse discharge cleaning technique to condition the vacuum vessel wall before exposure to high-temperature plasmas. The discharge cleaning technique [68] was developed by Robert Taylor and rapidly adopted by all the existing tokamak experiments. Ironically, as a combination of the discharge cleaning and Ti-gettering techniques was applied to the PLT device, low-Z (carbon and oxygen) impurities were reduced to the point that these ions no longer cooled the edge plasma by radiation. The resulting edge plasma was sufficiently energetic to sputter excessive quantities of metal impurities from the tungsten "limiter" structure that defined the plasma radius and protected the vacuum vessel wall. A switch of limiter material to a high-purity graphite solved the problem for PLT [69] and helped launch a development program for low-Z refractory

materials that would be needed for "first-wall" structures in the succeeding generation of tokamaks [70].

The problems of impurity generation and impurity suppression in tokamaks were deemed of sufficient importance to the development of fusion that significant portions of experimental programs on many machines and several dedicated devices were devoted to impurity studies. The Oak Ridge Impurities Studies Experiment or ISX device at Oak Ridge tested a prototype of the PLT graphite limiters [71], developed a Cr-gettering technique [72], and investigated unpolar arcing as an impurity generation source [73]. Impurity transport and improved hydrogen injection techniques were studied at the Doublet tokamak at General Atomics [74]. Large-scale tests of magnetic diverter structures for controlling the transport of both hydrogenic and impurities species were performed on the PDX tokamak at Princeton [75] and the ASDEX tokamak at Garching, Germany [76]. The latter two tokamaks, which became operational in the early 1980s, required large vacuum vessels ($\sim$36 m^3) to contain the internal magnetic structures for the diverter and a ballast volume for pumping the plasma particles that were neutralized on the diverter target plate (Figure 8). The diverter volumes were equipped

Fig. 8.



The vacuum vessel of a diverter-pumped tokamak (PDX, 1980), Princeton Plasma Physics Laboratory.

with large arrays of titanium evaporation sources to provide up to 2×10^5 L/s of hydrogenic pumping capacity. The PDX device was also used to test a prototype nonevaporable getter array [77] (based on the Zr-Al alloy) and a large-area, toroidal limiter fabricated from graphite tiles [78]. Because of the large size of the PDX and ASDEX vacuum vessels compared to the plasma volumes (6:1), both machines became test beds for developing hydrogen glow discharge techniques for cleaning complicated vacuum vessel structures [79,80].

6.2.13 TOWARD THE BREAKEVEN DEMONSTRATIONS

The rapid pace of the fusion program in the mid to late 1970s compressed and overlapped the design and operational schedules of three generations of machines. This accelerated schedule was most evident in the history of the large tokamaks built to reach the Lawson criterion for energy breakeven. Before the first plasma discharge was fired in PLT (or in a similar-sized device, the T-10 tokamak at Kurchatov), preliminary designs were in place for three tokamaks that would be sufficiently large to reach energy confinement times near one second, and include sufficient plasma heating and fueling capability that the remaining Lawson conditions of plasma temperature and density could be simultaneously achieved. These three large devices became the flagship experiments of the United States, European and Japanese fusion programs: the Tokamak Fusion Test Reactor (TFTR) [52] at Princeton began operating in 1982; the Joint European Torus (JET) [53], built next to the Culham Laboratories in the United Kingdom, began operating in 1982; and the JT-60 [81] tokamak in Japan began operating in 1984. All three machines are still operating and have had their original programs extended to the mid- to late 1990s. TFTR and JET were both designed to operate with D-T mixtures, and therefore have the complication of dealing with radioactive gas handling [82,83] and the additional activation of the vessel [84] by the 14 MeV D-T neutrons. JT-60 operations were confined to hydrogen plasmas, and a recently completed upgrade of the machine (JT-60U) has allowed the use of deuterium fueling [85]. The complications and expense of D-T fueling have postponed significant D-T experimentation to the last operations period for both TFTR and JET. Late in 1992, the JET physicists fired a single D-T fueled discharge that produced a megawatt of fusion power for a fraction of a second [86]. Unraveling the plasma physics in these devices is difficult with a single discharge; however, this short-lived experiment did achieve a long-standing program goal of producing a significant quantity of fusion power, and the subsequent exercise of tracing the tritium inventory and decontaminating the vessel first-wall compo-

nents was very informative [87]. As presently scheduled, JET will return to an extended series of D-T experiments in 1996. TFTR began a two-year series of D-T fueled experiments [88] in 1993, and subsequently the machine will be decommissioned. Before these long-awaited D-T experiments in JET and TFTR, both machines achieved impressive plasma parameters on the Lawson chart with the use of 20–30 MW of plasma heating [89,90] (with both neutral beam and ion cyclotron RF sources) and frozen hydrogen pellet injection systems [91]. Plasma parameters within 80% of the minimum Lawson requirements have been achieved in addition to pushing the plasma temperature records to three or four times the achievements of PLT [89]. Recently, experiments in JT-60U [92] and the D III D [93] device (at General Atomics) have also achieved these high-temperature regimes.

These impressive results in the large tokamaks have required continued development of first-wall materials, first-wall structures, and the conditioning techniques for these materials needed for edge-plasma particle and impurity control. The TFTR vacuum vessel is protected with a toroidal array of graphite of over 2000 tiles [94]. Special discharge cleaning techniques were developed to remove the large quantities of H_2O that are absorbed by this structure during vents to atmosphere [95]. The highest performance discharges were achieved in TFTR after the graphite first-wall was induced to behave like a plasma pump by conditioning the structure with a regimen of special He discharges [96]. Before initiation of D-T discharges on TFTR, extensive studies of tritium trapping and removal mechanisms were developed to recycle tritium trapped in the vessel [97]. After operational experience was gained initially with the use of fine-grained graphite for first-wall armor, JET replaced graphite in high-flux areas with Be [90], and the TFTR, D III D, and JT-60U devices upgraded to carbon-fiber-composite materials [98]. After pioneering applications of a large-scale vapor deposition process successfully coated the entire first-wall of the TEXTOR device in Germany with thin layers of boron-carbide [99], the technique was applied to other tokamaks (ASDEX, TFTR, D III D) as a means of *in situ* modification of first-wall properties.

The careful control of the surface properties of the first-wall structures in these large tokamaks would not be possible if high-reliability "second-wall" or vacuum containment structures were not maintained. The complicated vacuum vessels must satisfy many constraints, including UHV standards for leak-tightness despite hundreds of access ports for diagnostic instrumentation and plasma heating equipment, and structural integrity to withstand cyclic loading due to bakeout cycles and large eddy-current-induced forces [100]. The need for high-capacity, noncontaminating vacuum pumping systems for these large vessels drove the design of large (>3500 L/s) turbo pumps [101], and the associated high-power neutral beam heating systems drove the design of large He cryocondensation pumps with speeds exceeding 10^6 L/s for hydrogenic species [102].

6.2.14 THE NEXT STEP IN FUSION

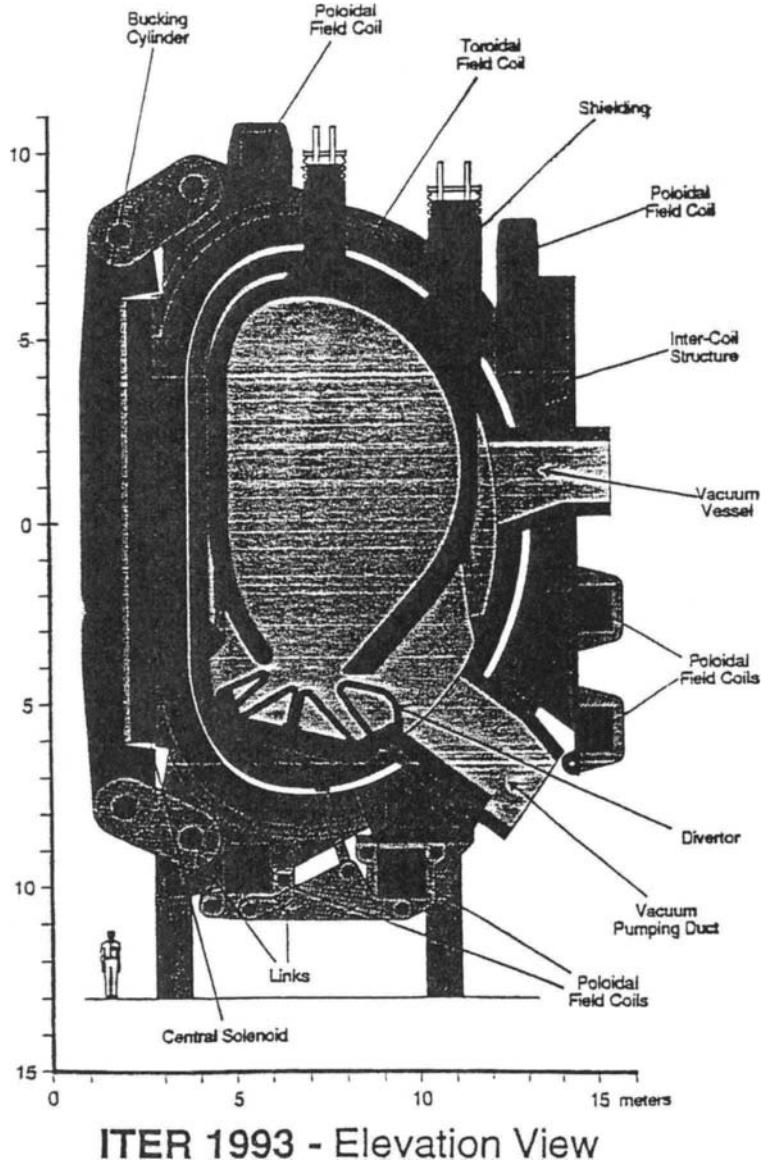
So where will the fusion program go next after the D-T demonstration experiments on JET and TFTR are complete? The international fusion research community and their supporting government agencies have been debating the next step for more than a decade. An era dedicated to demonstrating the scientific feasibility of fusion energy is drawing to a close — after a considerably longer time and larger investment than was envisioned by the scientists who launched Project Sherwood. Some argue that the next step must address engineering feasibility of a tokamak reactor if the goal of producing affordable fusion energy is to remain in the foreseeable future. Others argue that this next step is impractical and unaffordable, and the program should explore different physics regimes that might result in a more appealing reactor design.

