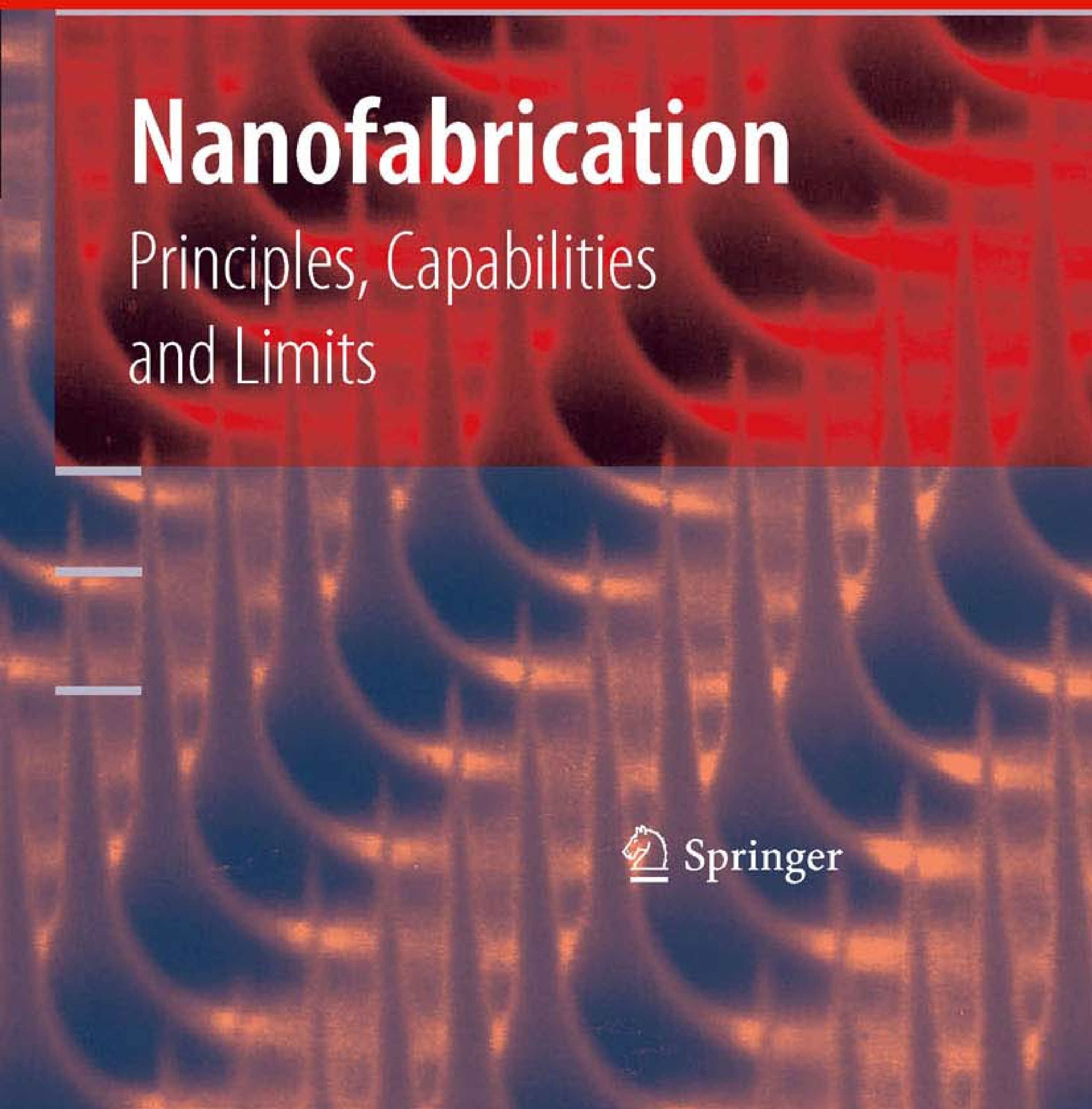
Zheng Cui

Nanofabrication

Principles, Capabilities and Limits



Nanofabrication

Zheng Cui

Nanofabrication

Principles, Capabilities and Limits

Zheng Cui
Rutherford Appleton Laboratory
Science and Technology Facility Counsil (STFC)
Harwell Science and Innovation Campus
Didcot, OX11 0QX
United Kingdom

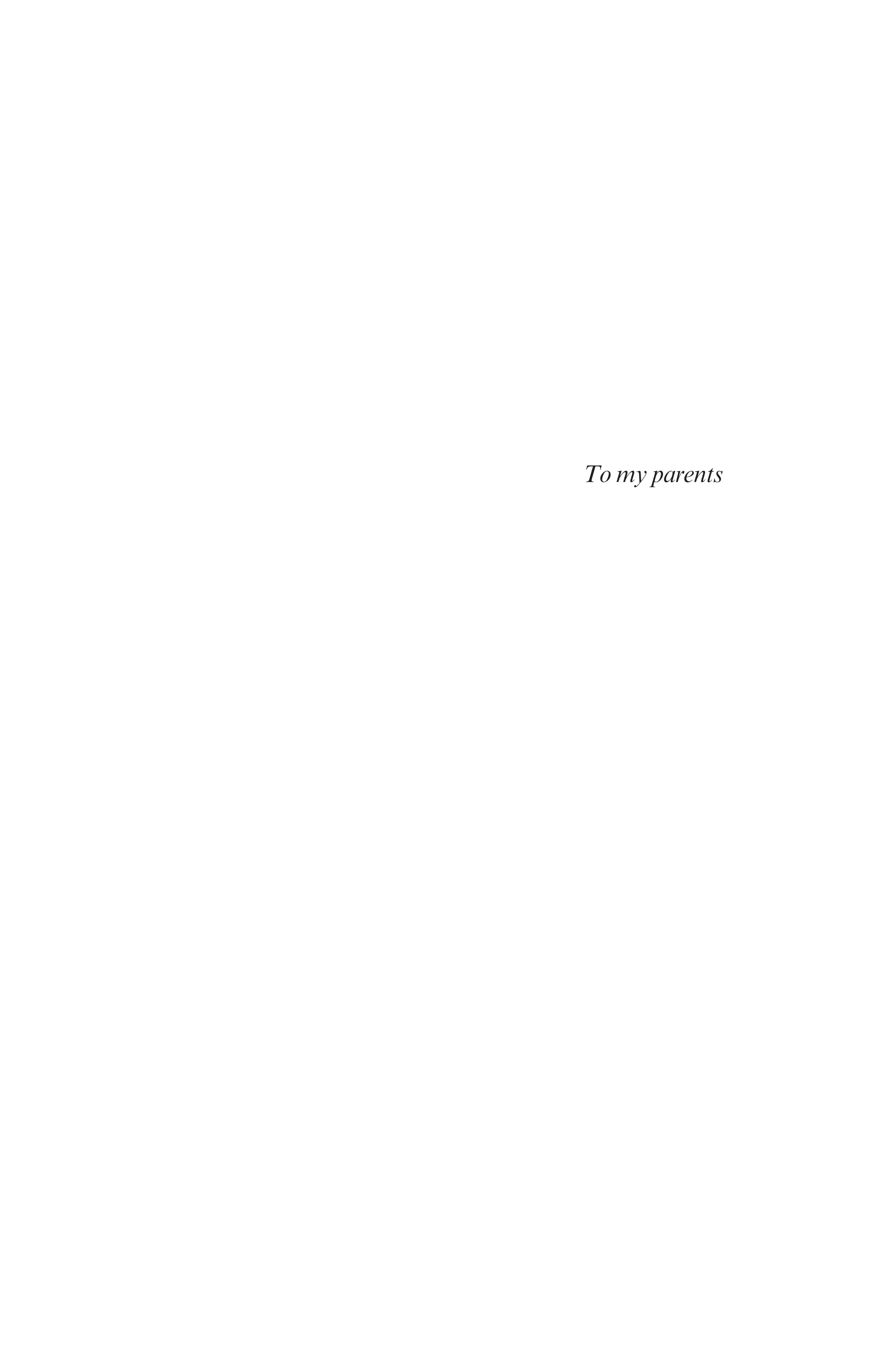
ISBN: 978-0-387-75576-2 e-ISBN: 978-0-387-75577-9
DOI: 10.1007/978-0-387-75577-9

Library of Congress Control Number: 2008924169

© 2008 Springer Science+Business Media, LLC
All rights reserved. This work may not be translated or copied in whole or in part without the written permission of the publisher (Springer Science+Business Media, LLC, 233 Spring Street, New York, NY 10013, USA), except for brief excerpts in connection with reviews or scholarly analysis. Use in connection with any form of information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed is forbidden.
The use in this publication of trade names, trademarks, service marks, and similar terms, even if they are not identified as such, is not to be taken as an expression of opinion as to whether or not they are subject to proprietary rights.

Printed on acid-free paper

springer.com



To my parents

Preface

The idea of writing a book focusing on nanofabrication came after I gave a series of lectures in 2007. First I was invited to give a lecture at the Nanoelectronics Research Centre in the Peking University, China, in January 2007. The title of my lecture was "Nanofabrication: Capabilities and Limits". Then I was invited to give a presentation with the same title at the UK Institute of Physics Conference on "Nanoelectronics-Materials and Technologies" in February 2007. The audiences, who were mostly working on new nanoelectronic devices, were keen to know which technology they could use to make the nanostructures they need. In April 2007, I was on a lecture tour in Taiwan where I lectured at National Taiwan University, Taiwan National Nanodevice Laboratory and National Cheng-Kung University. Wherever I lectured, researchers and students were showing great interest in nanofabrication technologies and concerned about their capabilities and limits. My collaboration with researchers in China, particularly with the Institute of Physics, Chinese Academy of Sciences, where a number of research groups are devoted to nanomaterials and devices, has also exposed me to frequent questions such as what are the capabilities and limits of various current nanofabrication technologies. Although I have published a book, titled "Micro-Nanofabrication Technologies and Applications" (Springer, 2006), my previous book, including a number of other books published so far, has not made clear which technologies can truly claim to be nanofabrication technologies. The descriptions of nanostructure fabrication are often buried in those of microstructure fabrication. It is apparent that a book on this subject is desirable.

It is no surprise that nanofabrication is drawing wide spread interest, as nanotechnology has been a buzz word in the last few years across the world. Not only the traditional physical science community is drawn into nanotechnology, but also researchers in chemistry, medicine, biology, energy and environment sciences are all becoming interested in nanotechnology. Though nanoscience and nanotechnology are broadly defined as research and exploitation of structures and systems at sub-100 nm dimension, the distinction between microfabrication and nanofabrication is not too obvious. Many traditional technologies for fabricating microscale structures are also capable of making

sub-100 nm structures, such as optical lithography, though many advances in the optical lithography taking place over the years have made the sub-100 nm patterning possible. On the other hand, some of the technologies are inherently "nano" rather than "micro", such as fabrication based on scanning probes. I have written the book especially with the sub-100 nm dimensional scale in mind, and have selected only those technologies that are capable of making sub-100 nm structures, or emphasized on the aspects of sub-100 nm structuring capabilities for those technologies which have been traditionally for microfabrication. As the book title suggested, I have described not only the nanofabrication capabilities for each of the selected technology, but also their limits which are either about the ultimate patterning resolution or about the practicality as a fabrication technology.

Nanofabrication is a multi-disciplinary field. A book in this nature is often written by multiple authors. I was able to take on this task on my own because of my many years working with various microfabrication and nanofabrication technologies. The Central Microstructure Facility, Rutherford Appleton Laboratory, where I have worked for 15 years, has been a place with sole job to provide micro and nanofabrication solutions to researchers of many scientific disciplines. My connection with Laboratory of Microfabrication at the Institute of Physics, Chinese Academy of Science, in the last 6 years as a visiting professor also has given me the opportunity to work with many scientists in nanoscience and nanotechnology areas. Though I cannot claim to be an expert in every nanofabrication technology described in the book, my experience and insight into how a nanostructure should be made give me the advantage of knowing which is a viable nanofabrication technology and which is not. This helped me to select only those most appropriate technologies to be included in the book. I hope the readers of the book should gain useful information and advice when deciding which technology they should choose for their specific applications.

In the era, when any information can be "googled" from internet, who would still need a book? In fact, more titles of professional books are published than ever before. Though internet has provided the convenience of finding required information at a stroke of keyboard, it still needs a trained eye to spot the correct and best knowledge among the vast amount information presented to someone. I must admit that internet has helped me a great deal in gathering information and preparing for the book. I hope that with my trained eyes I have filtered the useful information through to my book. In addition, the book has provided references, web links and explanation of terminologies, which should serve as a "street map" for readers to find their way through in the internet "maze".

My gratitude goes to many of my colleagues and friends working at the Central Microstructure Facility, Rutherford Appleton Laboratory, UK, at the Laboratory of Microfabrication, the Institute of Physics, Chinese Academy of Sciences, and at the Nanoelectronics Centre of Peking University, China, where I also serve as a visiting professor. Many of their published work have been cited

in the book, including work published by many other experts in different fields. My final thanks go to my family, my wife Ling Wang who herself is also an experienced engineer working in the micro and nanofabrication field, and my daughters, Helen and Katy, who have supported me all along.