The next-step fusion device that is feeding this debate is a megaproject called ITER, the International Tokamak Engineering Reactor (Figure 9). This device is currently entering the engineering design phase after a three-year conceptual design effort was completed by teams from Europe, Japan, Russia, and the United States. The goals of the machine include demonstration of a 10^3 second plasma burn with power extraction from a Li-blanket system. Effective first-wall protection, diverter pumping, tritium recovery, and remote maintenance systems would have to be demonstrated. The machine would probably cost somewhere between the estimated costs of the LHC ($4B) and the SSC ($11B). Clear thinking by the scientific community and sensible planning by the government-industry partnerships who will build the next-step fusion device (and the next large accelerator) are needed to keep these important fields moving forward.

ACKNOWLEDGMENTS

The author acknowledges the able assistance of the following colleagues with the preparation of this paper. C. Sinclair at Jefferson Laboratory, D. Carroll at Princeton Plasma Physics Laboratory, S. Wofford at Lawrence Livermore National Laboratory, E. Johnson at Brookhaven National Laboratory, K. Wilson at Sandia National Laboratory, and A. Mathewson at CERN. This work was supported by U.S. DOE Contract No. DE-AC05-84ER40150.

Fig. 9.



The ITER vacuum vessel (1993), Princeton Plasma Physics Laboratory.

REFERENCES

1. M. S. Livingston, *High-Energy Accelerators* (Interscience, New York, 1954).
2. G. K. O'Neill, *Phys. Rev.*, **102** (1956) 1418.
3. D. W. Kerst et al., *Phys. Rev.*, **102** (1956) 590.
4. G. K. O'Neill and E. J. Woods, *Phys. Rev.*, **115** (1959) 659.
5. G. Lewin and D. Mullaney, *1963 Vacuum Symp. Trans.* (Pergamon, New York, 1964), p. 176.
6. C. Bernardini, G. F. Corazza, G. Ghigo, and B. Touschek, *Nuovo Cimento*, **18** (1960) 1293.
7. E. L. Garwin, 3 BeV colliding beam vacuum system, *SLAC Memorandum* (August 1963).
8. O. Gröbner, A. G. Mathewson, H. Störi, and P. Strubin, *Vacuum*, **33** (1983) 397.
9. SLAC Storage Ring Group, *Proc. 8th Intern. Conf. on High Energy Accelerators CERN, 1971* (European Organization for Nuclear Research, Geneva, 1971), p. 145.
10. CERN Report No. AR/Int. SG/64-9 (1974).
11. R. S. Calder, *Vacuum*, **24** (1974) 437.
12. A. Mathewson and M. H. Achard, *Proc. 3rd Intern. Vacuum Congress, Vienna, 1977*, Vol. 2, edited by R. Dobrozemsky et al. (Berger, Vienna, 1977), p. 1217; also see A. Mathewson, *Vuoto*, **17** (1987) 102.
13. E. Fischer, *J. Vac. Sci. Technol.*, **2** (1965) 142.
14. E. Fischer, *J. Vac. Sci. Technol.*, **9** (1972) 1203.
15. R. Calder and G. Lewin, *Brit. J. Appl. Phys.*, **18** (1967) 1459.
16. A. G. Mathewson, in *Proc. Intern. Symp. on Plasma Wall Interactions, Julich, 1976* (Pergamon, Oxford, 1977), p. 517.
17. E. Fischer, *Jpn. J. Appl. Phys.*, Suppl. 2, (1974) 199.
18. A. G. Mathewson, M. Andritschy, A. Grillot, O. Gröbrer, and P. Strubin, in *Surface Conditioning of Vacuum Systems*, AIP Conference Proc. No. 199 (American Institute of Physics, New York, 1990), pp. 110–123.
19. H-P. Reinhard et al., *Proc. 11th Intern. Vacuum Congress, Cologne* (1989).
20. J. LeDuff, *Nucl. Instr. Meth.*, **A239** (1985) 83.
21. H. J. Halama, in *Vacuum Design of Synchrotron Light Sources*, edited by Y. G. Amer, S. D. Bader, A.R. Kraus, and R.C. Niemann, AIP Conference Proceedings No. 236 (American Institute of Physics, New York, 1991), pp. 39–51.
22. T. Kobari, M. Matsumoto, T. Ikeguchi, and S. Neda, in Reference [34], pp. 347–354.
23. M. Andritschky, O. Gröbner, A. G. Mathewson, F. Schumann, P. Strubin, and R. Souchet, *Vacuum*, **38** (1988) 933.
24. J. C. Schuchman, J. Godel, W. Jordan, and T. Oversluizen, *J. Vac. Sci. Technol.*, **16** (1979) 720; also see J. C. Schuchman, *J. Vac. Sci. Technol.*, **A1** (1983) 196.
25. H. Ishimaru, *J. Vac. Sci. Technol.*, **A2** (1984) 1170.
26. Y. Morimoto, T. Shirakura, K. Konishi, S. Takahashi, S. Yokouchi, and S. H. Bee, in Reference [34], pp. 110–116.
27. W. Schurman, *Physica*, **36** (1967) 137.
28. F. M. Penning and K. Nienhuis, *Phillips Tech. Rev.*, **11** (1949) 116.
29. M. D. Malev and F. M. Trachtenberg, *Vacuum*, **23** (1973) 403; **25** (1975) 211.
30. H. Hartwig and J. Kouptsidis, *J. Vac. Sci. Technol.*, **11** (1974) 1154.
31. C. Benvenuti, *Nucl. Instr. Meth.*, **205** (1983) 391; C. Benvenuti and F. Francia, *J. Vac. Sci. Technol.*, **A8** (1990) 3864.
32. R. F. Niemann, R. Benaroya, M. Choi, R. J. Dortwest, G. A. Goeppner, et al., in Reference [34], pp. 84–101.
33. S. H. Be, S. Yokouchi, T. Nishidono, Y. Morimoto, K. Watanabe, et al., in Reference [3], pp. 102–109.
34. W. C. Turner, *J. Vac. Sci. Technol.*, **A14** (1996) 2026.

35. A. G. Mathewson, in *Proc. 1993 Particle Accelerator Conference* (Institute of Electrical and Electronics Engineers, Piscataway, NJ, 1993), pp. 3828–3832.
36. C. Benvenuti, R. S. Calder, and G. Passardi, *J. Vac. Sci. Technol.*, **13** (1976) 1172.
37. M. G. Rao, private communication.
38. M. F. Fedorova, *Sov. Phys. Tech. Phys.*, **8** (1963) 434.
39. D. Hollenbach and E. E. Salpeter, *J. Chem. Phys.*, **53** (1970) 79; *Astrophys. J.*, **163** (1971) 155.
40. J. C. Schuchman, Reference [34], pp. 300–312.
41. Y. Kojima, in *Proc. 1991 Particle Accelerator Conference* (Institute of Electrical and Electronics Engineers, Piscataway, NJ, 1991) p. 16.
42. H. F. Dylla, L. R. Doolittle, and J. F. Benesch, in *Proc. 1993 Particle Accelerator Conference* (Institute of Electrical and Electronics Engineers, Piscataway, NJ 1993), pp. 748–752.
43. S. Mitsunoba et al., *Proc. of the 5th Workshop on RF Superconductivity*, DESY Rep. No. DESY M-92-01 (April 1992); p. 84.
44. A. Matheisen et al., in Reference [56], p. 44.
45. G. Cavallari, C. Benvenuti, P. Bernard, D. Bloess, E. Chiaveri, et al., in Reference [55], pp. 806–808.
46. D. Proch, in Reference [55], pp. 758–762.
47. H. F. Dylla, J. J. Bisognano, D. Douglas, C. W. Leemann, P. Liger, et al, in *CEBAF 1992 Summer Workshop*, AIP Conference Proc. No. 269 (American Institute of Physics, New York, 1993), pp. 80–84.
48. J. Mammosser, in *Proc. 6th Workshop on RF Superconductivity*, Newport News, VA 1993 (CEBAF, Newport News, VA, 1994).
49. R. H. Siemann, in Reference [55], pp. 532–536.
50. M. Ross, in Reference [55], pp. 522–526.
51. J. D. Lawson, *Proc. Phys. Soc.*, **B70** (1957) 2.
52. R. J. Hawryluk et al. in *Proc. 11th Intern. Conf. on Plasma Physics and Controlled Nuclear Fusion Research, Kyoto November 1986*, Paper No. IAEA-CN-47/A-I-3 (International Atomic Energy Agency, Vienna, 1987).
53. E. Bertolini, in *Proc. 11th Symposium on Fusion Engineering* (Institute of Electrical and Electronics Engineers, Piscataway, NJ, 1986).
54. A. S. Bishop, *Project Sherwood—The U. S. Program in Controlled Fusion* (Addison-Wesley, Reading, MA, 1958).
55. B. Rose, A. E. Taylor, and E. Wood, *Nature*, June 14, 1958, pp. 1630–1632.
56. Lyman Spitzer et al., *Problems of the Stellarator as a Useful Power Source* (unpublished PPPL report, June 1954).
57. J. L. Bromberg, *Fusion: Science, Politics and the Invention of a New Energy Source* (MIT Press, Cambridge, MA, 1982), pp. 130–150.
58. L. A. Artsimovich et al. *Experiments in Tokamak Devices*, Nuclear Fusion, Special Supplement 1969 (International Atomic Energy Agency, Vienna, 1969).
59. N. J. Peacock, D. C. Robinson, N. M. J. Forest, P. D. Wilcock, and V. V. Sannikov, *Nature*, November 1, 1969, pp. 488–490.
60. G. G. Kelley, M. Roberts, and J. F. Clarke, *Conceptual Design of the Ormak Facility*, Oak Ridge National Laboratory Report ORNL-TM-2821 (unpublished, December 3, 1969).
61. B. Coppie and J. Rem, *Sci. Am.*, **227** (1972) 65.
62. R. L. Hirsch, *Doublet II CTR Experiment at Gulf General Atomic* (unpublished, January 21, 1970).
63. H. P. Furth, *Nucl. Fusion*, **15** (1975) 487.
64. H. Eubank et al., in *Proc. 7th Intern. Conf. on Plasma Physics and Controlled Fusion Research*, Innsbruch, 1978 (International Atomic Energy Agency, Vienna, 1979), p. 11.
65. R. A. Ellis, H. P. Eubank, R. Goldston, R. R. Smith, and T. Nagashima, *Nucl. Fusion*, **6** (1976) 524.
66. P. E. Stott, C. C. Daughney, and R. A. Ellis, *Nucl. Fusion*, **15** (1975) 431.
67. R. R. Parker, D. S. Pappas, and E. Apgar, *Bull. Am. Phys. Soc.*, **21** (1976) 1141.