Oxford Zheng Cui

Contents

1 Introduction 1
1.1 Nanotechnology 1
1.2 Nanofabrication 2
1.3 Purpose of the Book 4
References 5

2 Nanofabrication by Photons 7
2.1 Introduction 7
2.2 Principle of Optical Projection Lithography 8
2.3 Optical Lithography at Shorter Wavelengths 13
2.3.1 Deep UV 13
2.3.2 Extreme UV 16
2.3.3 X-ray 20
2.4 Optical Lithography at High NA 22
2.5 Optical Lithography at Low k_1 Factor 30
2.5.1 Off-Axis Illumination (OAI) 33
2.5.2 Phase-Shifting Mask (PSM) 34
2.5.3 Optical Proximity Correction (OPC) 38
2.5.4 Photoresists 44
2.5.5 Design for Manufacturing (DFM) 53
2.5.6 Double Processing 54
2.6 Near-Field Optical Lithography 57
2.7 Interferometric Optical Lithography 63
2.8 Maskless Optical Lithography 67
References 73

3 Nanofabrication by Charged Beams 77
3.1 Introduction 77
3.2 Focusing Charged Particle Beam 79
3.2.1 Charged Particle Optics 79
3.2.2 Sources 81
3.2.3 Aberrations 85

3.3 Scattering and Proximity Effect . . . 89
3.3.1 Electron Scattering . . . 89
3.3.2 Proximity Effect and Correction . . . 94
3.3.3 Effect of Secondary Electrons . . . 99
3.3.4 Low-Energy E-Beam Lithography . . . 101
3.3.5 Ion Scattering . . . 105
3.4 Resist Materials and Processes . . . 106
3.4.1 Sensitivity of Resist Materials . . . 107
3.4.2 Contrast of Resist Materials . . . 112
3.4.3 Resolution Enhancement Processes . . . 114
3.5 Ion Sputtering and Redeposition . . . 117
3.6 Charged Particles Projection Lithography . . . 120
References . . . 124

4 Nanofabrication by Scanning Probes . . . 129
4.1 Introduction . . . 129
4.2 Principles of SPMs . . . 130
4.3 Exposure of Resists . . . 133
4.3.1 Field Electron Emission . . . 133
4.3.2 Exposure of Resist by STM . . . 135
4.3.3 Exposure of Resist by NSOM . . . 139
4.4 Local Oxidation Lithography . . . 140
4.5 Additive Nanofabrication . . . 143
4.5.1 Field-Induced Deposition . . . 143
4.5.2 Dip-Pen Nanolithography . . . 145
4.6 Subtractive Nanofabrication . . . 147
4.6.1 Electrochemical Etching . . . 147
4.6.2 Field-induced Decomposition . . . 148
4.6.3 Thermomechanical Indentation . . . 149
4.6.4 Mechanical Scratching . . . 150
4.7 High-Throughput SPL . . . 153
References . . . 156

5 Nanofabrication by Replication . . . 161
5.1 Introduction . . . 161
5.2 Thermal Press Nanoimprint . . . 162
5.2.1 Nanoimprint Stamps . . . 164
5.2.2 Nanoimprint Polymers . . . 166
5.2.3 Demolding . . . 169
5.2.4 Alignment . . . 172
5.3 Room Temperature Nanoimprint . . . 173
5.4 UV-Cured Nanoimprint . . . 176
5.4.1 Transparent Stamps . . . 176
5.4.2 UV Curable Polymers . . . 180

5.4.3 Step-and-Flash Imprint Lithography 183
5.4.4 Alignment Through Transparent Stamps 187
5.4.5 Combined Nanoimprint and Photolithography 189
5.5 Reverse Nanoimprint 192
5.6 Soft Lithography 196
5.6.1 Soft Stamps 197
5.6.2 Microcontact Printing 199
5.6.3 Replication by Capillary Force 202
References 206

6 Nanoscale Pattern Transfer 211
6.1 Introduction 211
6.2 Additive Pattern Transfer. 212
6.2.1 Thin Film Deposition 212
6.2.2 Pattern Transfer by Lift-Off 218
6.2.3 Pattern Transfer by Plating 226
6.2.4 Damascene Process 229
6.2.5 Pattern Transfer by Stencil Mask 231
6.3 Subtractive Pattern Transfer 234
6.3.1 Isotropic Wet Etching 234
6.3.2 Anisotropic Wet Etching 236
6.3.3 Reactive-Ion Etching (RIE). 238
6.3.4 Process Control in Nanoscale RIE 242
6.3.5 RIE by Inductively Coupled Plasma 249
6.3.6 Critical Issues in RIE. 253
6.3.7 Ion Milling. 258
References 262

7 Indirect Nanofabrication 267
7.1 Introduction 267
7.2 Sidewall Lithography 268
7.3 Dimensional Subtraction and Addition 272
7.3.1 Lateral Subtraction 272
7.3.2 Lateral Addition 278
7.3.3 Vertical Subtraction. 280
7.3.4 Nanosphere Lithography 283
7.4 Multistep Processing 285
7.5 Super Resolution Patterning 287
References 292

8 Nanofabrication by Self-Assembly 295
8.1 Introduction 295
8.2 Self-Assembly Processes 296
8.3 Guided Self-Assembly 304

8.3.1 Surface Topography 305
8.3.2 Surface Wetting 305
8.3.3 Electrostatic Force. 308
8.3.4 Magnetic Force 310
8.4 Building Blocks of Future Nanosystems 312
8.4.1 DNA Scaffold 313
8.4.2 Carbon Nanotubes 316
8.4.3 Block Copolymers 321
8.4.4 Porous Alumina. 324
References 329

Index 335

Chapter 1
Introduction

1.1 Nanotechnology

There has never been a technology which has caught such wide attention and imagination in a short period of few years as nanotechnology. It was envisioned in 1959 by Richard Feynman, the Nobel Laureate in physics, in his famous prediction "there are plenty of rooms at the bottom", and the word "nanotechnology" was coined by Professor Norio Taniguchi in 1974 and then popularized by K. Eric Drexler in 1980s in his book *Engines of Creation: The Coming Era of Nanotechnology*. Nanotechnology was really coming out of science fiction to become reality following the invention of scanning tunneling microscope (STM) which made it possible to see and manipulate atoms. The pervasive use of nanoparticles and carbon nanotubes in the last few years greatly fueled the global enthusiasm in nanotechnology. It has become the hottest pursued technology worldwide in recent years.

Nanotechnology deals with materials and systems at or around nanometer scale (1 billionth of a meter). Nanotechnology, according to Professor George Whitesides of Harvard University, is just "A word, not a field" [1]. It is a collective term for a set of technologies, techniques, and processes, rather than a specific science or engineering discipline. There are many subject areas under the banner of nanotechnology, such as nanoelectronics, nanomaterials, nanomechanics, nanorobotics, nanomagnetics, nanophotonics, nanobiology, nanomedicine, etc.: the list goes on. Any scientific discipline can attach the "nano-" label to it to become a part of nanoscience and nanotechnology. Yet what in common for the vast different nanotechnology disciplines is the dimensional scale. There has been a general consensus that nanotechnology is to control and restructure matters at 1–100 nm dimension and the creation of materials, devices, and structures around this dimension [2]. It has been found that many materials and structures at below 100-nm dimensional scale can have properties and characteristics dramatically different from their bulk forms. Therefore, the 100-nm dimensional scale has set the boundary between nanotechnology and all other microscale, mesoscale, and conventional macroscale technologies.

Z. Cui, *Nanofabrication*, DOI: 10.1007/978-0-387-75577-9_1,

© Springer Science+Business Media, LLC 2008

The key to nanotechnology is the making of various nanostructures, which specifically refers to making structures less than 100 nm in at least one dimension. For one-dimensional nanostructures there are thin films of a few atomic layers thickness. At three-dimension there are nanoparticles, carbon nanotubes, and nanowires. For thin films, manufacturing processes have long been established, such as molecular beam epitaxy (MBE) and atomic layer deposition (ALD) processes. Nanoparticles have already been manufactured in kilos or tons at industrial scale. Large-scale manufacturing of carbon nanotubes is also commercially available. Apart from nanoscale thin films and nanoparticles, a large category of nanostructures are two-dimensional, or quasi three-dimensional nanostructures which are built layer upon layer from two-dimensional structures. Typical examples include planar transistors, magnetic or photonic elements, and nanomechanical elements, etc. They are the basic building blocks to form a variety of functional nanodevices and systems. Nanofabrication, which differs from manufacturing of thin films and nanoparticles, is the key technology to enable this category of nanosystems.

1.2 Nanofabrication

Nanofabrication evolves from microfabrication. Since the advent of first transistor in 1947, microelectronics and integrated circuit (IC) industry has been the main driving force to continuously push fabrication technologies to their new dimensional limit. Following the famous “Moore’s Law”, the semiconductor industry was able to double the density of transistors on a unit area of silicon chip in every 18 months. In fact, the shrinking of IC feature size is faster than predicted by Moore’s law. According to prediction of the first edition of International Technology Roadmap for Semiconductors (ITRS) published in 1993 [3], the minimum circuit feature should have been 100 nm by 2007. The actually achieved minimum circuit feature in mass production is 65 nm, which is one full generation ahead [4], and volume manufacturing of 45-nm ICs is already on the horizon. From thousand transistors on a chip in 1970s to multi-billion transistors on a chip today, such a feat could only be possible thanks to the constant innovation in microlithography-based IC manufacturing technologies. The minimum feature size in an IC has reduced from 250 nm a decade ago to 45 nm today. Functional complementary metal-oxide semiconductor (CMOS)-based memory IC with minimum circuit feature of 32 nm has been demonstrated in December 2007 [5]. According to the 100-nm dimensional mark, the current generation of IC is already in the nanotechnology regime. Nanofabrication is already taking place in semiconductor manufacturing.

While semiconductor industry has spent billions of dollars to develop more and more sophisticated equipments and technologies to downsize IC features, they have the sole purpose of mass production of ICs. The semiconductor industry has been sticking to optical lithography all along in the past half

century, because it has been the only technology which has the capacity of patterning over a hundred wafers per hour at desired circuit feature dimension. The research community of nanoscience and nanotechnology is, on the other hand, quite content with vast variety of other less expensive nanofabrication tools and techniques. This has been the way that nanofabrication technologies were developed in the last few decades, in parallel to the huge investment in semiconductor manufacturing technologies. It is a fact that academic researchers rarely have the access to state-of-the-art microfabrication or nanofabrication tools, except via semiconductor foundry services with fewer choices of fabrication processes. They often have to work with whatever they have got, trying to make various nanostructures at affordable cost. Many new techniques and processes have therefore been developed. Low-cost nanofabrication has been the main feature of these technologies. One of the excellent examples is the development of nanoimprinting technology. The technology was developed in the mid-1990s for patterning sub-100 nm-scale structures in laboratories where expensive optical lithography tools could not be afforded [6]. The nanoimprinting technique has the parallel patterning capability of optical lithography, but much less expensive than an optical lithography system. Since then, many variants of the technology have been developed, such as working at room temperature and low pressure, using UV-curable polymers, and using soft printing masters [7], etc. They may not be as good as the state-of-the-art optical lithography technology. However, the low-cost feature certainly outweighs its shortcomings when it is used in a research laboratory.

Nanofabrication is not equivalent to nanomanufacturing. The two distinctive features of any manufacturing technology are volume and yield. The typical example is the manufacturing of ICs. To meet the volume requirement, processing has to be parallel. This is the sole reason that optical lithography has survived through so many generations of ICs. Other alternative technologies will have to be able to pattern features in parallel fashion, whether it is done by charged particles (electron or ion projection lithography) or by stamps (nanoimprinting lithography). Though optical lithography is already able to pattern ICs at 32-nm feature size, there are very few technologies available as candidates further down the dimensional scale to succeed optical lithography. Another key issue is that any nanofabrication technology that can be qualified as a manufacturing technology will have to establish a complete supporting infrastructure. For example, optical lithography cannot stand alone to serve the IC manufacturing industry. There have to be photoresist suppliers, photomask manufacturers, and a whole set of inspection/characterization technologies to accompany the optical patterning technology. The same applies to the other so-called “next-generation lithography” (NGL) technologies, such as extreme UV (EUV) lithography [8], step-and-flash nanoimprinting lithography (SFIL), and maskless lithography (ML2) [9]. They all have to have the supporting infrastructure in place before they can enter the IC manufacturing arena. To meet the yield requirement, a nanofabrication technology has to be reliable, repeatable, and have the lowest defect level. Few of emerging nanofabrication

technologies are able to meet this requirement. Even a mature technology may stumble over the reliability criterion. A good lesson was learnt from X-ray lithography. X-ray lithography is by nature a parallel patterning technique. Its much shorter wavelength is ideal for succeeding optical lithography as the next-generation patterning technology for IC manufacturing. However, the difficulty in making X-ray masks and related manufacturing reliability issue finally sealed the fate of the technology.

Though very few nanofabrication technologies would eventually become true nanomanufacturing technologies, it has not stopped many new nanofabrication processes being developed every year. Nanoscience researches require nanostructures to be made. In these applications, volume and yield are not critical issues. This is the area where creativity and ingenuity have led to many new and unconventional ways of making a nanostructure, whether directly or indirectly. A noticeable example is the use of scanning probes for various types of nanofabrication, either optically, electronically or chemically, mechanically [10]. For a few nanostructures or devices, a scanning probe system is a good alternative, with easy setup and guaranteed nanoscale resolution. At below 10-nm scale, all other nanofabrication technologies have seen their limits. Molecular self-assembly becomes the new force in the field [11]. Though there is a long way for these unconventional nanofabrication technologies to become conventional, they definitely have their niche places in the nanoscience and nanotechnology research and development community, and who knows that some of the technologies may one day be developed into industrial-scale manufacturing technologies.

1.3 Purpose of the Book

This book is not intended to be a comprehensive handbook for all the fabrication technologies that have been invented for making microscale or nanoscale structures. Many books have been published to cover various aspects of microfabrication and nanofabrication technologies, including a book published in 2006 by the author [12]. Most of the books published so far have been primarily dedicated to describing various microfabrication technologies, though methods for making sub-100 nm structures were mentioned here and there. For anyone who is interested in nanotechnology and would like to know more about how nanoscale structures and devices are made, there are confusions as to where to find the appropriate technologies that are truly capable of making structures at sub-100 nm dimensional scale, because the descriptions of most of these nanofabrication technologies are buried and mixed with other microscale fabrication technologies.

The present book aims to present a one-stop description at introduction level on most of the technologies that have been developed which are capable of making structures below 100 nm. These include conventional photon-based lithography (Chapter 2), charged beam-based lithography such as electron

beam and ion beam (Chapter 3), and conventional pattern transfer technologies such as deposition and reactive-ion etching (Chapter 6). The book also gives considerable space for a whole set of unconventional nanofabrication technologies, such as nanofabrication by scanning probes (Chapter 4), nanofabrication by replication of stamps (Chapter 5), and various ingenious indirect nanofabrication techniques that bring structure dimension right down to sub-10 nm scale (Chapter 7). The final chapter of the book is devoted to nanofabrication by self-assembly which is emerging as the ultimate route for construction of nanodevices and nanosystems (Chapter 8). The principle and capability of each technology, specifically for making sub-100 nm structures, is introduced and illustrated, with minimum mathematics involved, because it is the author's belief and experience that learning microfabrication or nanofabrication is very much a practical and hands-on process. It is important to know the first principles. The rest is very much down to practicing and mastering the various technologies by using them.

Though there are vast number of literature and published research papers which give details of each nanofabrication technology, they tend to present a positive picture of the technology rather than pointing out what are the shortcomings or telling the negative side of the story. A nanofabrication practitioner knows from his/her own experience that there are no perfect one-suits-all fabrication technologies. Each technology has its own weakness. For example, the state-of-the-art optical lithography, though capable of sub-100 nm patterning, is too expensive and not easily accessible; electron beam lithography is of high resolution and high flexibility but low throughput; scanning probes are simple and low-cost systems for sub-100 nm patterning, but the pattern area is small, speed is low, and patterning depth is shallow; nanoimprinting is best suitable for making high-density periodical patterns; soft lithography is limited in sub-100 nm resolution because of the softness of stamp and lateral diffusion of inks; molecular self-assembly has to be guided in order to make any useful pattern structures, etc. The present book tries to present a balanced view. For each technology introduced, both its advantages and weaknesses are described, so that the readers are made aware of both positive and negative aspects of a nanofabrication technology and aware of its limits. There are carefully selected references at the end of each chapter to help readers of the book to explore further for more detailed information. The book is intended to be a practical guide and first-hand reference for anyone who wishes to learn about different nanofabrication techniques or to consult with when facing to choose a suitable technique from a "tool box" for a specific application.

References

1. *Harvard's George Whitesides on Nanotechnology: "A Word, Not a Field"*. Science Watch, 2002 (July/August). **13**(4).
2. *US National Nanotechnology Initiative (NNI)*. [cited; Available from: http://www.nano.gov/].