68. L. Oren and R. J. Taylor, *Nucl. Fusion*, **17** (1977) 1143.
69. E. B. Meservey, V. Arunasalam, C. Barnes, K. Bol, N. Bretz, et al., *J. Nucl. Mater.*, **93/94** (1980) 267.
70. W. B. Gauster, J. A. Koski, and R. D. Watson, *J. Nucl. Mater.*, **122/123** (1984) 80.
71. R. A. Langley, R. J. Colchin, R. C. Colchin, R. C. Isler, M. Murakami, J. E. Simpkins, et al., *J. Nucl. Mater.*, **85/86** (1979) 215.
72. P. K. Mioduszewski, J. E. Simpkins, P. H. Edmonds, R. C. Isler, E. A. Lazarus, et al., *J. Nucl. Mater.*, **128/129** (1984) 884.
73. P. K. Mioduszewski, R. E. Clausing, and L. Heatherly, *J. Nucl. Mater.*, **85/86** (1979) 963.
74. M. A. Mahdavi, C. J. Armentrout, F. P. Blau, G. Bramson, N. H. Brooks, et al., *J. Nucl. Mater.*, **111/112** (1982) 355.
75. D. M. Meade et al., in *Proc. Eight Intern. Conf. on Plasma Physics and Controlled Fusion Research* (International Atomic Energy Agency, Vienna, 1981) Vol. 1, p. 665.
76. M. Keilhacker et al., in Reference [88], Vol. 2, p. 351.
77. J. L. Cecchi, R. J. Knize, H. F. Dylla, R. J. Fonck, D. K. Owens, and J. J. Sredniawski, *J. Nucl. Mater.*, **111/112** (1982) 305.
78. R. J. Fonck, M. Bell, K. Bol, K. Brau, R. Budny, et al., *J. Nucl. Mater.*, **111/112** (1982) 343.
79. H. F. Dylla, S. A. Cohen, S. M. Rossnagel, G. M. McCracken, and Ph. Staib, *J. Vac. Sci. Technol.*, **17** (1980) 286; H. F. Dylla, *J. Vac. Sci. Technol.*, **A6** (1988) 1276.
80. W. Poschenrieder, G. Standenmaier, and Ph. Staib, *J. Nucl. Mater.*, **93/94** (1980) 322.
81. JT-60 Team, *Plasma Phys. Contr. Fusion*, **28** (1986) 185.
82. P. H. LaMarche, H. F. Dylla, G. D. Martin, and S. Raftopoulos, in *Proc. 13th Symposium on Fusion Engineering, Knoxville*, 1989 (Institute of Electrical and Electronics Engineers, Piscataway, NJ, 1990), p. 963.
83. The JET Team, *J. Nucl. Mater.*, **196/198** (1992) 3.
84. E. E. Bloom, R. W. Conn, J. W. Davis, R. E. Gold, R. Little, et al., *J. Nucl. Mater.*, **122/123** (1984) 17.
85. H. Horiike et al., *Fusion Eng. Des.*, **16** (1991) 285.
86. The JET Team, *Nucl. Fusion*, **32** (1992) 187.
87. P. Andrew, C. J. Caldwell-Nichols, J. P. Coad, K. J. Dietz, J. Ehrenberg, et al., *J. Nucl. Mater.*, **196/198** (1992) 143.
88. R. J. Hawryluk et al., Princeton Plasma Physics Laboratory Report No. PPPL-2977 (Princeton Plasma Physics Laboratory, Princeton, NJ, February 1994).
89. J. D. Strachan et al., *Phys. Rev. Lett.*, **58** (1987) 1004.
90. The JET Team, *J. Nucl. Mater.*, **176/177** (1990) 3.
91. S. L. Milora, *J. Vac. Sci. Technol.*, **A7** (1989) 925.
92. H. Ninomiya and the JT-60 Team, *Phys. Fluids*, **B4** (1992) 2070.
93. D. N. Hill et al., in *Proc. 13th Intern Conf. on Plasma Physics and Controlled Fusion Research, Washington, DC 1990*, Vol. 3 (International Atomic Energy Agency, Vienna, 1990), p. 487.
94. J. L. Cecchi, *J. Nucl. Mater.*, **93/94** (1980) 28.
95. H. F. Dylla and the TFTR Team, *J. Nucl. Mater.*, **145/147** (1987) 48.
96. H. F. Dylla et al., *Nucl. Fusion*, **27** (1987) 1221.
97. H. F. Dylla and K. L. Wilson, eds., *Tritium Retention in TFTR*, PPPL Report No. 2523, and Sandia National Laboratories Report No. SAND-88-8212 (April 1988).
98. A. Miyahara and T. Tanake, *J. Nucl. Mater.*, **155/157** (1988) 49.
99. J. Winter, *J. Nucl. Mater.*, **176/177** (1990) 14.
100. W. G. Redan, *J. Vac. Sci. Technol.*, **20** (1982) 1173.
101. R. Frank and E. Usselmann, *Vakuum Technik*, **25** (1976) 48.
102. L. E. Valby, in *Proc. of the 7th Symposium on the Engineering Problems of Fusion Research, Knoxville 1977* (Institute of Electrical and Electronics Engineers, Piscataway, NJ, 1977), p. 1040.

Index

A

Ablation mechanism, pulsed laser deposition, 698–699
Abrasion resistance, 786
Abrasives, for cleaning, 570
Absorption, and construction materials, 447
Abundance sensitivity, mass analysis, 299
AC tungsten inert gas welding, 547
Acceptance ellipse, 316–317
Acetone, 571
Active storage, 591
Actuation of valves, 400–401
  manual, 401
  remote, 401
  sealing of actuating motion, 400
Adhesion, loss of, causes, 786
Adsorption, 40–41
  and construction materials, 447
  cryogenic pumps, 160–162
  heat of, 42
  sticking coefficient, 43–44
  surface adsorption isotherms, 45–47
Air firing, for cleaning, 576
Air pollution measurement, mass analysis, 291–292
Alconox, 520
Alkaline cleaners, 572–574
Almeco 18, 520
Alodyne, 454–455
Alpha ionization, 307–308
Alternating Gradient Synchrontron (AGS), 478
Aluminum
  chamber wall thickness requirements, 521, 533–535
  chemical finishes, 520
  cleaning, 569
  in construction of vacuum chamber, 453–455
  corrosion, 538–541
  cylinders, 533–534
  demountable seals, 512–518
  flat plates, 533
  joint design, 546–547
  mechanical finishes, 518–519
  and outgassing, 509–512
  peak stresses/strains, 535
  protective coatings, 541
  radiated heat, absorption of, 536–537
  relative thickness, 534–535
  repairing defects, 548
  as seal metal, 481
  thermal gradients, 536
  thermal time constant, 537–538
  weight/thermal expansion, 535
  welding, guidelines for, 541–548
Aluminum alloy
  chemical composition limits, 531–532
  mechanical properties of, 525
  and outgassing, 486
  polishing, 559
  types/designations, 521–525
Alutone, 520
Amchem, 520
Amklene, 520
Amplified sputtering, 626
Anhydrous alcohol, 585
Anisotropic etching, 635–638
Anodized aluminum, cleaning/sealing of, 560
Apiezon, 451
Appendage pump, 221

Argon, 518–519
  and diode pump, 223
Argon ion beam etching, 682
ASA (ANSI) flanges
  components with, 431
  features of, 421
Aspect ratio dependent etching, 639, 650
Atmospheric pressure, pumpdown from, 494
Atomic desorption, 53–54
Auger electron spectroscopy, 735, 737–741
Auger process, 735
Auger spectrum, 735
Auger transition, 734
Automotive industry, sputter technology in, 624
Auxiliary pumps, diaphragm pumps as, 87–88
Avogadro's law, 9–10
Avogadro's number, 10, 41
Axial sources, ion generation, 306

B

Backing, 116
Backing pumps
  diaphragm pumps as, 87–88
  molecular drag pumps, 202
  and tolerable forepressure, 128–129
  turbomolecular pumps, 202
Backstreaming
  accidental backstreaming, 140–142
  and contamination, 592–593
  and fluid films, 143
  mechanical pump effects, 142
  minimizing, 593
  and power failure, 593
  primary/secondary, 137–138
  and pump speed, 138–139
  and pumping fluid loss, 142–143
  rate versus pressure, 140–142
  surface migration, 140
  vapor jet pumps, 137–143
Baffles, vapor jet pumps, 132–133
Bakeout
  and outgassing, 507
  seal disturbance during, 516
  time frame for, 597
Bandgap engineering, 696
Barrel reactors, plasma etching, 644–645
Batch inlet, 302
Bayard-Alpert hot cathode ionization gauges, 273–276
Beaded-edge face-sealed flanges, 411–412
Beam stacking method, 791
Bearings
  and feedthrough, 437
  impact on process, 437
  lifetime, equation for, 436
Bellows, of valves, 400
Beryllium oxide, 458
  as thermal conductor, 457–458
Bias sputtering, 623–624
Biochemical analysis, mass analysis, 292
Blowers, vacuum blowers, 97–115
Blowoff, particulate removal, 569–570
Boat materials, roll-to-roll vacuum coating, 769–770
Bohm sheath criteria, 632–633
Bonnet seals, 402
Boron nitride, 458, 563
Bourdon gauges, 263
Boyle's law, 9, 365
  kinetic theory derivation, 14
Branauer-Emmett-Teller (BET), 44, 46, 48, 49
Brass, in vacuum materials, 461
Brinell hardness, 513
Broad beam ion source, 672–673
Butterfly valves, 397, 398, 407
Butyl rubber, as seal material, 472
Bypass ranging, mass flow controller, 379–380
Bypass valves, vacuum blowers, 109–110

C

Cadmium, avoiding, 461
Calibration
  mass spectrometers, 364–369
  of vacuum gauges, 286–287
Camphor, 446
Capacitance gauges, 778
Capillary action, 47–48
Capture pumps, 60
Carbon, 563
Carbon dioxide snow, 570
Cathodic arc deposition, 774
Ceramic bearings, turbomolecular pumps, 192
Ceramics
  cleaning, 569
  as electrical insulator, 457
  microcracks, reducing, 561
  outgassing, 501–504
  selected properties of, 449

Chain blower systems, vacuum blowers, 112–113
Channel electron multipliers, 324–325
Charge exchange, ion generation process, 350–351
Charles's law, 9
  kinetic theory derivation, 14–15
Chassis ground, mass flow controller, 384
Chelating agents, for cleaning, 575
Chemical depth profiling, electron spectroscopies, 740
Chemical etching
  chemical etchants, types of, 564
  for cleaning surfaces, 570–571
  for stripping, 563
  wet and dry, 571
Chemical finishes, aluminum, 520
Chemical ionization, 308
Chemical kinetics analyzer, 373
Chemical memory, 726–727
Chemical polishing, stainless steel, 557
Chemical shift, electron spectroscopies, 736–737
Chemical vapor deposition, 624
  mass analysis, 372
  roll-to-roll vacuum coating, 775
Chemically assisted ion beam etching, 682
Chemically unstable materials, 461
Chemisorption, 40, 42, 447
Chevron configuration, 325
Chlorinated hydrocarbon nonpolar solvents, 571–572
Chlorofluorocarbon solvents, 572
Circular tube, viscous gas flow, 35–36
Citrinox, 520
Classical scattering, meaning of, 752
Clausing factors, conductance, 38–39
Clausius-Clapeyron plot, 157
Cleaning, 553, 567–601, 598
  abrasives, 570
  aluminum, 569
  and assembly process, 587
  ceramics, 569
  chelating agents, 575
  chemical etching, 570–571
  conditioning, 596–598
  contaminant films, 570–577
  copper, 569
  detergent cleaning, 574
  and documentation, 599–601
  drying, 584–585
  external cleaning, 567–586
  gases for, 576–577
  glass, 569
  and handling, 589–590
  in situ cleaning, 591–596
  liquids for, 575–576
  mechanical processes in, 581
  monitoring of, 585–586
  particulate removal, 569–570
  pH adjusters, 575
  piranah solution, 575
  plasma cleaning, 571, 577
  polymers, 569
  and processing environment, 587–589
  and process-related contamination, 596
  reactive cleaning, 575–578
  rinsing, 578–580
  safety factors, 586
  saponifiers for, 572–574
  soaking, 580
  solvents for, 571–572
  spraying, 581
  stainless steel, 568–569
  and storage/packing, 590–591
  and surface tension, 574
  and system-related contamination, 592–596
  ultrasonic cleaning, 582–584
  vapor cleaning, 581–582
  See also Polishing
Cleanrooms, 589
Clothing, for contamination control, 588
Cluster tools, plasma etching tools, 666
Coarse pumps, 60
Coaxial feedthrough, 442
Coefficients
  of diffusion, 21
  of thermal conductivity, 21
Cold-bore machines, 798–800
  development of, 798–800
Cold-caps, vapor jet pumps, 146
Cold-cathode ionization gauges, 276–278
Colossal magneto resistance materials, 706–707
Combat, 563
Compression ratio
  diaphragm pumps, 87
  vacuum blowers, 101–102, 107–108
Compressors, cryogenic pumps, 153–155
Concrete, as vacuum material, 460

Condensation, 48–49
cryogenic pumps, 159–160
meaning of, 48–49
Conditioning, 596–598
to avoid splashing, 703
desorption, 596–597
sputter cleaning, 598
Conductance
clausing factors, 38–39
conductances in parallel, 33
conductances in series, 33–34
elements of, 32–33
Knudsen's formulation, 37–38
vacuum valves, 402–403
ConFlat flanges, 418–421, 430, 474–476
aluminum, 514–515
Constant speed case, 66–67
Constant temperature method, 266
Constant throughput case, 67
Constrained flanges, 516–517
Construction materials
and absorption, 447
and adsorption, 447
ceramics, selected properties of, 449
and diffusion, 447
for electrical conductors, 455–456
for electrical insulators, 457–458
and evaporation, 446–447
lubricants, 458–460
materials to avoid, 460–461
metals, selected properties of, 450
organic materials, 460
and outgassing, 447
and permeation, 449
polymeric materials, selected properties of, 448
and solubility, 447, 449
for thermal conductors, 457–458
for thermal insulators, 457–458
vacuum chamber materials, 451–455
Contaminant films, cleaning, 570–577
Contamination
versus clean surface, 567
and cleanrooms, 589
contaminants, types of, 568, 588, 596
from desorption, 596
oil contamination, 594–595
and plasma-enhanced chemical vapor deposition, 718
and plasma-etching, 640–641
and processing environment, 587–588
process-related contamination, 596
and stainless steel, 568–569
system-related contamination, 592–596
Continuous coating, advantages of, 784
Continuous Electron Beam Accelerator Facility, 800
Control valves, types of, 397–398
Controller error, mass flow controller, 381–382
Convection gauges, 269–270
Convectron, 269
Cool gas feedback blowers, vacuum blowers, 111
Copper
cleaning, 569
in construction of electrical conductors, 455–456
polishing, 560–561
as seal metal, 481
Corrosion
aluminum, 538–541
reducing rate of, 785–786
thin films, 785–786
Corrosive gases, turbomolecular pumps, 203–204
Cracks
avoiding and welding, 542–546
cracking patterns, 344
Crisco, 461
Critical dimension (CD control), 639
Critical fluids, 572
Critical temperature, 10
Cross-beam sources, ion generation, 306
Crossover, cryogenic pumps, 163–164
Crossover rating, 164
Crossplume configuration, to avoid splashing, 704
Crucible materials, roll-to-roll vacuum coating, 769–770
Crucible rotation, Czochralski silicon crystal puller, 441–442
Cryogenic pumps, 149–181
adsorption, 160–162
automated regeneration, 169–171
benefits of, 179–180
capacity of, 164–165
chamber evacuation, 173–174
combined with turbomolecular pumps, 205–207
components of, 149–151

compressors, 153–155
condensation, 159–160
crossover, 163–164
development of, 151–152
8-inch pump, characteristics of, 156
electrical heaters, 168–169
evaporation systems, 174–175
flow regimes, 156–157
future view of, 181
ion implantation, 177
limitations of, 180
partial regeneration, 171–172
pressure sensors, 168
refrigerator operation, 152–153
regeneration, 165–167
sputtering systems, 176–177
temperature indicators, 167–168
thermal radiation, 162
for ultra-high vacuum, 217–218
used in combination, 776–777
uses of, 155–156
vapor pressure, 157–159
water-vapor specific pumps, 177–178
Cryotraps, 776
C-Stellarator fusion device, 792
Cubic zirconia, as thermal insulator, 458
Cylinders, aluminum, 533–534
Cylindrical mirror analyzer, 741–742
Cyropumps, 776–777
Czochralski silicon crystal puller, crucible rotation, 441–442

D

Dalton's law of partial pressure, 10
Dark space, 630
DC bias, plasma etching, 648–649
DC diodes, sputter deposition, 613–614
DC ion source, 674–675
DC tungsten inert gas welding, 547
Decorative products, roll-to-roll vacuum coating, 764
Degree of vacuum, 4
Deionized water, for rinsing, 578–579
Delta seal, 478
Demountable seals, 512–518
CF flanges, 514–515
constrained flanges, 516–517
diamond-shaped seals, 515
materials, 516
metal C-rings, 515
seal disturbance during bakeout, 516
seal effectiveness, criteria for, 513
unconstrained flanges, 518
wire seals, 515
Densification, 688
Deposit buildup, removal of, 563
Deposit diamondlike carbon, 689–690
Deposition
cathodic arc deposition, 774
chemical vapor deposition, 775
ion-beam-assisted deposition, 687–689
ion beam direct deposition, 689–690
ion beam sputter deposition, 683–687
and magnetron sputtering, 618–620
organic materials, 775
pulsed laser deposition, 694–707
sputter deposition, 611–624
Desiccants, types of, 591
Desorption, 49–52, 596–597
electron-and ion-stimulated desorption, 53–54, 597–598
photodesorption, 598
thermal desorption, 50–52, 596–597
ultrasonic desorption, 53, 598
Detergent cleaning, 574
Diamond, as thermal conductor/electrical insulator, 457
Diamond-shaped seals, 515
Diamondlike carbon, and pulsed laser deposition, 706
Diaphragm gauges, 263–264
Diaphragm pumps, 84–96
as backing/auxiliary pumps, 93–95
multistage pumps, 86
operating principles, 84, 86–87
parts of, 84–85
performance factors, 91–92
pumping speeds, 87–88
single-stage pumps, 85–86
standalone operation, 92–93
state-of-the-art design, 87–91
uses of, 84
Diaphragm valves, 393, 396, 407
Diffusion
coefficient of, 21
and construction materials, 447
to harden surfaces, 565–566
and nonmetallic materials, 490
outgassing by, 497–500

Diffusion pumps, 776
for ultra-high vacuum, 217
See also Vapor jet pumps
Diffusion sources, plasma etching, 663–664
Diode sputtering
DC diodes, 613–614
RF diodes, 614–615
Direct reading gauges, 260–264
Discrete dynode multipliers, 324
Dispersion forces, 41
Documentation
laboratory/engineering (L/E) notebook, 599–600
manufacturing processing instructions, 600
process specifications, 600
and surface preparation, 599–601
Dremel tool, 548
Dry brushing, 569
Dry lubricants, 459
Dry pumps, 235
Drying, in cleaning process, 584–585
Dynamic process, calibration of gauges, 287
Dynamic seals, 463
Dynamic secondary ion mass spectrometry, 748–749

E

Efficiency, blower pumping, 101
Ejectors, 120
Elastomer seals, 472–474
See also Polymer seals
Elastomers, and outgassing, 488–489, 497–501
Electric light sources, mass analysis, 297–298
Electric quadrupole analyzers, 309–319
Electrical conductors, materials for, 455–456
Electrical insulators, materials for, 457–458
Electrical products, roll-to-roll vacuum coating, 764
Electro-optical products, roll-to-roll vacuum coating, 765
Electron beam heating, for evaporation, 768–769
Electron bombardment heating, sublimation from, 247
Electron cyclotron resonance sources, plasma etching tools, 662–664
Electron energy distribution function, 630
Electron impact, ion generation, 303–307
Electron impact emission spectroscopy, uses of, 296
Electron inelastic mean free path, 735
Electron spectroscopies, 732–745
Auger electron spectroscopy, 735, 737–741
chemical depth profiling, 740
chemical shift, 736–737
cylindrical mirror analyzer, 741–742
elemental depth profiling, 738–740
elemental mapping, 741
hemispherical sector analyzer, 742–744
operational factors, 733–734
qualitative analysis, 734–735
quantitative analysis, 737–738
surface sensitivity mechanism, 735–736
X-ray photoelectron spectroscopy, 732, 733, 736, 737–740
X-ray sources, 744–745
Electron stimulated desorption, 53–54, 597–598
ion generation process, 348–350
Electronic device fabrication, mass analysis, 296
Electropolishing
aluminum alloys, 559
stainless steel, 556–558
Electrostatic chucks, plasma etching tools, 666
Elemental depth profiling, electron spectroscopies, 738–740
Elemental mapping, electron spectroscopies, 741
Ellipsometry, plasma diagnostic, 641
End point detection, plasma etching, 666
Etching. *See* specific etching methods
Evacuation, 22–23
cryogenic pumps, 173–174
evacuation rate, 24
process pumping, 74–75
Evaporation
basics in process of, 767–768
and construction materials, 446–447
cyrogenic pumps, 174–175
heating as source of, 768–769
mass analysis, 371
roll-to-roll vacuum coating, 766–769
Exfoliation, and splashing, 702