3. *International Technology Roadmap for Semiconductors (ITRS)*. [cited; Available from: http://www.itrs.net/].
4. Mack, C., *Predicting the Future in the Past*. Semiconductor International, 2007 (September).
5. *TSMC Reports Foundry's First 32-Nanometer Technology with Functional SRAM*. Semicopnductor International, 2007(December).
6. Chou, S.Y., P.R. Krauss, and P.J. Renstrom, *Imprint of sub-25 nm vias and trenches in polymers*. Appl. Phys. Lett., 1995. **67**(21): p. 3114.
7. Xia, Y. and G.M. Whitesides, *Soft lithography*. Angew. Chem. Int. Ed., 1998. **37**: pp. 550–575.
8. Bjorkholm, J.E., *EUV lithography—the successor to optical lithography?* Intel Technol. J. 1998. **Q3'98**: p. 1.
9. Colburn, M., et al., *Step and flash imprint lithography: A new approach to high-resolution patterning*. Proc. SPIE 1999. **3676**: p. 379.
10. Tseng, A.A., A. Notargiacomo, and T.P. Chen, *Nanofabrication by scanning probe microscope lithography: A review*. J. Vac. Sci. Technol., 2005. **B23**(3): p. 877.
11. Zhirnov, V.V. and D.J.C. Herr, *New frontiers: Self-assembly and nanoelectronics*. IEEE Comput., 2001. **34**(1): pp. 34–43.
12. Cui, Z., *Micro-Nanofabrication Technologies and Applications*. 2006, Springer.

Chapter 2
Nanofabrication by Photons

2.1 Introduction

Since the planar processing for integrated circuits (ICs) manufacturing started nearly half century ago, optical lithography, often called photolithography, has become the convenient choice for making planar microstructures. In optical lithography, a mask or photomask, also called reticle, is imaged onto a flat substrate surface coated with a thin layer of polymer material called photoresist. The photon energy is focused onto the photoresist, causing polymer chain scission or cross-linking. The mask pattern is then delineated into the photoresist after development.

There is a whole spectrum of photons, as shown in Fig. 2.1, which can be explored for lithographic patterning. In the early days when the features size of ICs were a few micrometers, optical lithography was done with visible light and was no different from copying photos in a photography workshop. As the circuit feature dimension shrank, illuminating light with shorter wavelength was employed. This was the time when mercury lamps were used as the illumination source, with UV-wavelength emission at G-line ($\lambda = 436$ nm) and I-line ($\lambda = 365$ nm) being the choices of wavelengths for exposure of photoresists. In the early days of IC manufacturing, 1-μm feature dimension had once been considered as a formidable barrier for optical lithography using the aforementioned illuminating wavelength [1]. Then excimer lasers at deep UV (DUV) wavelength $\lambda = 248$ nm (KrF excimer laser) and $\lambda = 193$ nm (ArF excimer laser) became the new illumination sources for optical lithography. New photoresists, particularly the chemically amplified (CA) resists, were developed to suit the shorter-wavelength irradiation. Next from 193 nm further going down the wavelength were vacuum UV (VUV) at 157 nm, extreme UV (EUV) at 13 nm, and X-ray at 1 nm. The technical difficulty inevitably increases with the shorter-wavelength illumination being used for optical lithography. Some of technical barriers have been insurmountable or the solutions were too expensive, as there is always an economical leverage in determining the champion technology. The time came when a new limit at around 130 nm was predicted for the exit of optical lithography and the entry of next-generation lithography (NGL) techniques. However, optical lithography has marched on.

Z. Cui, *Nanofabrication*, DOI: 10.1007/978-0-387-75577-9_2,

© Springer Science+Business Media, LLC 2008

With many improvements and innovations, such as a series of resolution enhancement techniques (RETs) and using high numerical apertures (NAs), optical lithography was able to achieve patterning circuit structures well below the illumination wavelength and is now poised to go down the dimensional scale as far as 32 nm. Apart from the mainstream development in optical lithography, there was always an army of researchers who were trying to find low-cost alternatives: interferometric lithography and near-field optical lithography are among the few. These techniques are trying to beat mainstream optical lithography not only on the cost front but also in delivering sub-100 nm patterning capability. There is also the maskless optical lithography technique, which promises to deliver sub-100 nm patterning without using the increasingly expensive masks.

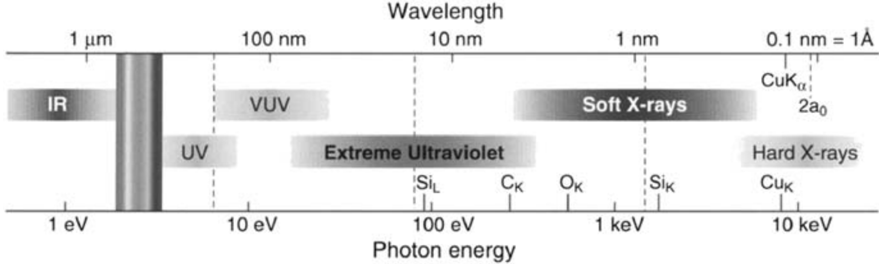


Fig. 2.1 Complete spectrum of photon wavelength (Reprint from [2] with permission)

Optical lithography has been the subject of many published literatures including a book published by the author [3]. The basic issues of optical lithography have been well described in these literatures. This chapter will not simply repeat those basics but review the photon-based lithography techniques with emphasis on those issues related to producing sub-100 nm structures on a photon-sensitive polymer layer. While the basic principles of various optical lithography techniques are still being mentioned, most of the technical details about the basic optical lithography processes have been omitted. The content has been structured around the central theme of sub-100 nm patterning, the capabilities and limits of current photon-based lithographic techniques, and the ways of overcoming these limits.

2.2 Principle of Optical Projection Lithography

Although IC manufacturing started in the early days with contact optical lithography, it soon migrated to projection optical lithography. In contact mode, an optical mask is in direct contact, or having a proximity gap, with a photoresist-coated substrate, and features in the mask are imaged onto the

photoresist in 1:1 ratio without reduction. Projection optical lithography can project a demagnified image onto photoresist (5:1 or 10:1), capable of patterning much smaller features than contact optical lithography as well as eliminating the mask damage and contamination problems associated with contact optical lithography. For submicron patterning, projection optical lithography prevails.

The basic projection optical lithography system is schematically shown in Fig. 2.2, where the mask is illuminated by a light source and the light passing through the transparent patterns in the mask is focused onto the photoresist layer. The theory of optical imaging in projection photolithography can be very complicated. A comprehensive mathematical description of the theory is presented in a book written by Alfred K. Wong [4]. However, the concept of projection optical lithography can be understood easily. A simple conceptual view of diffraction-limited projection optical lithography is shown in Fig. 2.3.

A mask feature, which is a transparent opening, can be considered as a slit from which passing light is diffracted in the same way as a Young's single-slit diffraction. Diffraction is a phenomenon which involves the bending of waves

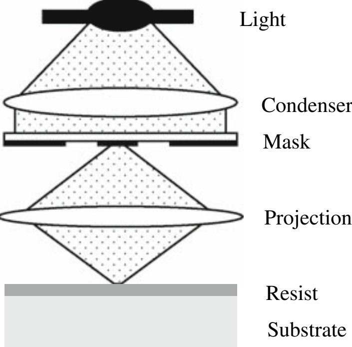


Fig. 2.2 Schematic of projection optical lithography

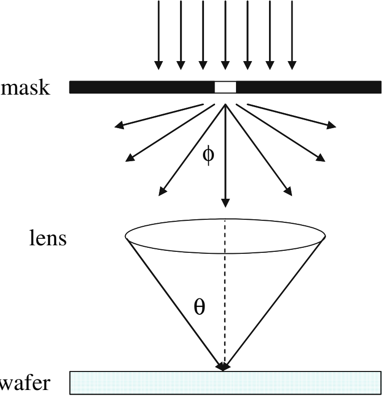


Fig. 2.3 Conceptual view of optical projection imaging

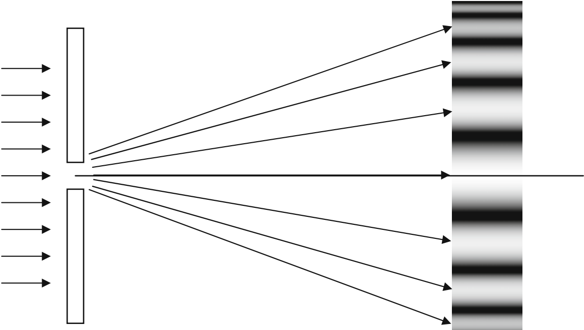


Fig. 2.4 Single-slit diffraction

around obstacles. It is described by the Huygen's Principle that every point on a wave front acts as a source of tiny wavelets that move forward with the same speed as the wave. These wavelets interfere with each other, causing light intensity variations with maxima and minima which are bright and dark fringes, as shown in Fig. 2.4. The angle of each dark fringe relative to the source can be determined by a simple equation (Eq. 2.1):

$$sin\phi = m\frac{\lambda}{a}, m = 1, 2, 3, \ldots\ldots \quad (2.1)$$

In Eq. 2.1, ϕ is the diffraction angle of light passing through the slit, λ is the wavelength of light, and a is the width of the slit. It is apparent that the spread or expansion of light, represented by the diffraction angleϕ, is the function of wavelength and the width of the slit. For a fixed wavelength, the smaller the slit the wider the spread of light, as shown in Fig. 2.5a, where diffracted light intensity distributions are plotted for three different widths of the slit. On the other hand, light with shorter wavelength can have narrower distribution of diffraction, as shown in Fig. 2.5b. Because of the wide spread of light from a smaller slit opening, the final image occupies much larger area on an image plane than it should be for a wider slit. Referring the slit to a feature opening on a mask, if two of such slits are close to each other, then they may merge into each other, making them irresolvable. Figure 2.6a shows exactly such a case where two light images have merged into one because the slits where the lights come from are too close to each other. They can only be resolved as separate spots provided the two slits are well apart, as shown in Fig. 2.6b

Light diffraction from a slit, or a mask feature, is only half of the process in a projection optical lithography. Referring to Fig. 2.2, the light that comes out of a mask opening has to be focused by a lens system onto a substrate surface. An optical lens is always of limited diameter. The diameter of a lens aperture determines how much or how wide the diffracted light from the mask can be accepted into the lens system. The lens diameter is defined by the NA, expressed

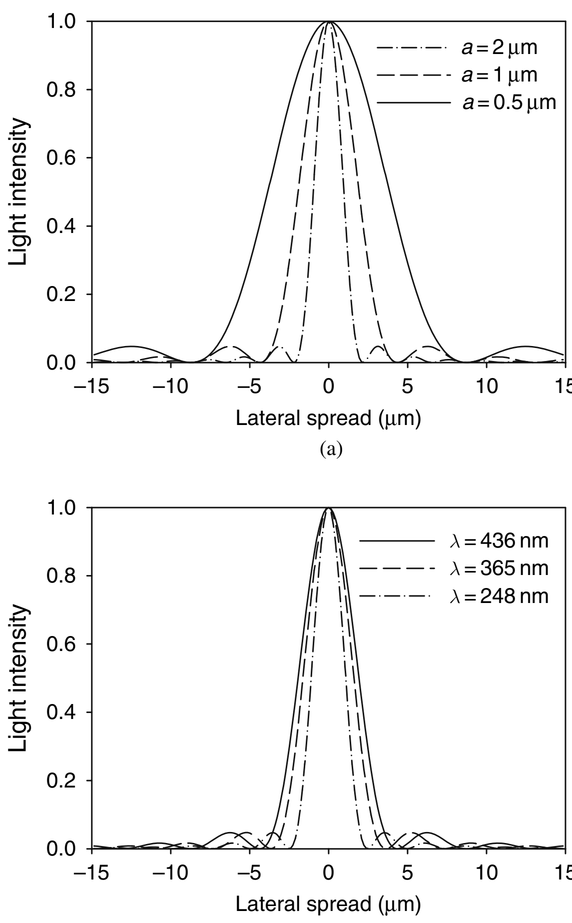


Fig. 2.5 Light intensities of three single-slit diffractions: (**a**) at 436-nm wavelength, and (**b**) for 1-μm slit width

as Eq. 2.2, where n is the refractive index of the medium between the lens and an imaging plane, and θ is the half-angle of the maximum cone of light that can enter or exit the lens or the angle of outmost ray of light entering or exiting the lens:

$$\mathrm{NA} = n \sin \theta. \tag{2.2}$$

Skipping the complicated optical imaging theory, a simple equation can define the resolution limit of a projection lithography system, as expressed in Eq. 2.3, where λ is the illumination wavelength, NA is the numerical aperture of imaging optical system, and k_1 is a factor related to a specific imaging process ($k_1 < 1$):

$$R = k_1 \frac{\lambda}{\mathrm{NA}}. \tag{2.3}$$

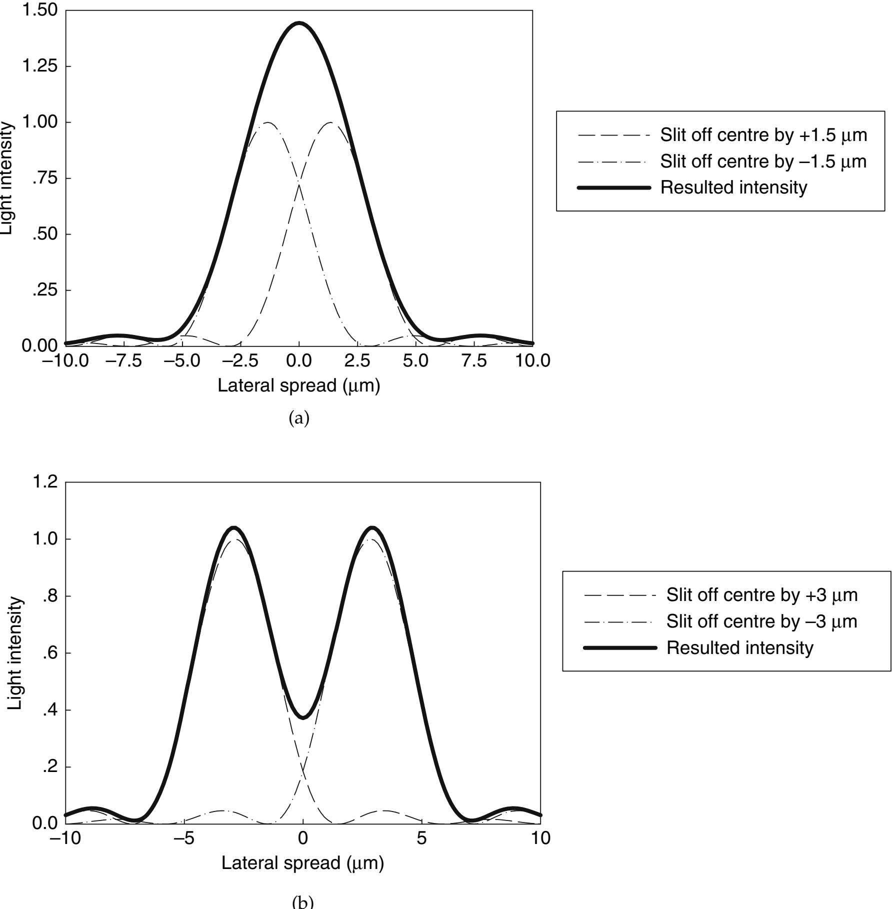


Fig. 2.6 Light intensities of diffraction from separated slits: (**a**) irresolvable image, and (**b**) resolvable image