F

Faraday collectors, 323–324
Feedthroughs
magnetic fluid feedthroughs, 433–443
materials used, 439

Ferrite magnets, 229
Fickian kinetics, 48
Fick's laws, 50, 51
Field ionization, 308
Filament materials, ion generation, 304
Fillet joints, 545
Filter geometry, 318
Fine pumps, 60
Fingerprint spectra method, 750
Flanges
  ASA (ANSI) flanges, 421
  beaded-edge face-sealed flanges, 411–412
  comparison of systems, 411
  components with ASA (ANSI) type flanges, 431
  components with ISO-type flanges, 427–430
  components with knife-edge-sealed flanges, 430
  constrained flanges, 516–517
  JIS flanges, 421–423
  knife-edge-sealed flanges, 418–421
  O-ring-sealed flanges, 412–417
  rectangular flanges, 425
  selecting system, 410
  unconstrained flanges, 518
  vacuum valves, 402–403
  wire-sealed flanges, 423–424
Flat plates, aluminum, 533
Flourel, 465
Fomblin, 474
Forepumping, 116
Forward recoil scattering, 753
Fractionation patterns, 344
Fragmentation, ion generation process, 344–345
Fragmentation analysis, gas identification, 358–364
Free molecule flow, 29
Free outgassing rate, 485
Freon, 585
Fusion. *See* Magnetic fusion

G

Gas analysis, residual gas, 135–136
Gas chromatograph-mass spectrometry, 374
Gas composition/purity analysis, mass analysis, 292–293
Gas correction factor, mass flow controller, 385
Gas flow, 26
  description of, 26–27
  gas flow, viscous gas flow, 26
  gas flow rate, 29–30
  intermediate gas flow, 26
  molecular gas flow, 26, 29, 37
  nature of, 25–26
  throughput Q, 30
  turbulent gas flow, 26, 27–28
  viscous gas flow, 28, 35–36
  volumetric flow rate, 30
Gas identification
  fragmentation analysis, 358–364
  isotope analysis, 355–358
Gas laws
  derivation of, 14–15
  ideal gas law, 14–15
  listing of, 9–10
Gas rarefaction, and magnetron sputtering, 621
Gas release, from surfaces, 54
Gas secondary ion mass spectrometry, 308
Gas stick, mass flow controller, 383
Gas temperature rise, calculation of, 106–107
Gas throughput, 23
Gas transport, throughput, 61–62
Gases
  characteristics of, 8–9
  for cleaning, 576–577
  nonideal gases, 10
  in vacuum system, 337
Gaskets, 412
  and outgassing, 506–507
Gate valves, 393, 395, 406, 407
Gauges
  Bayard-Alpert hot cathode ionization gauges, 273–276
  bourdon gauges, 263
  calibration of, 286–287
  capacitance gauges, 778
  cold-cathode ionization gauges, 276–278
  convection gauges, 269–270
  diaphragm gauges, 263–264
  direct reading gauges, 260–264
  hot-cathode ionization gauges, 270–276
  indirect reading gauges, 265–282
  ion gauges, 778
  liquid-wall gauges, 260–263
  molecular drag gauges, 279–282
  Pirani gauges, 265–266
  resonance gauges, 278–279
  thermal conduction gauges, 778
  thermal conductivity gauges, 265–270

Gauges (*continued*)
thermistor gauges, 268
thermocouple gauges, 266–268
triode hot cathode ionization gauges, 271–273
General gas law, 10
Getter pumps, 220, 242–252
getter-ion pumps, 75–76
nonevaporable getter pumps, 220, 247–252
sublimation pumps, 242–247
Gifford-McMahon cycle, 152
Glass
cleaning, 569
in construction of vacuum chamber, 451–452
glass coating, sputter technology in, 624–625
microcracks, reducing, 561
viewports, 444
Glass industry, mass analysis, 297
Gloves, materials for, 590
Glow bar, 598
Glyptal, 451
Gram molecular volume, 9
Graphite, 675
Greases, as lubricant, 458–459
Grid/open/RGA/UHV sources, ion generation, 306
Gridless sources, ion sources, 675
Grit blasting, 563

H

Handling, and cleaning surfaces, 589–590
Hardening of surfaces, 565–566
Heat of adsorption, 42
Helical inductive couplers, plasma etching tool, 659–660
Helical resonators, plasma etching tool, 660–661
Helicoflex seal, 477–478, 482
Hemispherical sector analyzer, 742–744
Henkel/Parker Amchem P3 Detergent, 520
Hi-potting, 240
High-curie point permanent magnet, 675
High-pressure diodes, plasma etching, 645–646
High vacuum, 6–7, 71–72, 78–79
sputtering, 609–626
High-vapor-pressure materials, 460–461
Holding pump, 80
Hollow shaft seals, 441
Hollow-anode-enhanced triodes, plasma etching tool, 654–656
Horizontal introduction system, 756
Hot-cathode ionization gauges, 270–276
Hydrogen, 447, 449, 456
Hydrogen peroxide, for cleaning, 576
Hydrogen plasmas, for cleaning, 577
Hypertests, 785
HYSURF, 568

I

Ice cubes, 446
Ideal gas law, 14–15
Impedance measures, plasma diagnostic, 641
In situ cleaning, 591–596
Indirect reading gauges, 265–282
Indium, as seal metal, 481
Induced DC bias, plasma etching tools, 648–649
Inductive heating, for evaporation, 768
Inductively coupled plasmas, 658–662
Inlet systems, 300–303
batch inlet, 302
membrane inlet, 303
modulated inlet, 303
pressure reduction, 302
and range of pressure, 301
tee element, 302–303
Inlet throttling systems, vacuum blowers, 109
Intelligent gauge tubes, 286
Intermediate gas flow, 26
International Tokamak Engineering Reactor (ITER), 790, 810
Intersecting Storage Rings, 793, 794, 798
Intrinsic resputtering, 612
Ion analyzers, 308–323
electric quadrupole analyzers, 309–319
magnetic sector analyzers, 319–322
time-of-flight analyzers, 322–323
Ion beam assisted deposition
cyrogenic pumps, 175
process of, 687–689
Ion beam direct deposition, process of, 689–690
Ion beam etching, 678–683
argon ion beam etching, 682
chemically assisted ion beam etching, 682
cooling of substrate, 679–680
etch rate, 679

ion beam etching, 682–683
reactive ion beam etching, 682
Ion beam neutralization, 677–678
methods of, 677–678
Ion beam sputter deposition, 683–687
Ion beam technology, 672–691
broad beam ion source, 672–673
DC ion source, 674–675
development of, 672–673
gridless sources, 675
ion beam assisted deposition, 687–689
ion beam direct deposition, 689–690
ion beam etching, 678–683
ion beam neutralization, 677–678
ion beam sputter deposition, 683–687
Kaufman ion source, 673
RF ion source, 674
Ion bombardment-enhanced etching, 636
Ion detection, 323–325
channel electron multipliers, 324–325
discrete dynode multipliers, 324
Faraday collectors, 323–324
microchannel plate multipliers, 325
Ion detector, partial pressure measurement, 354–355
Ion gauges, 778
Ion generation, 303–308
alpha ionization, 307–308
axial sources, 306
chemical ionization, 308
cross-beam sources, 306
electron impact, 303–307
field ionization, 308
filament materials, 304
grid/open/RGA/UHV sources, 306
ion source configurations, 306
lathane hexaboride-coated base metals, 305–306
magnetic sector instruments, 307
photo ionization, 307
quadrupole instruments, 307
rhenium, 305
special construction ion sources, 307
thoria-coated cathodes, 305
thoriated tungsten, 305
tungsten, 304–305, 346–347
Ion generation process, 340–351
charge exchange mechanisms, 350–351
electron stimulated desorption, 348–350
fragmentation, 344–345
instrument outgassing, 345–346
ionization efficiency, 342
surface chemistry, 346–347
Ion implantation, cryogenic pumps, 177
Ion peening, 623, 688
Ion scrubbing, 597–598
Ion source configurations, ion generation, 306
Ionized sputter deposition, 626
Iris/blade valves, 407
ISO-K flanges, 412, 414
ISO-KF flanges, 412, 415, 417, 427, 428
ISO-MF flanges, 412, 414, 415, 417, 427
ISO-O-ring-sealed flanges
components with, 427–430
features of, 412–417
ISO-Universal flanges, 414
Isotherms, 42
Langmuir isotherm, 46
surface adsorption isotherms, 45–47
Isotope analysis, gas identification, 355–358
Isotope ratio analysis, mass analysis, 292

J

JIS flanges, 421–423
Joint design, aluminum, 546–547

K

Kaufman ion source, 673
Kinetic theory
basic assumptions in, 11
diffusion, 21
gas laws, derivation of, 14–15
molecular motion, 12–13
thermal conductivity, 21
Knife-edge-sealed flanges, 418–421
components with, 430
Knock on effects, sputtering, 609
Knudsen number, 26–27
Knudsen's formulation, conductance, 37–38

L

Laboratory/engineering (L/E) notebook, 599–600
Langmuir isotherm, 46
Langmuir probe, plasma diagnostic, 642–643
Large Electron Positron storage ring, 798
Large Hadron Collider, 790, 798

Large-scale processes
roll-to-roll vacuum coating, 761–787
ultra-high-vacuum technology, 789–810
Laser absorption spectroscopy, plasma diagnostic, 642
Laser induced fluorescence, plasma diagnostic, 643
Laser reflection interferometry, plasma diagnostic, 641
Lathane hexaboride-coated base metals, ion generation, 305–306
Lawson criteria, magnetic fusion, 802
Lead
avoiding, 461
as seal metal, 481
Lead-tin solder, in vacuum materials, 461
Leak/needle valves, 311, 407
Leaks
detection, mass analysis, 293, 370–371
leak rates, 23
and pumpdown time, 67–68
Light-duty feedthroughs, 440
Linearity, mass analysis, 300
Liquid column manometer, 261–262
Liquid nitrogen, fluid films on LN traps, 143
Liquid seal, 481
Liquid sealed blowers, 112
Liquid target, to avoid splashing, 705
Liquid wall gauges, 260–263
Loschmidt's number, 10
Loss electron, 735
Low-carbon steel
cleaning/polishing, 558
porosity of, 558
Low-pressure etching, 636, 646–647, 650–651
Low-vacuum, 5
Low-vapor pressure materials, 446–447
Lubricants, 458–460
criteria in selection of, 458
dry lubricants, 459
metal-on-metal couples, 459–460
oils/greases, 458–459
uses in vacuum systems, 458
Lucite, 460

M

Macor, 451
Magnesium perchlorate, 591
Magnetic bearings, turbomolecular pumps, 193–194
Magnetic fields, and turbomolecular pumps, 203
Magnetic fluid feedthroughs
characteristics of, 437, 439
coaxial feedthrough, 442
comparison to other feedthroughs, 442–443
crucible rotation/Czochralski silicon crystal puller, 441–442
hollow shaft seals, 441
impact on process, 436–438
light-duty feedthroughs, 440
operating region guidelines, 435
uses of, 434–436
Magnetic fluids, and process environment, 436–437
Magnetic fusion, 801–810
future view, 810
historical view, 803–804
Lawson criteria, 802
Model C, 804–805
tokamaks, 802, 805–809
Magnetic sector analyzers, 319–322
Magnetic sector instruments, ion generation, 307
Magnetically enhanced reactive ion etcher (MERIE), 652
Magnetically enhanced triodes, plasma etching tool, 656–658
Magnetron sputtering, 616–623, 624–625
and deposition, 618–620
and gas rarefaction, 621
plasma etching, 652–653
process of, 616–618
reactive sputtering, 621–623
uses of, 624–625
Manufacturing processing instructions, 600
Martin, 454
Mass and energy sampling, plasma diagnostic, 642
Mass filter, partial pressure measurement, 351, 354
Mass flow controller, 378–387
bypass ranging, 379–380
control system design, 381–382
control valve design, 380–381
controller error, 381–382
electrical requirements, 383–384
gas line components, 383

gas type effects, 385
installation of, 382–383
orientation effects, 385
sensing techniques, 378–379
temperature effects, 384
troubleshooting, 386–387
Mass flow measurement, 376–387
mass flow controller, 378–387
measurement principles, 377–378
units of mass flow, 376–377
Mass flow rate, 30
Mass number, 290
Mass spectrometer calibration, 364–369
orifice flow technique, 367–369
pressure rate-of-rise technique, 366–367
static volume expansion, 365
transfer standards, 370
Mass spectrometers, 282, 290–303
chemical kinetics analyzer, 373
development of, 335–336
dipole type, 318–319
electron impact emission spectroscopy, 296
and expected gases in vacuum system, 336–340
gas chromatograph-mass spectrometry, 374
gas identification, 355–364
inlet systems, 300–303
ion generation process, 340–351
monopole type, 318
multipole type, 318
partial pressure measurement, 351–355
performance factors, 299–300
research applications, 372–374
secondary ion mass spectrometry, 308, 373
temperature desorption spectroscopy, 373
time-of-flight mass spectrometry, 299, 373–374
uses of, 291–299, 336, 370–372
Matching network, plasma etching tools, 648, 662, 664
Material Safety Data Sheets, 586
Maxwell-Boltzmann distribution, 12
Mean free path, 17, 19
meaning of, 4
Measurement
calibration of vacuum gauges, 286–287
gauges, 258–282
ion analyzers, 308–323
ion detection, 323–325
ion sources, 303–308
mass flow measurement, 376–387
mass spectrometers, 282, 290–303
vacuum-measuring instruments, 282–286
Mechanical finishes, aluminum, 518–519
Medical applications, mass analysis, 293
Medium-vacuum, 5
Membrane inlet, 303
MESA flanges, 425
Metal C-rings, 515
Metal films
and pulsed laser deposition, 706
sputter technology for, 625
Metal seals, 464, 474–482
historical view, 474–476
seal metals, 481
situations for use of, 464
and stored energy, 476–478
theoretical factors, 479–481
types of, 474–479
Metal-on-metal couples, 459–460
Metallurgical applications, mass analysis, 294
Metals
outgassing, 501–504
selected properties of, 450
Metglas, 461
Microchannel plate multipliers, 325
Microsharpening, 683
Microwave interferometry, plasma diagnostic, 642
Mirror effect, 650–651
Mirror machines, 803
Model C, fusion device, 804–805
Modular scaleup, 724–725
Modulated inlet, 303
Molar mass, 377
Mole, 9
Molecular beam epitaxy, 696
Molecular density, meaning of, 4
Molecular drag gauges, 279–282
Molecular drag pumps, 184, 195–197, 200
backing pumps, 202
combined with turbomolecular pumps, 197–200
design of, 195
trade names for, 200, 201
ultimate pressure, 195–196
Molecular gas flow, 26, 29
equations for, 37

Molecular motion, 12–13
impacts with chamber wall, 19
molecular mean free path, 17, 19
thermal transpiration, 20–21
time to form monolayer, 20
Molecular sieve, 234
Molybdenum, 456
Molybdenum disulphide, as seal metal, 481
Multi-fastener flanges, 414
Multiblade valves, 399
Multistage systems
diaphragm pumps, 86
vacuum blowers, 112–113

N

Near-threshold sputtering, 609
Needle valves, 311, 407
Net pumping speed, 65
Next Linear Collider project, 801
Nickel, 456
as seal metal, 481
Nitride films, and pulsed laser deposition, 707
Nonevaporable getter pumps, 220, 247–252
combined with sputter-ion pump, 252
self-contained pumps, 250–251
for ultra-high vacuum, 220
uses of, 247–249, 249–252
Nonideal gases, 10

O

Off-axis substrate configuration, to avoid splashing, 704–705
Oil contamination, 594–595
Oils, as lubricant, 458–459
Optical coatings, sputter technology for, 625
Optical emission spectroscopy, plasma diagnostic, 642
Optical products, roll-to-roll vacuum coating, 764
Optimum performance curve, 316
Organic materials, 460
deposition of, 775
and pulsed laser deposition, 707–708
Orifice flow technique, mass spectrometer calibration, 367–369
Orifice geometry, 64
Original equipment manufacture (OEM) support, plasma etching tools, 666–667
O-ring-sealed flanges, 412–417
Outgassing
and aluminum, 509–512
and aluminum alloy, 486
baked system for, 511–512
and bakeout temperatures/times, 507
and chamber dimensions, 493
and chamber pressure, 491–493
and construction materials, 447
and cyropump, 173
decreasing rate of, 506–507
by diffusion of absorbed gas to exposed surface, 497–501
and elastomers, 488–489
experimental data on, 501
free outgassing rate, 485
and gaskets, 506–507
ion generation process, 345–346
mass analysis, 294
and metals/ceramics, 501–504
and net pumping speed, 484–485
and nonmetallic materials, 490–491
outgassing rates, 23
and plastics, 488–489
and plastics/elastomers, 497–501
polymer seals, 468–471
polymers, 561–562
pressure versus time in, 495–497
and properties of materials, 485–491
and pumpdown from atmospheric pressure, 494
and pumpdown time, 68–69
pumpdown with, 495
rate during second pumpdown, 504–506
reported rates, 508
and stainless steel, 486
unbaked system for, 510–511
and water, 447
Oxide films, and pulsed laser deposition, 707
Oxygen-free, high-conductivity copper, in construction of electrical conductors, 455–456
Oxygen plasmas, for cleaning, 577
Ozone
for cleaning, 576–577
and O-ring cracking, 466

P

Packaging products, roll-to-roll vacuum coating, 764

Parallel, conductances in, 33
Parallel-plate triodes, plasma etching tool, 654
Partial pressure, 10
Partial pressure measurement, 290, 351–355
  ion detector, 354–355
  ion source, 351
  mass filter, 351, 354
Particle filter, to avoid splashing, 703–704
Particulates
  defects, avoiding and welding, 542
  filtration from environment, 588–589
  and plasma etching, 640–641
  process-related contamination, 596
  removal, methods of, 569–570
Pascal, 3
Passive storage, 591
Penning discharge, 222
Penning gauge, 278
Permeability, 23
Permeation
  and construction materials, 449
  and nonmetallic materials, 490
  polymer seals, 470–471
Perovskites, and pulsed laser deposition, 706–707
Phillips ionization gauge, 278
Phosphorous, in vacuum materials, 461
Phosphorous pentoxide, 591
Photo ionization, 307
Photoactivation, 52–53
Photodesorption, 598
Photon-stimulated desorption, 53
pH adjusters, for cleaning, 575
Picture framing, 726
Pinhole flaking, 568
Pipe flow geometry, 64
Piranah solution, 575
Pirani gauges, 265–266
Plasma cleaning, 571, 577
Plasma diagnostics
  ellipsometry, 641
  impedance measures, 641
  in situ rate measurements, 641
  Langmuir probe, 642–643
  laser absorption spectroscopy, 642
  laser-induced fluorescence, 643
  laser reflection interferometry, 641
  mass and energy sampling, 642
  microwave interferometry, 642
  optical emission spectroscopy, 642
Plasma disk source method, plasma etching, 663
Plasma enhanced chemical vapor deposition, 711–728
  densities on plasmas, 714
  development of, 715–716
  materials deposited by, 715
  operational factors, 711–717
  particulate control, 718
  remote deposition, 722–723
  scaleup, 723–727
  uses of, 716–717
Plasma etching, 628–667
  contamination control, 640–641
  and damage, 640
  end point detection, 666
  etch profile control (anisotropy), 635–638
  high etch rate, 635
  ion-bombardment-enhanced etching, 636
  ion-enhanced etching, 636
  layers in semiconductor integrated circuits, 634
  low-pressure etching, 636
  physical processes in, 628–631
  selectivity, 638–639
  sheath formation, 631–633
  steps in, 633
  trends related to, 665
  uniformity of etch, 639–640
  uses of, 628
  water temperature effects, 637–638
Plasma etching tools
  advanced etch reactors, 649–667
  barrel reactors, 644–645
  cluster tools, 666
  development of tool, 643–644
  diffusion source configuration, 663–664
  diode configuration limitations, 649
  electron cyclotron resonance sources, 662–664
  electrostatic chucks, 666
  helical inductive couplers, 659–660
  helical resonators, 660–661
  high-pressure diodes, 645–646
  hollow-anode-enhanced triodes, 654–656
  induced DC bias, 648–647
  inductively coupled plasmas, 658–662
  low-pressure etching, 646–647, 650–651
  magnetic enhancement, 650–651, 652–653
  magnetically enhanced reactive ion etcher (MERIE), 652