From the above equation, one can intuitively work out that shorter wavelength results in narrow distribution of light diffraction, as proved in Fig. 2.5b. Therefore, more lights can be focused by the lens of fixed diameter (NA). The same applies to larger NA at a fixed wavelength; that is, more diffracted lights can be admitted into the lens and focused onto a substrate if the NA of lens is bigger. Both can result in improved resolution in imaging, because higher-ordered diffractions (light of large diffraction angle) passing the lens result in more details of the mask feature to be imaged onto the substrate. The effects of wavelength and NA can be quantitatively evaluated by computer simulation of partially coherent imaging in projection optical lithography. Figure 2.7a and b are the simulated light intensity of a mask feature of 0.5-μm linewidth at

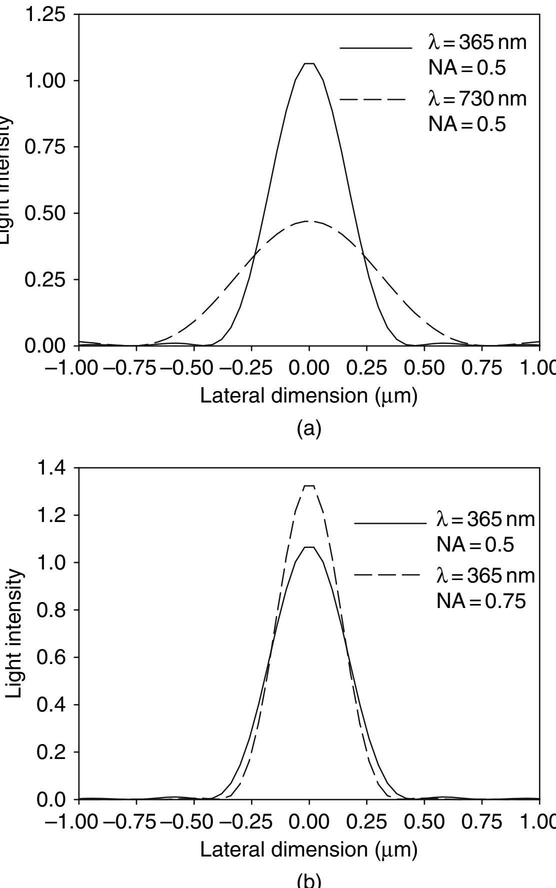


Fig. 2.7 Computer-simulated light intensities imaged at (**a**) different wavelengths and (**b**) different NAs

different illumination wavelengths and different NAs. The simulation was performed by the optical lithography simulation program COMPARE [5].

The k_1 in Eq. 2.3 is a factor that can further enhance the resolution of the imaging process by the collective action of many techniques. The following sections will describe how the resolution of optical lithography has been improved over the years by using shorter wavelengths, larger NAs, and reduction of the k_1 factor.

2.3 Optical Lithography at Shorter Wavelengths

2.3.1 Deep UV

In projection optical lithography, the mask is the object to be imaged onto a substrate which is the image plane. The light passing through the mask is the

light from the illumination source in an optical lithography system. Therefore, the wavelength appearing in Eq. 2.3 is the wavelength of illumination source. In the early days of optical lithography, mercury lamps were the illumination source. The spectra line at wavelength of 436 nm, called G-line, was first utilized for illuminating the mask. As the mask features shrank, shorter wavelength at 365 nm (I-line) became the choice of illuminating source. Further demands in imaging small mask features made the mercury lamp unsuitable as the illumination source, because there is simply not enough photon energy at shorter wavelength to meet the requirement of volume production of ICs.

Fortunately excimer lasers can meet the requirements of both high photon energy and shorter wavelength. Excimer lasers are pulsed gas discharge lasers which produce light output in the ultraviolet region of the spectrum. The commonly used excimer laser sources are listed in Table 2.1, each emitting different wavelength of UV light.

Table 2.1 Excimer lasers and their relative power

Wavelength	Active gases	Relative power
157 nm	Molecular Fluorine (F_2)	10
193 nm	Argon fluoride (ArF)	60
248 nm	Krypton fluoride (KrF)	100
308 nm	Xenon chloride (XeCl)	50
351 nm	Xenon fluoride (XeF)	45

The KrF excimer laser has the highest relative power and was the obvious choice of illumination source to go down the scale of wavelength for volume IC manufacturing. The introduction of excimer laser marked the beginning of DUV optical lithography. The next illumination source for DUV optical lithography was the ArF excimer laser at 193-nm wavelength. It could have gone down even shorter wavelength with excimer lasers. There are the molecular fluorine (F_2) excimer laser emitting photons at 157-nm wavelength and Ar_2 excimer laser at 126-nm wavelength. The semiconductor industry started concerted effort from 1998 to develop 157-nm optical lithography technology, aiming at IC feature size of 70 nm.

For optical lithography at 157-nm wavelength, there are three critical technology issues: material for making optical lenses, transparent and radiation-durable pellicle for masks, and photoresists. All these technical barriers stemmed from high absorption of photons at short wavelengths such as 157 nm or below. The lens materials at 248-nm and 193-nm wavelengths used to be fused silica, which is no longer suitable at 157 nm because of the high absorption or poor transmission. The choice of low-absorption material at 157 nm is CaF crystal. As CaF is no longer a glassy material like the fused silica, the crystalline nature of the material can create a phenomenon called "birefringence", or double refraction, which is the decomposition of a ray of

light into two rays (the ordinary ray and the extraordinary ray) when it passes through the crystalline material [6]. For lenses made of crystalline material, the intrinsic birefringence has two main effects. The first is different refraction for the two polarization components at the lens surfaces, which results in a ray bifurcation at each lens. The second is that each polarization component accumulates a different phase traversing the birefringent crystal, resulting in a phase-front distortion. The second effect is considered more troublesome for lithography. Both effects contribute to blurring of the image, limiting the achievable resolution. In order to reduce the birefringence to its minimum, the growth of CaF crystal has to be controlled as isotropic as possible, to be more like a glassy material than a crystal material. This has turned out to be an insurmountable task.

The high-absorption nature at 157-nm wavelength has prevented the use of pellicle which is a thin film stretched on a frame used to protect a photomask from particle contamination. In modern IC manufacturing pellicle has become an integral component of photomask. Without pellicle protection, a photomask can easily get particles on its surface and form distorted images on the wafer, creating defects on a chip. Conventional pellicle materials are too absorptive and not durable at high dose of 157-nm laser irradiation. The same problem applies to photoresists. Resists designed for 248-nm and 193-nm lithography are too absorptive at 157 nm. New resist materials have to be developed.

Given the huge efforts needed to make the 157-nm wavelength optical lithography production worthy and the fact that 157 nm can only last one generation of IC manufacturing, the semiconductor industry had to ponder whether to pursue 157 nm or to go directly to much shorter wavelength which is the EUV at 13 nm. The picture became clear in 2004 when immersion optical lithography at 193 nm started to mature and could well take over 157-nm wavelength for 70-nm IC manufacturing or even below [7]. Optical lithography at 157 nm became unnecessary. On the other hand, EUV could play a key role in several generations of IC ahead from 32 nm to 22 nm or below, because of its much shorter wavelength. With semiconductor industry major players' backing, such as Intel, the effort to develop shorter-wavelength optical lithography went to EUV, instead of the 157-nm excimer laser. Table 2.2 has listed the minimum IC feature sizes in the history of semiconductor industry development, and the corresponding wavelengths of optical lithography. There is a big jump from 193 nm to 13 nm. Entering 13-nm wavelength regime, optical lithography has stepped into a completely new era.

Table 2.2 Reduction of wavelengths in optical lithography

Minimum feature size (nm)	7000–1000	1000–350	350–180	180–32	32 and below
Lithography wavelength (nm)	436 (G-line)	365 (I-line)	248 (DUV)	193 (DUV)	13 (EUV)

2.3.2 *Extreme UV*

Extreme UV at 13-nm wavelength is in strict sense no longer optical irradiation. It is often called soft X-ray. To differentiate it from hard X-ray lithography, the semiconductor industry still considers EUV lithography (EUVL) as an optical lithography. In many aspects, EUVL retains the look and feel of optical lithography as practiced so far. However, conventional refractive optics ceases to function at EUV because of its strong absorption in virtually all materials. Instead, reflective optics has to be used for both mask and focusing optics. The basic system of EUVL is schematically shown in Fig. 2.8. It shows that the EUV radiation from an EUV source is collected by a set of reflective mirrors (illumination optics) and projected onto a reflective mask. The reflected mask image is then focused by another set of reflective mirrors (projection optics) and projected onto a substrate to expose the EUV resist layer. In addition, EUV imaging must be carried out in a near-vacuum condition to reduce atmospheric absorption of EUV light.

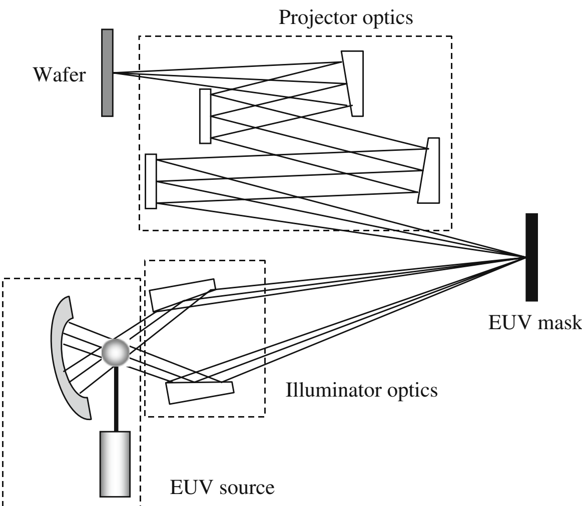


Fig. 2.8 Schematic of EUVL system

There are four critical components in EUVL: EUV source, EUV projection optics, EUV mask, and EUV resist. The state-of-the-art developments of the four critical components are described as following.

2.3.2.1 EUV Source

Extreme UV radiation can be generated by two methods: plasma and synchrotron radiation source (SRS). Synchrotron radiation source is a large-scale instrument which can only be built and owned by a national government and is mainly for fundamental scientific research. Industrially dedicated EUV sources for lithography purpose are mostly based on plasma-generated EUV radiation. There are two

types of plasmas that can generate EUV-wavelength radiation: laser-produced plasma (LPP) and discharge-produced plasma (DPP). Some elements such as Li, Xe, and Sn, after ionized, can emit short-wavelength photons when electrons in the atomic structure jump from higher energy band to lower energy band. For example, ionized Li atom has a radiation spectrum from 10 nm to 22 nm, with the highest intensity of spectra line at 13.5 nm [8]. This has been the reason that the central wavelength for EUVL is fixed at 13.5 nm. There are a number of key technical issues in the development of plasma-based EUV sources. So far the most important issue has been how to achieve high power in order to meet the exposure throughput of 120 wafers per hour. This is equivalent to EUV power of 115 W at intermediate focus [9]. This seems to be a formidable task because the current EUV sources can only achieve power level of 10 W at intermediate focus. It was projected to reach 30 W in 2007 and 115 W in 2009 [10]. It is still uncertain whether the 115-W target is achievable by 2009. Output power of EUV source is only part of the problem. There are other critical issues, such as the lifetime of collector optics due to contamination caused by debris generated in the source [11].

2.3.2.2 EUV Optics

The key component of EUV optics is the multilayer reflective mirror. Extreme UV reflectivity of any single material at near-normal incidence is very low. In order to achieve reasonable reflectivity at near-normal incidence, the mirror surface must be coated with multilayer thin-film coatings known as distributed Bragg reflectors. These coatings consist of a large number of alternating layers of materials having dissimilar EUV optical constants, and they provide a resonant reflectivity when the period of the layers is approximately $\lambda/2$ of EUV wavelength. The most developed and best understood reflective multilayers for EUV are made of alternating layers of Mo and Si, and they have the highest reflectivity at wavelengths of about 13 nm. Figure 2.9 shows the theoretical reflectivity of a 50-pair Mo(2.76 nm)–Si(4.14 nm) multilayer mirror at normal incidence of EUV [12]. In practice, around 68% reflectivity can be achieved by a multilayer Mo–Si mirror. The causes of decrease in reflectivity are absorption of EUV energy by impurities in the deposited film, index change due to formation of interdiffusion of layers, and scattering due to interface roughness. The major issues for EUV optics are the contamination control and lifetime under EUV irradiation.

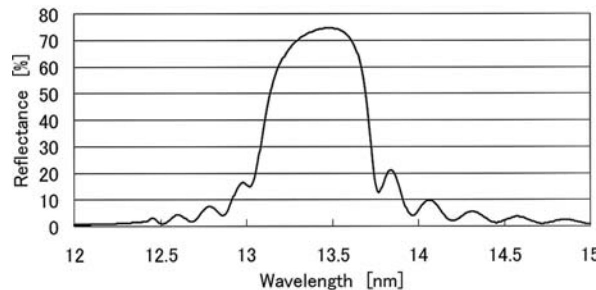


Fig. 2.9 Theoretical reflectivity of EUV radiation from multilayer Mo–Si mirror

2.3.2.3 EUV Mask

Extreme UV mask works the same way as the EUV mirror. It has to be reflective. There are two stages of processes involved in making EUV masks: first is the making of the mask blank and second is the patterning of the absorber layer. Figure 2.10 shows the cross-section of a EUV mask. The blank is similar to a EUV mirror except that there is an absorber layer on top of the multiple pairs of Mo–Si layers. The substrate must be a low thermal expansion material (LTEM) with flatness better than 50 nm and free of defects. A number of metals can be used as the absorber layer, such as Al, Cr, Ta, and W. Because EUV projection optics is reduction optics with reduction ratio of 4:1, the patterning of EUV absorber is less demanding. Conventional electron beam lithography and reactive-ion etching (RIE) are adequate for the absorber patterning. The most difficult part in EUV mask manufacturing is the defect inspection and repair. There are two groups of defects: the defects in a mask blank and defects in patterned absorber. Because EUV imaging relies on reflective surface of mask, any bumps can cause phase error in the final image. It has been characterized that a 65 nm × 1.5 nm (width × height) bump in the multilayered surface can cause 20% deviation of dimension to a 25-nm line [13]. Defects as small as 50 nm on the absorber patterns can be printable [14]. For repair of defects in conventional optical mask, focused ion beam (FIB) has been the universally used repair tool. However, FIB repair of opaque defect in EUV masks can easily cause both damage and gallium staining to the substrate, and gallium is very absorbent at EUV wavelength of 13.5 nm. In this case, electron beam repair method is preferred. Electron beam irradiation assisted by reactive gases can either etch or deposit materials with high resolution and high accuracy, which makes a perfect tool for defect repair in EUV masks [15].

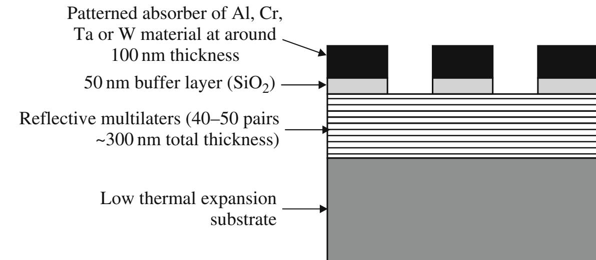


Fig. 2.10 Cross-section view of EUV mask

2.3.2.4 EUV Resists

Resists for EUVL are very much like those for DUV lithography, but required to have higher sensitivity and higher resolution capability, as well as low line edge roughness (LER). High sensitivity is required because of the low power level of EUV source. Even at EUV source power of 115 W (intermediate focus),

resist sensitivity better than 3 mJ cm^{-2} is necessary to ensure throughput of 100 wafers per hour. Such sensitivity is only possible for CA resists. However, the higher the resist sensitivity is, the greater the resist LER becomes. Some recent EUV exposure results showed 35-nm and 40-nm line/space patterns but with unacceptable LER [16]. Line edge roughness has been a critical issue for all sub-100 nm lithography, which will be discussed later in Section 2.5.4. For EUVL at 32-nm technology node and below, the LER has to be below 2.5 nm. The goal for EUV resists is to have sensitivity below 2.5 mJ cm^{-2} and LER below 2.5 nm. Unfortunately most of existing CA resists cannot satisfy both requirements, as shown in Fig. 2.11. Due to the difficulty encountered in improving EUV source power level, EUV resist development has become the top critical issue since 2006, in order to meet the target of inserting EUV into IC production by 2009. While the resist sensitivity and LER have to be improved simultaneously, another issue cannot be ignored, that is, the outgassing of resist during EUV exposure. Gaseous molecules released from resist upon EUV irradiation can contaminate the EUV mirror surface, hence degrading the EUV optics and reducing its lifetime.

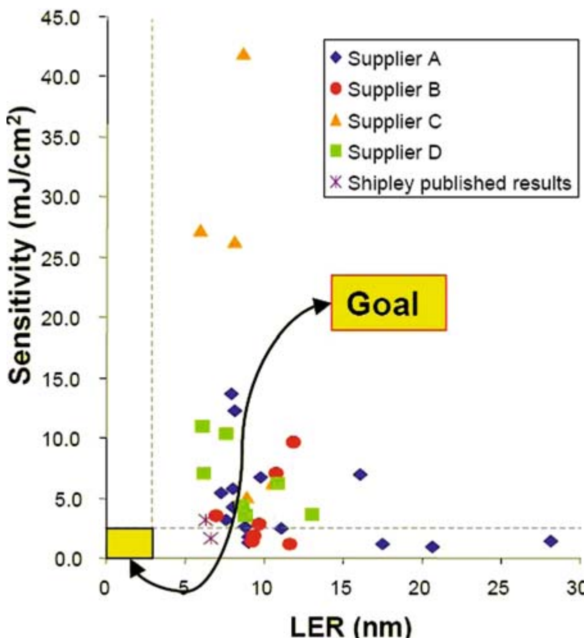


Fig. 2.11 Sensitivity vs LER for EUV resists (Reprint from [17] with permission)

Although EUV has been projected to become the main lithography technology in 2009 at 32-nm generation of IC manufacturing, there are still a series of formidable technical hurdles ahead. In addition, economics is another deciding factor. An industrial EUVL system can cost up to $50 M, compared with a

DUV lithography system which is about $20 M. The cost of ownership (CoO) is expected to be high too, which includes the cost of equipment as well as the cost of running the equipment. The CoO of EUVL will not be competitive, compared to 193-nm DUV lithography, unless a comparable or higher throughput can be achieved [18]. On the other hand, 193-nm immersion lithography is rapidly catching up with a range of innovations, as will be described in Section 2.4. Extreme UV lithography is now facing the threat of ever becoming a viable IC manufacturing tool.

2.3.3 X-ray

The ultimate short-wavelength photon radiation would be the X-ray. From the photon spectrum shown in Fig. 2.1, X-ray is the radiation wavelength spanning from 1 nm at the soft X-ray end toward 0.1 nm of hard X-ray. The fundamental difference between X-ray wave and optical wave is that X-ray can penetrate the majority of materials. Only materials with high atomic number can absorb X-ray. X-ray cannot be focused because the refractive index of all materials to X-ray is the same ($n \approx 1$). Therefore, no matter whether X-ray enters from free space to a material or from one material to another, there is no refraction. X-ray exposure of mask image can only be 1:1 with a proximity gap between mask and resist surface.

The proximity X-ray lithography is illustrated in Fig. 2.12, where an X-ray mask stops or lets through X-ray. The X-ray mask consists of a membrane of low-atomic-number material, such as silicon or silicon carbide, with patterned

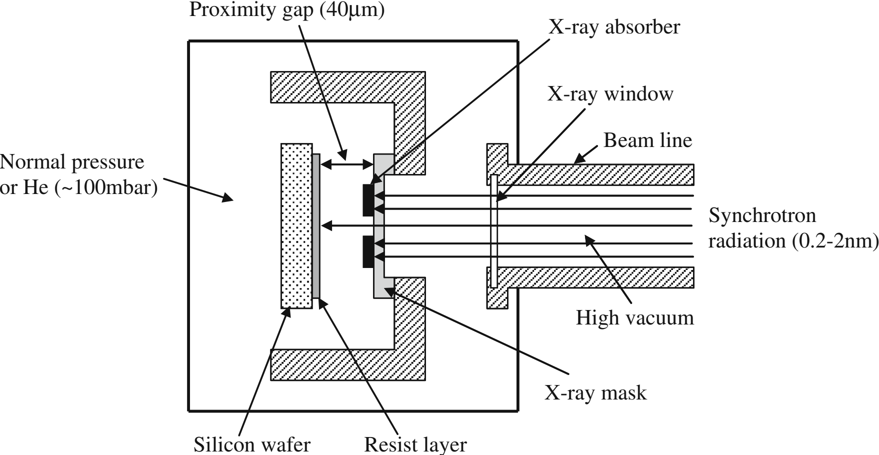


Fig. 2.12 Schematic of X-ray lithography

high-atomic number material as the absorber on top. Low-atomic-number materials are transparent to X-ray, while high-atomic number materials, such as Au, W, Ta, and other heavy metals, can effectively block X-ray. For the X-ray of 1-nm wavelength, the silicon membrane is of 1–2-μm thickness and absorber thickness is in the range of 300–500 nm. The exposure depth in resist will be about 1 μm and the proximity gap between X-ray mask and substrate is in the range of 5–50 μm [19]. The proximity gap is necessary because of the poor mechanical strength of thin-membrane X-ray mask. The large thin membrane is very fragile and cannot have any mechanical contact with the exposure surface, which rules out the possibility of contact lithography.

Because of its short wavelength, X-ray lithography once was a very hopeful candidate technology to replace optical lithography for IC manufacturing. IBM started the technology development in 1980. Motorola, NEC, NTT, and Mitsubishi all had invested in the technology. The beginning and mid-1990s were the heydays of X-ray lithography R&D. As early as in the mid-1990s when optical lithography was still struggling to get down to 0.25-μm circuit dimension, X-ray lithography had already achieved 0.12-μm experimental Dynamic random access memory (DRAM) chips and 0.1-μm logic chips [20]. However, X-ray lithography eventually did not succeed to replace optical lithography. The biggest hurdle the high-resolution X-ray lithography faced is the fabrication of mask. As X-ray exposure is 1:1 pattern transfer, the heavy metal absorber has to have the same dimension as the resist feature. In addition, local stress caused by heavy metal absorber and internal stress of supporting membrane all can have influences on the pattern resolution. Besides, any defects in absorber pattern will have a serious effect on exposure patterns. In optical lithography, mask defects smaller than illumination wavelength will not be imaged on resist. In X-ray lithography, any mask defects will be transferred to resist because of the 1:1 imaging process.

Another challenge for ultra high-resolution X-ray lithography is the accurate control of mask–wafer gap. High-resolution X-ray lithography requires the gap to be controlled within 10 μm. Variation of the gap will cause variation of image resolution. This requires the flatness of silicon wafer to be within ±0.25 μm and flatness of the mask within ±0.5 μm [21]. These are very stringent requirements for a large-size silicon wafer and a large area of membrane. The gap control is also required to be within ±0.5-μm accuracy. Bearing in mind that the mask membrane is only 1–2-μm thick and very easy to crack, it makes the gap control a lot harder.

It is no doubt that X-ray lithography is capable of sub-100 nm resolution. However, there are some fundamental limits which make X-ray lithography difficult to go below 50-nm resolution. The main limiting factors are photoelectrons and Auger electrons generated by X-ray irradiation in resist, and X-ray diffraction. High-energy X-ray irradiation will generate considerable number of photoelectrons and Auger electrons inside a resist material. These low-energy electrons are more efficient to interact with resist polymer molecules, causing exposure of the resist. Upon acquiring energy from X-ray, these electrons scatter to some distance within the resist layer. The consequence of

electron scattering is the expansion of exposure pattern dimension, hence degradation of lithography resolution. Although X-ray has very short wavelength, there is still diffraction effect which becomes worse for highly coherent X-ray sources. Synchrotron radiation source (SRS) is such a coherent source because the divergent angle of X-ray emission from SRS is normally less than 1 mrad. Diffraction may cause proximity effect in exposed pattern, very much the same as in optical lithography. Reducing the gap between X-ray mask and resist-coated substrate can reduce the diffraction effect. However, it becomes extremely difficult to control the gap below 5 μm for the reason explained previously. The best resolution of X-ray lithography demonstrated in laboratory is about 30 nm [22]. Further improvement in resolution has to rely on playing tricks with diffraction effect [23].

On top of all the technical challenges is the economical factor. For industrial-scale manufacturing, the X-ray "light" has to be from a SRS. Construction of a large-scale SRS costs over $100 M, or a compact-size SRS costs about $20 M. Although one SRS can have 10–20 beam lines equivalent to 10–20 optical steppers, they all rely on proper working condition of the SRS. If the SRS breaks down, all the beam lines will stop working. Had the optical lithography not kept the pace of shrinking down in IC feature dimension, the semiconductor industry would have had no choice but to adopt X-ray lithography. However, optical lithography has been continuously advancing. New resolution record has been broken time and again beyond anyone's dream a decade ago. Fabrication of optical masks is much simpler and easier than fabrication of X-ray masks. X-ray lithography has eventually failed to win favors from the semiconductor industry to become the mainstream manufacturing technology, though it is still being used as a mature microlithography means for small batch production of ICs and applications in many other fields of micro- and nanotechnology, if a SRS is within easy reach.

2.4 Optical Lithography at High NA

The second parameter influencing the resolution of optical lithography, in addition to the wavelength as seen in Eq. 2.3, is the NA. As explained earlier, NA is equivalent to the diameter of a lens. While lights passing through a small mask feature (transparent opening) diffract in wide angle as illustrated in Fig. 2.4, a lens with large diameter can collect more diffracted lights and focus onto the image plane. It is crucial to collect the wide-angle diffracted lights because they contain the information about the details of the mask feature. The light diffraction through a small feature has the characteristic that the finer the details of the feature, the further the lights diffract from the central optics axis; or, in term of Fourier optics, smaller spatial detail can be referred to as being of higher "spatial frequency". Such a characteristic is best explained in Fig. 2.13a and b.

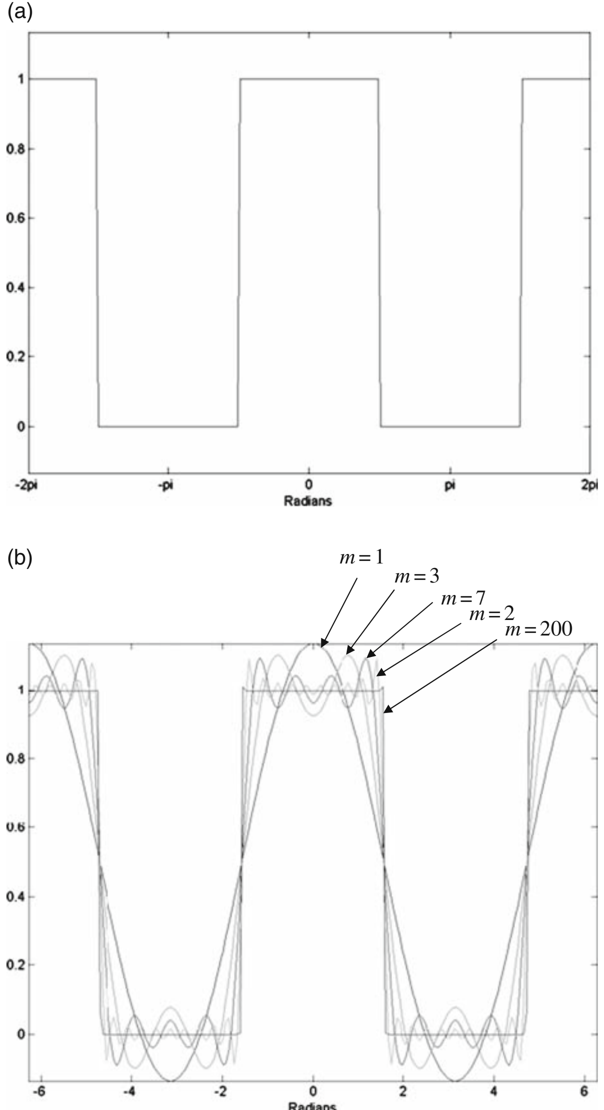


Fig. 2.13 (**a**) Mask image without diffraction. (**b**) Mask image with diffraction. Reprint from [24] with permission

Figure 2.13a is square wave function, representing slits or a series of line features in a mask, which have the ideal distribution of light intensity (without diffraction). After diffraction at the image plane, the square wave function is transformed into a series of sinusoidal wave functions from the base spatial frequency ($m = 1$ in Eq. 2.1) to higher orders ($m > 1$ in Eq. 2.1). The super-imposed sine waves of different spatial frequencies become the image of original

square function (object) at the image plane. The more the high spatial frequency waves are included, the better the final image approximates the original object, as shown in Fig. 2.13b. As seen in the plots, the superimposed image of up to 200 sinusoidal waves is almost identical to the original square wave function. It is therefore understandable that only an optical lens system with larger NA can allow diffracted lights of much higher spatial frequency to reach the image plane, resulting in better image of the mask features. High NA also means more lights coming through the lens system, hence a brighter image on the image plane. Computer simulation of optical imaging can confirm the influence of NA, as shown in Fig. 2.14, where a group of 0.2-μm lines are imaged through a projection optical system at 248-nm wavelength and with three different NAs (NA = 0.5, 0.6, 0.8). The lines imaged at high NA are not only resolved better (high image contrast) but also brighter (high image intensity).