Plasma etching tools (*continued*)
magnetically enhanced triodes, 656–658
magnetron sputtering, 652–653
matching network, 648, 662, 664
original equipment manufacture (OEM) support, 666–667
parallel-plate triodes, 654
plasma disk source method, 663
plasma generation, control of, 651–652
reactive ion etchers, 646–647
spiral inductive couplers, 661–662
substrate bias, control of, 651–652
Plasma potential, 630
Plasma-related applications, mass analysis, 294–295
Plastics
and outgassing, 488–489
outgassing, 497–501
Plexiglass, 460
PNEUROP flanges, 414
Poiseuille's equation, 35–36
Polishing, 554
aluminum alloys, 559
chemical polishing, 557
copper, 560–561
electropolishing, 556–558
fine/course abrasive powder, 555
low-carbon steel, 558
stainless steel, 556–558
Polyimide
as electrical insulator, 457
as seal material, 472
Polymer seals, 463–472
compression, 466–467
design factors in use of, 472–474
limitations in use of, 463–464
O-ring material selection, 471–472
outgassing, 468–471
permeation, 470–471
selection criteria, 465–466
trade names, 464–465
Polymers
cleaning, 569
outgassing, 561–562
properties of, 448
Polymethylmetacrylate, for vacuum chamber, 460
Polystyrene divinylbenzene, 460
Polytetrafluoroethylene (PTFE), as electrical insulator, 457
Polyurethane, as seal material, 472
Polyvinyl chloride (PVC), in vacuum materials, 460
Poppet valves, 392–393, 394, 405, 407
Pores, 47–48
Porosity, avoiding and welding, 542
Positive ray analyzer, 336
Power common, mass flow controller, 384
Precipitated calcium carbonate, 570
Pressure
gas, 8, 15–16
and outgassing, 491–492, 495–497
Pressure gauges
capacitance gauges, 778
ion gauges, 778
roll-to-roll vacuum coating, 778
thermal conduction gauges, 778
Pressure measurement, gauges, types of, 258–282
Pressure rate-of-rise technique, mass spectrometer calibration, 366–367
Pressure ratio
lighter gases, 133
vapor jet pumps, 132
Pressure regions, 3–5
Pressure stability, vapor jet pumps, 146–148
Pressure units, conversion factors, 4
Primary backstreaming, 137–138
Printing products, roll-to-roll vacuum coating, 765
Process control, mass analysis, 295–296
Process pumping, 74–75
Process-related contamination, 596
Process specifications, 600
Processing environment, and surface preparation, 587–589
Processing of parts, 566–567
and surface chemistry, 566–567
and surface morphology, 566
Product dehydration systems, vacuum blowers, 114–115
Protective coatings, aluminum, 541
Pulsed laser deposition, 694–707
ablation mechanism, 698–699
advantages to, 700–701
and crystallinity enhancement, 700–701
development of, 694–695
and diamondlike carbon, 706
future view of, 708
large-area uniformity, 701–702

limitations of, 701–705
and metal films, 706
modifications to system, 696–697
multilayer growth, 701
and nitride films, 707
operational factors, 695–697
and organic materials, 707–708
and oxide films, 707
and perovskites, 706–707
and reactive deposition, 701
and semiconductors, 706
splashing, 702–705
and target/substrate geometry, 696
types of lasers, 695–696
Pump protection valves, 407
Pumpdown, 65–69
constant speed case, 66–67
constant throughput case, 67
initial, from atmospheric pressure, 494
and outgassing, 68–69, 494–495, 504–506
in presence of leak, 67–68
in presence of outgassing, 68–69
second pumpdown, 504–506
time frame, 65–66
Pumping chamber, diaphragm pumps, 84–85
Pumping speed, 23, 64–65
diaphragm pumps, 87–88
geometries used, 64
and outgassing, 493
sputter-ion pumps, 229–231
turbomolecular pumps, 189
vacuum blowers, 103–104
vapor jet pumps, 122–126
Pyralin, 457
Pyre ML, 457
Pyrex, 452

Q

Quadrupole instruments, ion generation, 307
Quadrupole ion trap, 319
Quadrupole mass filters, 316
Quartz viewports, 444–445

R

Radial seal, 473
Radial sources, ion generation, 306
Radioactive gases, turbomolecular pumps, 204
Radiated heat, aluminum, absorption of, 536–537
Rate sensor, mass analysis, 296
Reactive cleaning, 575–578
Reactive ion beam etching, 682
Reactive ion etchers, plasma etching, 646–647
Reactive sputtering, 621–623, 772–774
Rectangular flanges, 425
Refrigerator operation, cryogenic pumps, 152–153
Regeneration, cryogenic pumps, 165–167, 180
Reinburg reactors, 645–646
Release agents, for stripping, 563
Remote deposition, 722–723
Removable surfaces, 562–565
stripping, 562–564
Residual gas analyzers, 298, 492
Resistive heating, for evaporation, 768
Resolving power, mass analysis, 300
Resonance gauges, 278–279
Rexolite, 460
Reynolds number, 26
RF diodes, sputter deposition, 614–615
RF ion source, 674
Rhenium, 456
ion generation, 305
Rinsing, in cleaning process, 578–580
Robotic arms, coaxial feedthrough, 442
Roll-to-roll vacuum coating, 761–787
abrasion resistance, 786
adhesion, causes of, 786–787
boat materials, 769–770
cathodic arc deposition, 774
chemical vapor deposition, 775
corrosion rates, decreasing, 785–786
crucible materials, 769–770
deposition materials, 766
development of, 761–762
electron beam heating, 768–769
environmental problems of, 785–787
evaporation, 766–769
film measurement after deposition, 784
inductive heating, 768
monitoring/control, 783–784
organic materials, deposition of, 775
pressure gauging, 778
resistive heating, 768
sputtering, 770–774
steps in process, 765
substrates (webs), 779–783

Roll-to-roll vacuum coating (*continued*)
types of units for, 762–763
uses of, 763–765
vacuum pumps, 775–777
Rotameters, 377
Rough pumping sputter-ion pumps, 233–236
Rutherford backscattering spectrometry, 672, 751–757
instrumentation, 753–754
operational factors, 751–753

S

Safety factors
cleaning, 586
tolerable forepressure, 129–130
vapor jet pumps, 129–130
Sample introduction system, 755–757
horizontal introduction system, 756
instrumentation, 756
pumpdown time, 756
Sanford, 454
Saponifiers
for cleaning, 572–574
types of, 572–574
Sapphire viewports, 445
Scaleup
modular scaleup, 724–725
plasma-enhanced chemical vapor deposition, 723–727
Scan rate, mass analysis, 300
Scroll pumps, 235
Sealed gas sample analysis, uses of, 296–297
Seals
bonnet seals, 402
demountable seals, 512–518
dynamic seals, 463
elastomer seals, 472–474
metal seals, 464, 474–482
polymer seals, 463–472
static seals, 463
within valves, 402
Secondary backstreaming, 137–138
Secondary ion mass spectrometry, 308, 373, 745–751
dynamic SIMS, 748–749
operational factors, 745–748
static SIMS, 749–751
Selectivity, plasma etching, 638–639
Selenium, in vacuum materials, 461
SEMI flanges, 425, 426
Semiconductors
and pulsed laser deposition, 706
turbomolecular pumps, 206–207
Sensing techniques, mass flow controller, 378–379
Sequenced start systems, vacuum blowers, 108–109
Series, conductances in, 33–34
Sheath formation
plasma etching, 631–633
presheath, Bohm criteria, 632–633
Sheaths, nature of, 631
Shock-wave-induced droplet expulsion, and splashing, 702
Shutoff valves, types of, 391–396
Signal common, mass flow controller, 384
Silica gel, 591
Silver
avoiding, 461
as electrical conductors, 456
as seal metal, 481
Simple oxides, and pulsed laser deposition, 707
Single-wafer etchers, 649–650
Single-wafer tool, 624
Skull, 770
Slit valves, 393
Slush hydrogen pumping systems, vacuum blowers, 113–114
Snow-Floss, 570
Soaking, cleaning, 580
Soft start valves, 407
Solubility, and construction materials, 447, 449
Solvents
cautions in use of, 438
for cleaning, 571–572
densifing, 572
polar/nonpolar, 571
solubility parameter, 571
types of, 571–572
Somorjai's notation, 42
Sorption pump, 235–236
operation of, 236
Space missions, mass spectrometers, 336
Special construction ion sources, ion generation, 307
Spectrum fingerprints, 344
Spiral inductive couplers, plasma etching tool, 661–662

Splashing, 702–705
avoidance of, 703–705
causes of, 702–703
Split flow turbomolecular pumps, 200
Spraying
cleaning, 581
particulate removal, 570
Sputter cleaning, 598
Sputter deposition, 611–624
bias sputtering, 623–624
DC diodes, 613–614
ion beam, 683–687
magnetron sputtering, 616–623
mass analysis, 371–372
RF diodes, 614–615
Sputter etching, compared to sputter deposition, 611–612
Sputter ion pumps, 220–242
combined with nonevaporable getter pumps, 252
construction of, 227–229
operation of, 221–227
operational life of, 231–232
pressure measurement with, 232
pumping speed, 229–231
rough pumping sputter-ion pumps, 233–236
start-up/running of, 236–240
troubleshooting/maintenance of, 240–241
for ultra-high vacuum, 219
uses of, 232–233
Sputter plate, 176
Sputtering, 609–626
amplified sputtering, 626
bias sputtering, 623–624
and deposition, 611–624
and energy levels, 609–610
future view of, 626
high-rate sputtering, 771–772
industrial uses of, 624–626
intrinsic resputtering, 612
ionized sputter deposition, 626
knock on effects, 609
magnetron sputtering, 616–623, 624–625
materials/alloys for, 774
nature of, 609
near-threshold sputtering, 609
reactive sputtering, 621–623, 772–774
roll-to-roll vacuum coating, 770–774
sputter yields, 609–610
Sputtering systems, cryogenic pumps, 176–177
Square card diagram, 105
Stainless steel
alloys, types of, 554–555
cleaning of, 568–569
in construction of vacuum chamber, 452–453
corrosion rate, 540
mill plate/finishes, types of, 555–556
and outgassing, 486
polishing, 556–558
relative thickness, 534–535
Stanford Positron-Electron Asymmetric Ring (SPEAR), 793, 794, 796, 797–798
StarCell geometry, 224
Static process, calibration of gauges, 287
Static seals, 463
Static secondary ion mass spectrometry, 749–751
Static volume expansion, mass spectrometer calibration, 365
Stellarators, 474
Sticking coefficient, 43–44
Storage, 590–591
active storage, 591
passive storage, 591
Storage rings, 790–798
development of, 790–795
materials used, 797
pumping system developments, 796–798
Stripping, 562–564
methods for, 563
Sublimation pumps, 242–247
control units for, 247
electron bombardment heating, sublimation from, 247
operation of, 243–244
titanium spheres, sublimation from, 246–247
for ultra-high vacuum, 219–220
uses of, 244–245
wire sublimator sources, 246
Subsurface boiling, and splashing, 702
Sulfur, in vacuum materials, 461
Super-Conducting Super Collider, 790, 798
Superconducting RF accelerators, 800–801
development of, 800–801
Supercritical fluids, 572
Supermagnetron, 657
Surface adsorption isotherms, 45–47
Surface analytical methods
electron spectroscopies, 732–745
in-line use, 757