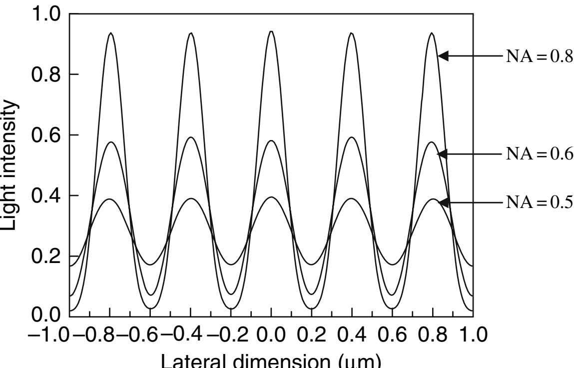


Fig. 2.14 Influence of NA on optical imaging

Although it is so obvious that high NA can improve the resolution of optical lithography, resolution is only one requirement for high-quality optical lithography. Another important requirement is the depth of focus (DOF), or commonly known as depth of field in photography. Depth of focus is also related to the wavelength of illumination and NA of an optical system by Eq. 2.4, where k_2 is a constant ($k_2<1$) related to a specific lithography system and photoresist process:

$$\mathrm{DOF} = k_2 \frac{\lambda}{(NA)^2}. \tag{2.4}$$

The Eq. 2.4 indicates that increasing NA has an effect on reducing the DOF. Figure 2.15 compares the image intensity of the same line features as in Fig. 2.14 at NA = 0.8 but with defocus. If there is a defocus, which can be set at different values in the aerial image simulation, the high NA results in much worse degradation of image quality. As can be seen in Fig. 2.15, the group of 0.2-μm lines is no longer resolvable.

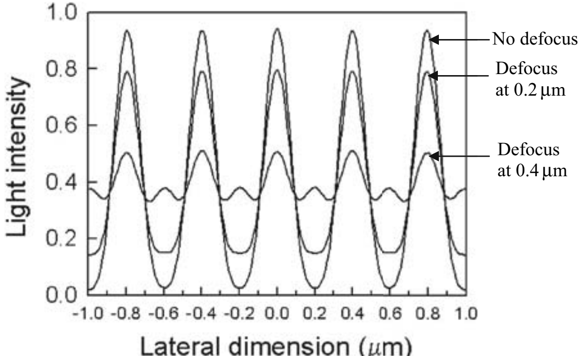


Fig. 2.15 Influence of DOF on optical imaging

The DOF defined by Eq. 2.4 is only valid for NA less than 0.5. For higher NA, the equation has to be modified and takes the form of Eq. 2.5 (where k_3 is a constant similar to k_2), which causes DOF to drop much more steeply at high NA value [25]:

$$\mathrm{DOF} = k_3 \frac{\lambda}{2(1 - \sqrt{1 - \mathrm{NA}^2})}. \tag{2.5}$$

It must be pointed out that DOF is actually a more important parameter in the lithography of ICs. Integrated circuits are always fabricated on large size of silicon wafers, normally in 8″ or 12″ diameter. The wafers of such sizes are impossible to be perfectly flat, plus there are topographies formed by circuit structures in previous process steps. One can imagine that if the DOF is very small the optical imaging can only be in focus in a very small range of height variation. There would be many places in a wafer where defocus would be inevitable because of the topography or wafer bowing.

Every layer of two-dimensional (2D) circuit layout on an IC chip has a group of critical minimum features which is called “critical dimension”, or CD. This CD has to be controlled within a certain range (less than ±10% variation from nominal dimension) during fabrication to ensure the final circuit is working properly. For example, for a 100-nm critical feature dimension, any errors during fabrication should increase or reduce the feature dimension no more than 10 nm. The DOF of a lithography system has to be large enough to ensure the CD variation is within ±10% everywhere on a wafer. For scientific experiments, resolution may be the only important target to achieve, as the researchers are only interested in a small device area where there are a few pattern structures with minimum dimensions. For large-scale production, high yield can only be possible when all the minimum circuit features across the entire wafer are controlled within the target CD variation. In this case, the highest resolution is only meaningful within certain DOF.

The increase in NA used to be restricted to ensure the DOF is above 1 μm at a specific wavelength, which would give much ease to IC processing. With advances in all aspects of optical lithography, particularly in wafer flatness, introduction of auto-focus and auto-leveling systems, thinner resist processes, and reduced surface topography via chemical–mechanical polishing (CMP), the requirement on DOF has been gradually reduced to below 0.5 μm. This has enabled gradual increase of NA over the years. Figure 2.16 shows the evolution of NA in optical lithography from 1980s up to early 2000. While the wavelength has increased from G-line to DUV at 193 nm, the NA also increased from 0.3 to above 0.8. The higher the NA the more complex, bigger, heavier, and more expensive the lens has become. In the 1980s, a G-line stepper lens at 0.35 NA weighed only 14 kg. By the mid-1990s, an I-line stepper lens at 0.63 NA weighed over 500 kg [26].

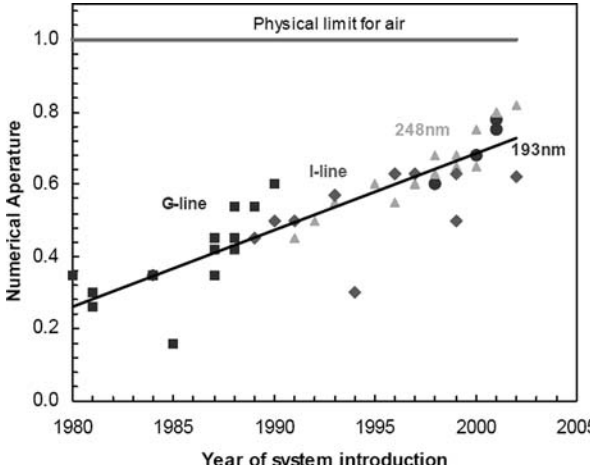


Fig. 2.16 Evolution of NA in projection optical lithography (Reprint from [27] with permission)

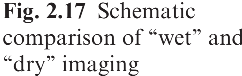
Further increase in NA, particularly above 0.8, will not only encounter the difficulty in lens design/building and leave optical lithography with very small DOF, but also have the problem of polarization effect at high oblique incidence of light, which degrades the image contrast [28]. High NA has forced optical lithography into an optical imaging regime where the approach angles of the imaging rays are so wide that the light coming out of the lens can no longer be considered as a scalar. Instead we must consider the vector components of light and take into account the separation seen between the electric and magnetic field components of light, the so-called s-polarization and p-polarization. When incident light is at a large oblique angle from air into a polymer resist layer, part of the light (s-polarization) can be reflected from the surface if the angle is greater than the Brewster's angle. Reflection of polarized light will reduce the photon energy that goes into the photoresist, causing low contrast of image and

low image intensity. The Brewster's angle θ_B at which the polarization reflection occurs is expressed as Eq. 2.6, when light from one medium with refractive index n_1 goes into another medium with refractive index n_2. In the case of optical lithography, light goes from air ($n_1 = 1$) to a photoresist layer (n_2).

$$tan(\theta_B) = \frac{n_2}{n_1}. \tag{2.6}$$

High-refractive-index photoresist will be able to tolerate imaging with high-NA optics. For example, if a photoresist has refractive index of 1.5, the Brewster angle θ_B will be 56°, which translates into NA of 0.83, based on Eq. 2.2. Theoretically, further increase in NA can be tolerated by increasing the refractive index of photoresist, ignoring the difficulty and cost of making high-NA lenses. However, the refractive index of photoresist is not a property that can be easily adjustable. One way to solve the polarization problem is to illuminate the mask with a deliberately polarized light source, for example, illumination containing only the p-polarization light. In this way, there will be no reflection of light from the photoresist surface, even with very high-NA optics. Improvement in DOF and lithography quality has been observed using polarized illumination at lens NA of 0.92 [29].

The most important development in the history of optical lithography is the introduction of immersion imaging, the so-called immersion optical lithography. Immersion imaging scheme came just at a time when optical lithography has hit the limit at 193-nm wavelength and the semiconductor industry is facing the choice of whether to go for the problematic 157-nm option or to go for the so-called "next-generation lithography" (NGL), which includes EUV, electron projection, ion projection, and nanoimprinting. Immersion imaging is not a new concept. Medical and biological research has for a century used immersion imaging in optical microscope to enhance the image resolution. It is obvious that if the refractive index n in Eq. 2.2 is not for air ($n = 1$) but for a liquid ($n > 1$), then the NA can be more than 1. The immersion imaging, also called "wet" imaging, is illustrated and compared with conventional "dry" imaging in Fig. 2.17. The liquid between the objective lens and imaging plane has effectively reduced the angle of light coming out of the lens. Although the NA of optical lens is the same as for dry imaging, the effective NA seen from the imaging plane side has increased to a greater value than actual NA of the lens. Table 2.3 shows the angle of refracted light in different media for a lens with a fixed NA. The light, which is used to be entering photoresist at 58° angle in the case of dry imaging, is now entering at 36° angle; in other words, the incident light with 58° angle in wet imaging, implies the NA value of about 1.22, which is greater than 1. However, the incident light angle cannot be increased by merely placing an immersion fluid between an objective lens and the image plane. The objective lens has to be specifically designed to allow larger angle of light interfacing with the immersion fluid. The most frequently used form of an immersion lens is a prism or hemispherical element, as the one shown in Fig. 2.17.



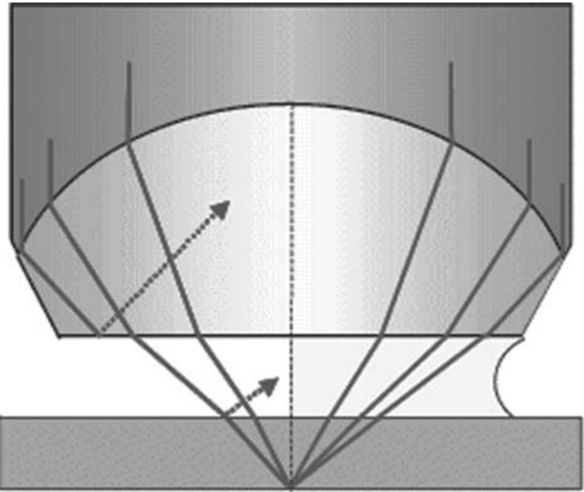


Fig. 2.17 Schematic comparison of "wet" and "dry" imaging

Table 2.3 Refractive angles of light in different media at NA = 0.85

Medium	Refractive index (n)	Angle of light (θ)
Lens	1.55	33°
Air	1	58°
Water	1.437	36°
Photoresist	1.7	30°

Immersion imaging at 193 nm has indeed made a leap forward in term of sub-100 nm IC manufacturing using optical lithography. Its potential to extend further into sub-50 nm resolution regime cannot be underestimated. For example, assuming $k_1 = 0.25$ and the NA of lens is 0.75, the dry lithography at 193-nm wavelength can achieve resolution of 65 nm, whereas the same optical system with immersion imaging using water can achieve resolution of 45 nm. While immersion imaging is such a simple and mature concept, it took the semiconductor industry so long to seriously consider the merits of replacing air with a liquid. This was because there were many technical difficulties to realize immersion optical lithography. In the past, it was not worth the effort to solve these technical difficulties, as conventional optics was more than adequate to meet the resolution requirement of IC production above 100 nm. Now, as optical lithography is approaching its end life, any measures that can revive it will be seriously considered.

One of the most serious technical difficulties is how to move the wafer stage rapidly relative to the objective lens while in water without stirring up bubbles. In order to achieve high throughput the wafer stage must accelerate rapidly and move accurately to each exposure field location, then settle and scan the image before stepping to the next location, all in a short period of time. Maintaining a

consistent bubble-free liquid between the lens and the wafer is very difficult. There are basically three approaches to the problem [27]. The first approach is to submerge the whole chuck, wafer, and lens in a pool of water. The issue with the pool approach is that a complex system of servo motors and laser interferometers is required to accurately move the chuck. Submerging this whole system would require significantly complex engineering. The second approach is to limit the pool size to the top of the chuck. This technique would keep all of the chuck control mechanisms dry but would add considerable mass to the chuck that must rapidly accelerate. The third and currently being implemented technique is to dispense the water between the lens and the wafer with a nozzle and rely on surface tension to maintain a water "puddle" while the wafer stage moves to each die location during step–scan process. Because the water puddle moves with the lens, no bubbles are stirred up during the step–scan process.

Another critical issue is the defects that are likely to form on photoresist surface due to water marks and drying stains. Water droplets left on the surface after exposure will interact a much longer time with the photoresist than the exposure itself. Photoresist molecules can diffuse into the droplet. When the droplet completely dries out, a watermark is left on the exposed surface, sometimes with a photoresist residue in the center. This results in partially or fully bridged patterns.

The immersion imaging has indeed saved the DUV optical lithography and pushed the IC manufacturing technology from 65-nm minimum feature dimension to 45 nm. The 45-nm technology has come out of R&D phase and entered volume production of ICs in 2007. Now the industry is researching the possibility of pushing the immersion imaging technology to 32 nm or even below. The current immersion imaging using water ($n = 1.44$) can achieve NA at 1.3 and is capable of 45-nm resolution. This is about the limit with water as the immersion liquid. Further improvement in resolution has to seek an immersion liquid with higher refractive index ($n = 1.8 \sim 2.0$). At present water has similar refractive index as the glass as shown in Table 2.3. To maintain index matching, high-refractive-index glass material is also needed for making the objective lens. Currently, the optical glass for making lenses and water as the immersion fluid are all well known for their physical properties. Introduction of new materials for both the immersion liquids and lens materials will lead optical lithography into unknown territory. These new materials have to meet the following requirements:

- low absorption at 193-nm wavelength,
- low index variation with regard to temperature (dn/dT)
- the viscosity and wetting characteristics of immersion liquid suitable for high-speed scanning,
- low staining to photoresist and low defect level,
- readily available at relatively low cost.

Research in this area is hotly pursued. In 2006 IBM demonstrated 29.9-nm line and space features achieved by 193-nm immersion optical lithography using a high-index fluid ($n = 1.64$) and a quartz crystal prism ($n = 1.67$) at the end of

objective lens to match the high-index fluid [30]. This has certainly proved that optical lithography can be done at even higher NA, and therefore has the potential to extend to 32-nm generation of ICs.

2.5 Optical Lithography at Low k_1 Factor

The third parameter influencing the resolution of optical lithography, as seen in Eq. 2.3, is the k_1 factor. The k_1 factor comes from Rayleigh Criterion. The Rayleigh Criterion states that the minimum resolvable detail of two images through an optical system is diffraction-limited. Two closely adjacent light spots are defined as resolvable when the first diffraction minimum of the light spot image coincides with the maximum of another light spot image. Such a definition of resolution for diffraction-limited imaging is illustrated in Fig. 2.18, where the second diffractive image has its maximum intensity coinciding with the minimum intensity of the first diffractive image. For diffraction through a circular aperture, the angle θ_R at which direction the first minimum intensity occurs at the image plane is related to the light wavelength λ and the diameter of the circular aperture d and expressed as Eq. 2.7:

$$\mathrm{Sin}(\theta_R) = 1.22\frac{\lambda}{d} \tag{2.7}$$

The minimum separation of the two image intensities, as shown in Fig. 2.18, would be the half of distance between the two minimums of diffractive image of single spot, which means the resolution limit for closely adjacent spots can be expressed as Eq. 2.8:

$$\mathrm{R} = 0.61\frac{\lambda}{\mathrm{Sin}(\theta_R)} \quad \text{for circular aperture} \tag{2.8}$$

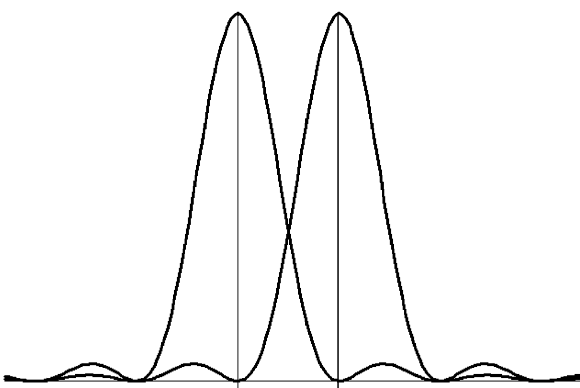


Fig. 2.18 Definition of Rayleigh Criterion

If there is a lens between the aperture and the image plane, the angle of diffraction from the aperture equals the angle of light projection from the lens onto the image plane. As described in Section 2.2, the NA of the lens can be expressed as Eq. 2.2. For the air medium between the lens and the image plane, the refractive index $n = 1$. Equation 2.8 then becomes the same as Eq. 2.3 which was introduced earlier in Section 2.2. Here, k_1 equals 0.61 for circular aperture. In the case of slit, the angular diffraction of the first minimum intensity for a single slit is expressed by Eq. 2.1 instead of Eq. 2.7. Therefore, k_1 equals 0.5. This theoretical limit refers to two line-type features brought adjacent to each other with a minimum separation that can be resolved on an image plane. In optical lithography the light images on photoresist surface are illuminated through a photomask. Two line features on a photomask cannot be brought completely adjacent to each other. There has to be a separation between the two features on a photomask. The highest density features on a photomask are equal line and space features or array of contact holes with space between two holes equal to the diameter of hole. In this case, the half-pitch (equal to the linewidth or diameter of contact hole) of the periodically repetitive pattern is the minimum feature to be imaged. The theoretical limit of k_1 factor for half-pitch pattern is 0.25. Figure 2.19 shows the light intensities (or commonly called "aerial image" in optical lithography community) of two line features (0.2-μm linewidth) in a photomask as they are separated by different distances. When the separation is half of the "half-pitch" ($D = 0.1$ μm), they are no longer resolvable.

It should be noted that the level of circuit integration in the past was measured by minimum feature dimension. For example, when speaking of 0.35 μm, 0.25 μm, or 0.18 μm IC technologies, they implicate the minimum linewidth of circuit feature in an IC chip. However, minimum linewidth cannot

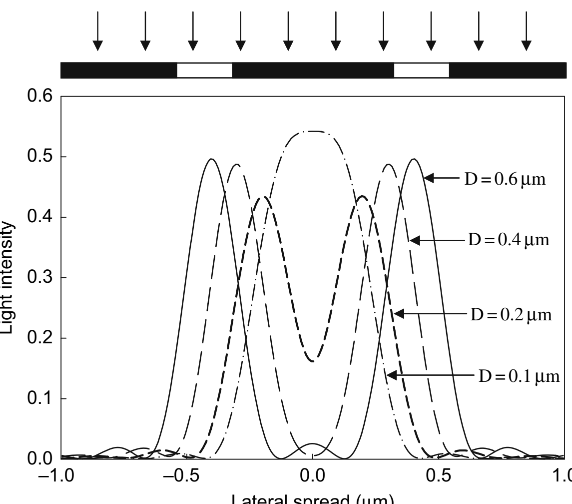


Fig. 2.19 Two separate line images (D is the separation distance between the centres of two line opening)

represent the true integration level of an IC because it does not indicate how dense the circuit features are in an IC. In fact, it is much easier to achieve high resolution for single isolated line than for dense line/space pattern. Since the 2001 version of International Technology Roadmap for Semiconductors (ITRS), the "half-pitch" definition was introduced and became the benchmark for new generations of IC technology [7].

Apart from developing shorter-wavelength illumination and high-NA optics, any other efforts in the development of optical lithography over the last 20 years have been concentrated on lowering the k_1 factor. Figure 2.20 shows the reduction of k_1 factor from G-line optical lithography in 1980s until 2000 when 193-nm DUV optical lithography was introduced. Such a trend coincided with the continuous shrinking of features dimension in IC manufacturing in the last 20 years.

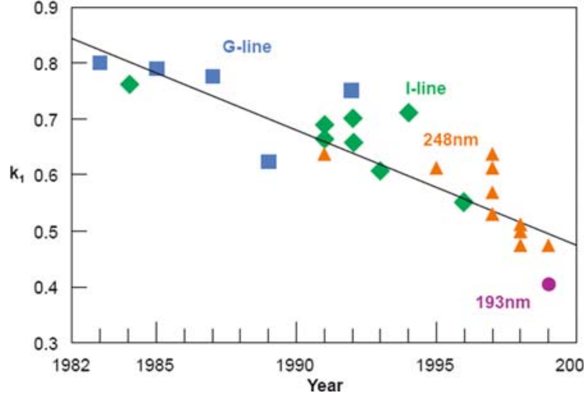


Fig. 2.20 Reduction of k_1 factor until 193-nm optical lithography was introduced (Reprint from [27] with permission.)

The k_1 factor encompasses any measures that can improve the resolution of imaging, not only in the optics but also the image resolution in photoresist layer. The improvements in imaging optics and photoresist processes have continuously driven down the k_1 factor. There have been innovations in every step of imaging process from top to bottom of an optical lithography system. They are all under the name of "resolution enhancement technique" (RET). Table 2.4 summarizes the RETs that have been developed to improve image resolution, contrast, and DOF. The following sections will introduce each of these innovative techniques.

Table 2.4 Summary of resolution enhancement techniques

Components of optical lithography	Innovations to reduce k_1 factor
Illumination optics	Off-axis illumination (OAI): annular, quadrupole, or programmable
Photomask	Phase-shifting masks (PSMs), optical proximity correction (OPC)
Photoresist layer	Top surface imaging (TSI), antireflective coating (ARC), double exposure or double patterning

2.5.1 Off-Axis Illumination (OAI)

Light illuminating through a photomask will be diffracted, as described in Section 2.2. The smaller the mask feature the larger the angle of diffracted light (Fig. 2.3). An optical system may not achieve high resolution if the large angular diffractions are cut off by finite NA of a lens. However, if the illumination is deliberately tilted through an off-axis aperture instead of an on-axis aperture, light passes through the mask feature in an oblique angle while the zero-order illumination is blocked. Figure 2.21 compares the conventional illumination (a) and off-axis illumination (b). In this way the high orders of diffraction can pass through the mask feature to be imaged. It should then enhance the image resolution.

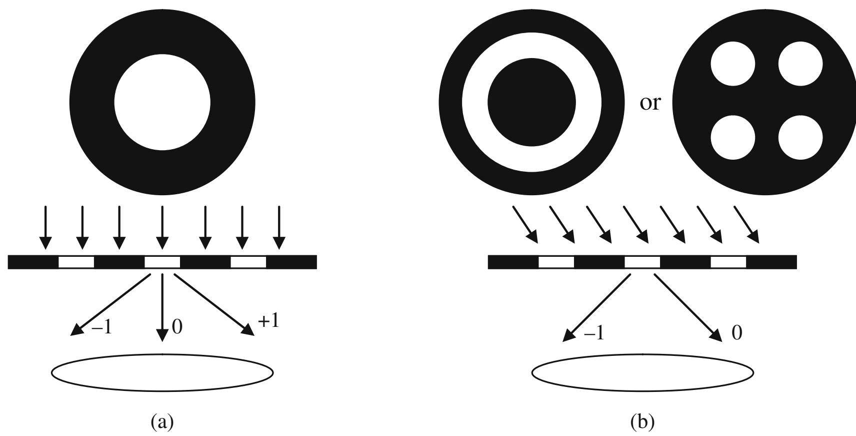


Fig. 2.21 Comparison of conventional and off-axis illumination schemes: (**a**) conventional aperture (**b**) off-axis apertures (annular or quadrupole)

Off-axis illumination was proposed in early 1990s. Two schemes of OAI were proposed, one was annular OAI [31] and another was quadrupole OAI, also called SHRINC (Super High Resolution by Illumi-Nation Control) [32]. Both schemes are illustrated in Fig. 2.21b. Quadrupole OAI is most effective for line and space (L/S) pattern, but less so for isolated features. With Quadrupole OAI, 0.275-μm L/S pattern can be resolved with I-line lithography. Quadrupole OAI is also dependent on the orientation of L/S patterns. It gives the best results for vertical or horizontal L/S pattern, but less optimum for L/S with 45° angle. Annular OAI, on the other hand, does not have the discrimination of pattern orientation, but the resolution enhancement effect with annular OAI is rather weak compared to quadrupole OAI.

Off-axis illumination is the easiest and requires the minimum cost, among other RETs, to implement in an optical lithography system. It certainly is the most widespread used RET in commercial optical lithography systems. Because the OAI effect is dependent on pattern density and orientation, the

state-of-the-art OAI apparatus are programmable. It can be tuned to a particular set of mask patterns to achieve the optimum performance.

2.5.2 Phase-Shifting Mask (PSM)

The concept of phase-shifting was proposed in 1982, whereby the light diffraction of a mask feature can be suppressed by introducing light wave with 180° phase difference surrounding the mask feature [33]. Figure 2.22 schematically shows how the light intensity distribution from a mask feature can be made narrower by introducing light of opposite phase in comparison with imaging of a conventional binary photomask. There are a number of ways to realize phase-shifting on a mask, depending on where and how to place the phase shifters. Figure 2.23 shows five different types of PSM in comparison with a conventional photomask. The shaded region on a PSM is 100% transparent to light but with 180° phase difference from unshaded pattern region, except the attenuated PSM (attPSM) (Fig. 2.23f) where the shaded region is only 5–15% transparent to light but has 180° phase-shifting. Different phase-shifting methods have different performance on resolution enhancement, with the alternating PSM and chromeless PSM being the strongest phase-shifting methods which produce the greatest improvement on imaging resolution. Figure 2.24 shows the simulated light intensity distributions (aerial image) of L/S pattern for conventional mask, alternating PSM and attPSM. The linewidth is 0.2 μm and exposed with 248-nm wavelength and 0.5 NA. It can be seen that the alternating PSM has the most significant improvement in aerial image over the imaging of conventional mask. The attPSM has improvement over conventional mask but not as pronounced as the alternating PSM.

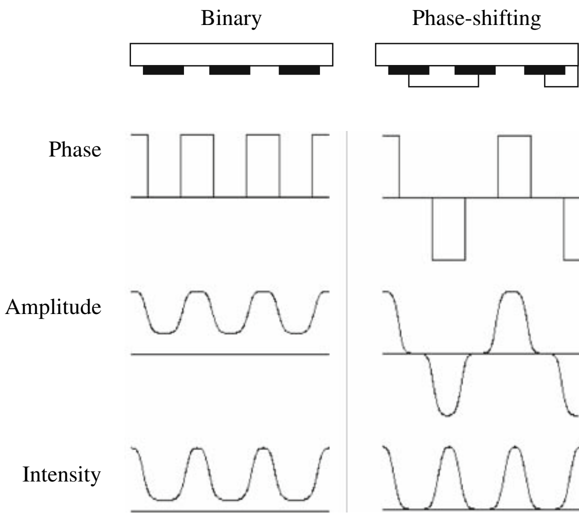


Fig. 2.22 The concept of PSM

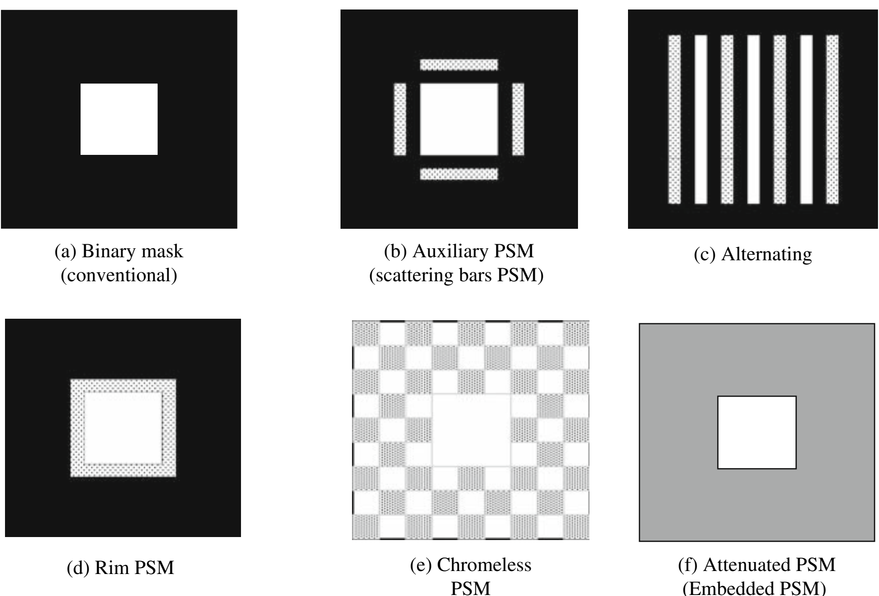


Fig. 2.23 Layouts of different PSM schemes (Reprint from [34] with permission)

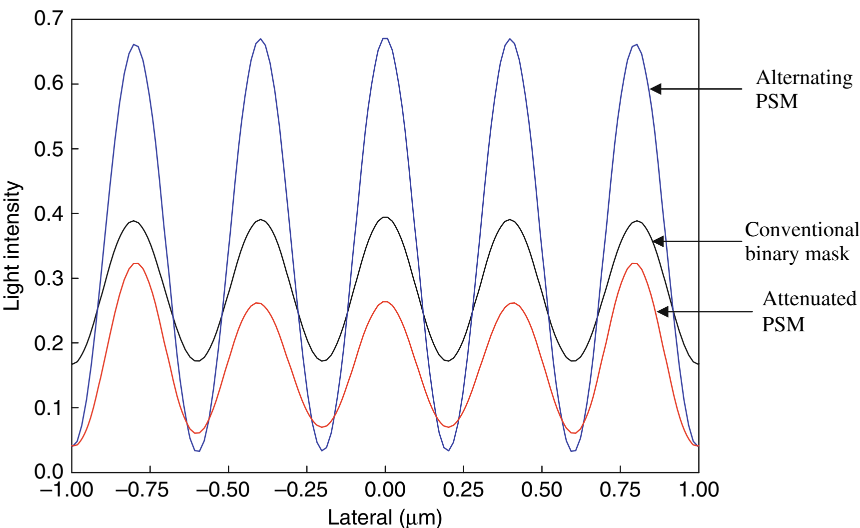


Fig. 2.24 Comparison of aerial images of conventional mask with alternating and attPSMs

Although the alternating PSM has superb performance in resolution enhancement, there are three critical issues with this type of PSM. First, alternating PSM is only suitable for patterns with high periodicity, but it is very difficult to apply the phase shifters systematically on a random design layout. Some memory cell layouts have been modified to be ideally suited for this approach, but this is not possible for logic applications like microprocessors or application-specific integrated circuits (ASICs). There are dedicated software tools which can deal with placing phase-shifting patterns to random design layout, but not all issues have been solved.