Surface analytical methods (*continued*)
Rutherford backscattering spectrometry, 751–755
sample introduction system, 755–757
secondary ion mass spectrometry, 745–751
Surface area, 44
Surface migration, backstreaming, 140
Surface modification, 554–565
aluminum alloys, 559–560
ceramics, 561
copper, 560–561
glass, 561
hardening of surface, 565–566
low-carbon steel, 558
polymers, 561–562
removable surfaces, 562–565
stainless steel, 554–558
Surface physics
adsorption, 40–41, 43–47
capillary action, 47–48
chemisorption, 40, 42
condensation, 48–49
desorption, 49–52, 53–54
gas release from surface, 54
photoactivation, 52–53
surface area, characteristics of, 44
Surface preparation
cleaning, 553, 567–601
processing of parts, 566–567
surface modification, 554–565
Surface roughness, necessity of, 554
Surface tension, and cleaning, 574
Surface-to-volume ratio, and pump performance, 74
Surf-riding mechanism, 632
System-related contamination, 592–596

T

Tailing contribution, mass analysis, 299
Tantalum, 456
Tee element, inlet systems, 302–303
Teflon, 457, 569
Televac, 269
Temperature, of gas, 9
Temperature desorption spectroscopy, 373
Terpenes, 572
Thermal conduction gauges, 778
Thermal conductivity
aluminum, 536
coefficient of, 21
Thermal conductivity gauges, 265–270
Thermal conductors, materials for, 457–458
Thermal desorption, 50–52, 596–597
Thermal insulators, construction materials, 457–458
Thermal radiation, cryogenic pumps, 162
Thermal siphoning, mass flow controller, 385
Thermal time constant, aluminum, 537–538
Thermal transpiration, 20–21
Thermistor gauges, 268
Thermocouple gauges, 266–268
Thin-film magnetoresistance (MR) read heads, 686
Thoria-coated cathodes, ion generation, 305
Thoriated tungsten, 305
Throttling valves, 398
Throughput
conversion factors, 31
gas transport, 61–62
maximum, 127–128
vapor jet pumps, 124, 127–128
Throughput pumps, 60
Throughput Q, 30
Time to form monolayer, 20
meaning of, 4
Time-of-flight mass spectrometry, 299, 322–323, 373–374
operational factors, 749–751
Titanium, 456
Titanium hydride, 226
Titanium spheres, sublimation from, 246–247
Tokamaks, 802, 805–809
future view, 810
historical view, 805–806, 808–809
impurity problem, 806–808
plasma interaction in vessel wall, 806
Tolerable forepressure
backing pump criteria, 128–129
discharge pressure, 128
and pressure ratio, 130–131
safety aspects, 129–130
vapor jet pumps, 128–132
for various gases, 131–132
Torr, 3
Transfer standards, mass spectrometer calibration, 370
Traps, vapor jet pumps, 133
Trichloroethylene, 571
Triode geometry, advantages of, 224–225
Triode hot cathode ionization gauges, 271–273
Triodes, plasma etching tools, 654–658

True splashing, 702
Tubulations, vacuum valves, 402–403
Tungsten, 456
  ion generation, 304–305, 346–347
Tungsten sputter, 456
Turbine-type pumps, 76
Turbomolecular pumps, 95, 183–212
  backing pumps, 202
  baking out, 202–203
  balancing/vibration, 194
  combined with cryopumps, 205–207
  combined with MDP pumps, 197–200
  compression, 188
  cooling of, 203
  corrosive gases, pumping of, 203–204
  design elements, 190–195
  development of, 184, 192
  drive systems, 194–195
  features/applications, 201
  guidelines for operation of, 211–212
  and magnetic fields, 203
  molecular drag pumps, 195–197
  operating principles, 183–184
  performance factors, 184–188
  pumping speed, 189, 190
  radioactive gases, pumping of, 204
  rotators/stators, 191–192, 194
  rotor suspension, 192–194
  in semiconductor manufacturing, 206–207
  special measures/applications, 208–210
  split flow pumps, 200
  trade names for, 200–201
  ultimate pressure, 190
  for ultra-high vacuum, 218–219
  venting of, 202, 205–206
Turbulent gas flow, 26, 27–28
Two-stage valves, 405

U

Ultimate pressure, 69–71
  molecular drag pumps, 195–196
  pump limit versus system limit, 70–71
  turbomolecular pumps, 190
Ultra-high vacuum, 7, 71–72
  cryogenic pumps, 217–218
  definition of, 214
  diffusion pumps, 217
  getter pumps, 220, 242–252
  sputter-ion pumps, 219, 220–242
  sublimation pumps, 219–220
  system design for, 215–216
  turbomolecular pumps, 218–219
Ultra-high-vacuum technology, 789–810
  cold-bore machines, 798–800
  development of, 789–790
  magnetic fusion, 801–810
  Next Linear Collider project, 801
  storage rings, 790–789
  superconducting RF accelerators, 800–801
Ultrapure water
  large volumes, making of, 580
  for rinsing, 578–580
Ultrasonic cleaning, 582–584
  cavitation bubbles, formation of, 582–583
  factors related to, 583
  and fixturing, 584
  limitations of, 583–584
Ultrasonic desorption, 53, 598
Unconstrained flanges, 518

V

Vacuum
  applications of, 5
  definition of, 3
  degree of, 4
  high vacuum, 6–7
  low-vacuum, 5
  medium-vacuum, 5
  pressure regions, 3–5
  ultra-high vacuum, 7, 215–252
Vacuum applications
  ion beam technology, 672–691
  mass analysis, 298
  plasma etching, 628–667
  plasma-enhanced chemical vapor deposition, 711–728
  pulsed laser deposition, 694–707
  sputtering, 609–626
  surface analytical methods, 732–757
Vacuum blowers, 97–115
  blower rotating speed control, 109
  bypass valves, 109–110
  components of, 97–100
  compression ratio, 101–102
  cool gas feedback blowers, 111
  discharge gas cooling, 108
  gas carryback, 103
  gas species, 102
  inlet throttling systems, 109
  liquid-sealed blowers, 112

maximum allowable compression ratio, 107–108
maximum allowable discharge temperature, 107
maximum allowable temperature rise, 106–107
multistage systems, 112–113
operating principles, 100–101
power requirements, 104–106
pressure level, 101–102
product dehydration systems, 114–115
pumping efficiency, 101–103
pumping speed calculations, 103–104
rotational speed, 102
sequenced start systems, 108–109
slush hydrogen pumping systems, 113–114
Vacuum chamber, removing air from, 59–60
Vacuum chamber materials, 451–455
aluminum, 453–455
glass, 451–452
steel, 452–453
Vacuum measuring instruments, 282–286
combination instruments, 282, 284
intelligent gauge tubes, 286
modular instruments, 284
single-measurement instruments, 282
Vacuum pumps
cryogenic pumps, 149–181
cryopumps, 776–777
diaphragm pumps, 84–96
diffusion pumps, 776
gas transport, 61–62
performance parameters, 62–64
process pumping, 74–75
pump system relationships, 73–77
pumpdown time, 65–69
pumping speed, 64–65
roll-to-roll vacuum coating, 775–777
turbomolecular pumps, 183–212
types of pumps, 75–76
ultimate pressure, 69–71
valved pumping system, 82–83
valveless systems, 80–82
Vacuum valves
actuation of valves, 400–401
butterfly valves, 397, 398, 407
conductance, 402–403
control valves, 397–398
diaphragm valves, 393, 396
flanges, 402–403
gate valves, 393, 395, 406, 407
installation of, 407–408
multiblade valves, 399
poppet valves, 392–393, 404, 405, 407
sealing of actuating motion, 400
seals within valves, 402
shutoff valves, 391–396
slit valves, 393
specialized valves, types of, 404–407
throttling valves, 398
tubulations, 402–403
Valved pumping system, 82–83
Valveless systems, 80–82
van der Waals forces, 41
Vapor cleaning, 581–582
system for, 581–582
Vapor contamination, 588
Vapor deposition, mass analysis, 371
Vapor drying, 585
Vapor jet pumps, 76
backing pump requirements, 128–129
backstreaming, 137–143
baffles, 132–133
cold caps, 146
design features, 144–145
discharge pressure, 128
fluid breakdown/purification, 136–137
gas type and speed, 125–126
gas type and tolerable forepressure, 131–132
operating range, 121
performance factors, 121–122
pressure ratio, 132, 133
pressure ratio effects, 130–131
pressure stability, 146–147
pumping fluid section, 133, 135
pumping mechanism, 117–120
pumping speed, 122–126
pumping stages, number of, 120
residual gas analysis, 135–136
safety factors, 129–130
size effects, 124–125
speed-efficiency capture probability, 123
throughput, 124, 127–128
tolerable forepressure, 128–132
traps, 133
ultimate pressure, 132–137
vapor/gas interaction, 120–121

Vapor pressure, 10
  cryogenic pumps, 157–159
  high-vapor-pressure materials, 460–461
  low-vapor pressure materials, 446–447
Vee-grove joints, 545, 546
Viewports
  glass viewports, 444
  mounting system for, 445
  quartz viewports, 444–445
  sapphire viewports, 445
Virtual leak, 473
Viscous gas flow, 28
  circular tube, 35–36
  rectangular duct, 36
Viton, 464–465, 569
Viton E60C, 465
Viton gaskets, 412
Volatile organic compounds, 572
Volume, gas, 8–9
Volumetric flow rate, gas flow, 30

W

Wall creep
  and contamination, 592–593
  minimizing, 593
Water, and outgassing, 447
Water vapor, 45
  cyrogenic pumps, 177–178
Webs, 779–783
  defects, types of, 783
  handling of, 780–782
  materials for, 779–780
  roll-to-roll vacuum coating, 762–763, 779–783
  temperature control, 782–783
Welding
  AC-TIG method, 547
  aluminum, 541–548
  crack defects, avoiding, 542–546
  DC-TIG methods, 547
  particulate defects, avoiding, 542
  porosity, avoiding, 542
Whiskers, removal from ion pump, 240–250
Wire seals, 515
Wire-sealed flanges, 423–424

X

X-ray photoelectron spectroscopy, 732, 733, 736, 737–740
X-ray sources, 744–745

Z

Zinc, 446
  avoiding, 461

This Page Intentionally Left Blank