The second issue is the phase conflict. Each 180° phase-shifting region has a boundary. If the phase shifter is adjacent to a transparent 0° phase region, as is drawn of the mask layout in Fig. 2.25, the transition between regions of 180° and 0° phase will always print out as a dark narrow line whether this is intended or not, because of the cancellation of light at the transition region. For most of IC designs, it is very difficult to avoid these undesirable features created by the phase conflict. There are two ways to overcome this problem. One way is to apply a second exposure with a trim mask to eliminate these dark lines. This approach will increase the exposure time and require an additional mask. Another way is to design some intermediate phase shifters at the transition region so that the transition of phase is not from 180° abruptly to 0°, but to 120°, 60°, and 0°, for example. Both methods require extra space around the critical features and hence are less suitable for densely packed pattern layouts, such as memory cells.

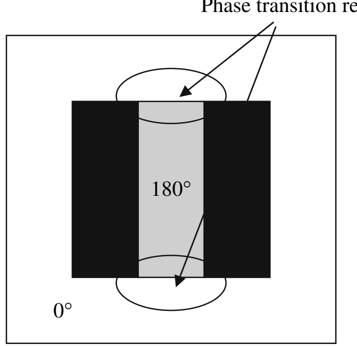


Fig. 2.25 Phase conflict in alternating PSM

The third issue is mask manufacturing complexity. Phase shift layer has to be produced by additional exposure and deposition or etching of the mask. This additional manufacturing step not only increases the mask cost, but also requires accurate alignment of two different patterns and increases the possibility of defects. Inspection and repair of phase defects are still critical issues. The overall cost of a PSM is at least 10 times higher than a conventional binary mask.

Although alternating PSM has the best performance, the first and most widely used PSM in optical lithography for IC manufacturing is the attPSM, solely because of its relative ease and low cost to make. Attenuated PSM is also

called half-tone PSM or embedded PSM. As shown in Fig. 2.23f, the attPSM is not much different from a conventional binary mask (Fig. 2.23a). The only difference is that the chrome layer on a conventional mask is replaced by a thin film which gives 5–15% light transmission with 180° phase shift. The transmission of attenuation film has to be carefully controlled as high transmission can cause side lobes which become unwanted features in photoresist [35]. Attenuated PSM is particularly suitable for isolated or low-density features, which makes it the best phase-shifting option for optical lithography of logic ICs or ASICs. Even though attPSM is not an aggressive or strong phase-shifting scheme, as has been demonstrated in Fig. 2.24, the economical factor alone has kept it alive all the way down to 65 nm node. As the photomask manufacturers pointed out, few IC manufacturers have required alternating PSMs. They are too expensive to make and the benefit out of that is not justifiable [36]. Process engineers can get by with attPSM and optical proximity correction (OPC) without using the alternating PSMs.

As mentioned earlier, alternating PSM is good for highly repetitive features such as the memory cell in DRAM-type ICs, but not for randomly distributed features which are common in logic-type ICs. To tackle this challenge, a variation of chromless phase-shifting scheme shown in Fig. 2.23e, called chromeless phase lithography (CPL), was developed in early 2000 [37]. Imagine that the transmission of phase-shifting layer on an attPSM becomes 100% instead of 5–15%, then at any edges of a mask feature there will be dark lines because of the cancellation of light where 0° phase and 180° phase regions meet. If the two phase edges come reasonably close, then the two dark lines become merged into one dark line, which is equivalent to a dark line produced by a conventional binary mask (clear-field mask) but with much higher contrast, because of the phase-shifting. This is illustrated in Fig. 2.26 where the aerial image of an opaque line from a binary mask is compared with that of an all-transparent mask but with a phase-shifting line. The all-transparent mask produces the same dark line aerial image as the conventional chrome-on-glass mask but has much higher image contrast. As it is known that negative-tone resist image in IC manufacturing is always realized by positive-tone photoresist combined with a clear-field mask, the CPL can well replace the conventional clear-field mask to create such negative-tone resist image with substantially improved resolution.

Dark aerial images from CPL are always formed by two phase edges in close proximity. The image formation is intrinsically symmetrical with no unwanted phase-edge images. The CPL layout needed to form a resist pattern on the wafer is not substantially different from that of a binary chrome mask. Chromeless phase lithography is intended to be a single-mask technique. It can be made by etching the quartz substrate of a conventional chrome-on-glass mask blank to create π-phase depth after the initial chrome patterning. The chrome is then removed in subsequent steps to reveal non-phase-shifting areas. Both "π-phase background areas" and "0-phase patterning areas" share the same quartz substrate. It looks like a conventional mask but without chrome patterns. By

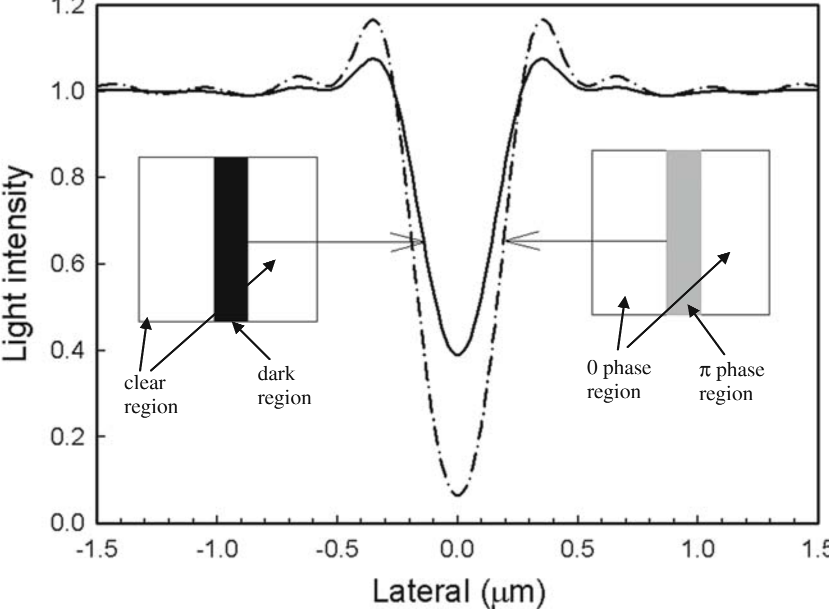


Fig. 2.26 Comparison of CPL with conventional binary mask

careful design of π-phase patterns, dark random patterns can be formed in photoresist. An example of CPL mask and its resist image are shown in Fig. 2.27 where irregular resist patterns are formed by phase-shifting features instead of chrome features. For design layouts with both large and small features on the same mask, a hybrid CPL mask can be used. A hybrid CPL mask contains both phase-shifting features and chrome features. For large patterns, chrome features can be used as it does not require high resolution. For small patterns, π-phase patterns can be used to improve image resolution. Because of the mixture of phase-shifting regions and chrome regions, two steps of processes are needed to fabricate the hybrid CPL mask. Chromeless phase lithoraphy performance can be further enhanced by high-NA imaging optics and off-axis illumination. In combination with OAI and 0.85 NA optics, 70-nm features with minimum pitch of 160 nm have been produced at 193-nm wavelength illumination, which corresponds to k_1 factor of about 0.3 [38].

2.5.3 Optical Proximity Correction (OPC)

Because of diffraction, fine details of an object could be lost through an optical imaging process. As illustrated in Fig. 2.13b, the fine details of an image are contained in those high spatial frequencies of diffracted lights. The fidelity of an

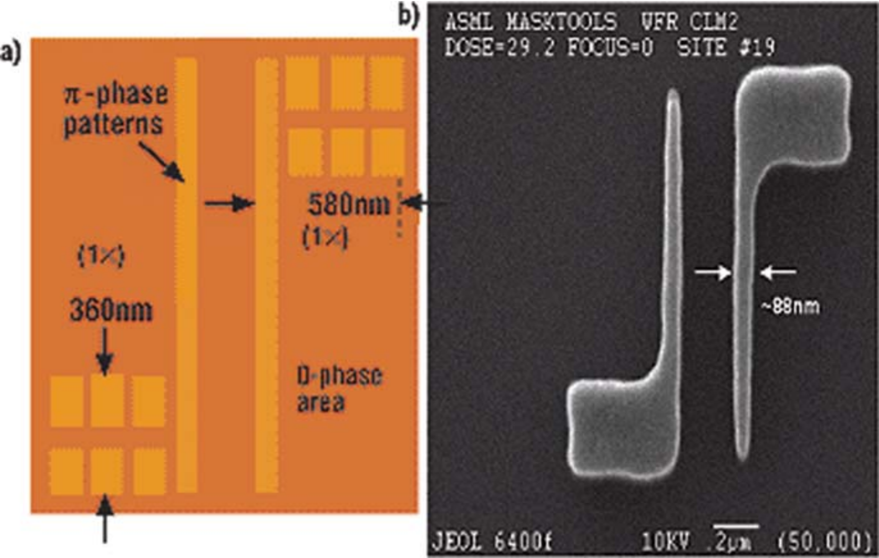


Fig. 2.27 CPL mask and printed resist patterns (Reprint from [38] with permission)

image can only be maintained if those diffractions of high spatial frequency can be imaged. Loss of fine details implies the image is no longer sharp but blurred. In optical lithography of IC patterns, this blur of mask pattern image is translated into corner rounding and line-end shortening and is called optical proximity effect. Such effect is demonstrated in Fig. 2.28a, where the optical image in the form of white intensity is shown to compare with the original mask pattern. Line-end shortening and corner rounding are apparent. Increasing NA and reducing illumination wavelength (λ) are able to reduce the proximity effect. However, as described in previous sections, there are limits as to how short wavelength and how high NA can be employed. Failing all that, one can just modify the original design layout to compensate the line-end shortening and corner rounding. This is the idea of OPC [39]. Optical proximity correction is to deliberately distort the original design to balance the image intensity. For example, overexposure region can be reduced by taking away part of the pattern and underexposure region can be enhanced by adding some extra features. Figure 2.28b shows the design pattern after OPC and its intensity image. Comparing the intensity images in Fig. 2.28a and b, the OPC has made the image closer to its mask design. Optical proximity correction can also be realized by modifying the transmission of different parts of a mask pattern using the so-called grey-scale mask technique [40]. Although the OPC design looks different from the original design, those added features are all sub-resolution, meaning that they are not resolvable by the imaging optics, hence will not be printable in photoresist.

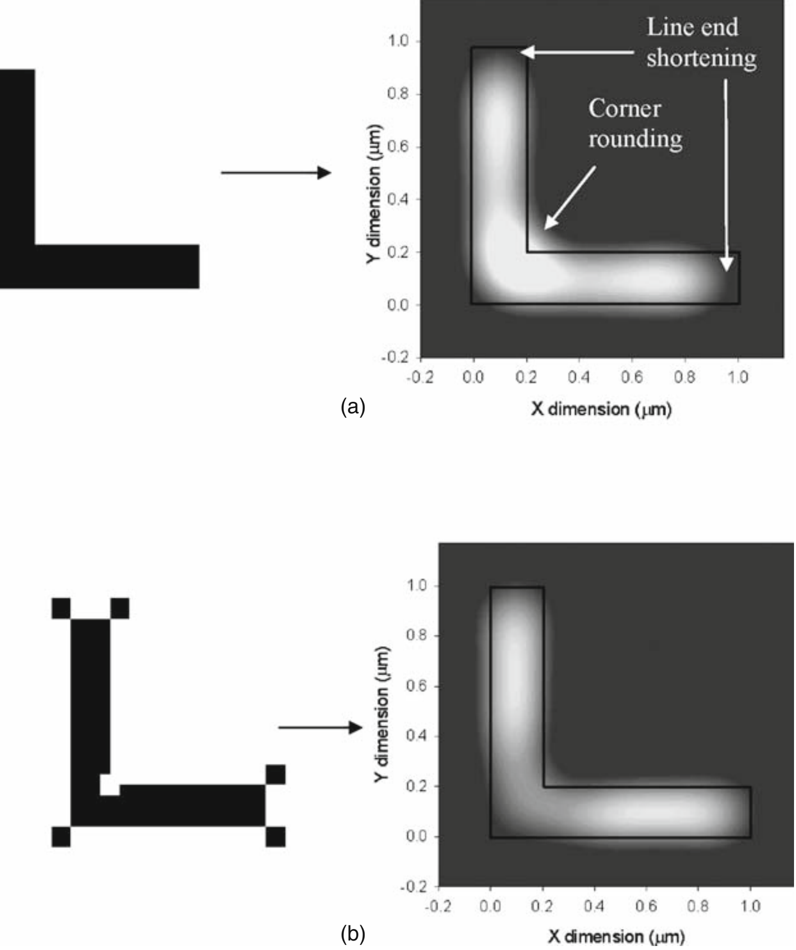


Fig. 2.28 (**a**) Optical proximity effect results in corner rounding and line-end shortening. (**b**) Optical proximity correction modifies the mask design to restore the pattern fidelity

Another interesting technique is to use the so-called scattering bars, or sub-resolution assist features (SRAFs) [41]. These are narrow lines or spaces placed adjacent to a primary feature in order to make a relatively isolated primary line behave lithographically more like a dense line. For example, Fig. 2.29 is a group of lines. The lower half of the central line can be considered as dense line because it is closely adjacent to other lines. The upper half of the central line, however, is an isolated line. Upon optical exposure, the resultant CD of the central line would be different in its upper section from its lower section. The upper section will be narrower than its lower section. This is the optical proximity effect. To make up the difference, one can either deliberately widen the upper section of the central line, which is called biasing, or add two narrow

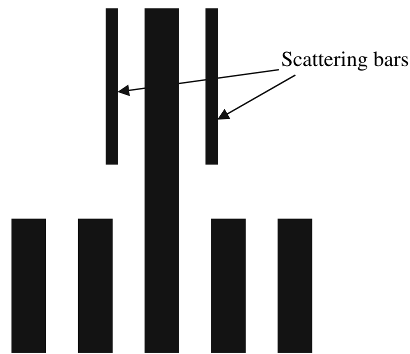


Fig. 2.29 Mixture of dense and isolated lines with added scattering bars

bars on both sides of the upper section, as shown in Fig. 2.29. These two narrow bars are called scattering bars or SRAF. They are too narrow to be imaged on photoresist (which is the reason why they are called sub-resolution features), but the light passing through these bars can compensate the low intensity of light of the isolated line, so that the CDs of both dense lines and isolated lines are of the same width. Sub-resolution assist feature is a complementary technique to the OPC methods mentioned earlier. Sub-resolution assist feature is only workable if there are spaces nearby to allow placing them.

Optical proximity correction has become a standard technique in sub-100 nm optical lithography. The implementation of OPC faces two challenges: one is the placing and routing of OPC features and the other is the fabrication of OPC mask. Optical proximity correction mask features are added after the IC design is done. It is a complicated task to determine where to place what type of OPC features or how to modify the original design. In the early days of employing OPC technique, placing the OPC features were based on some rules which were derived from experience and computer simulation of aerial images of some typical shapes [42]. Rule-based OPC has the advantage of fast correction because patterns are modified according to a set of rules. However, the accuracy is poor as the number of rules are limited, which cannot cover all the possible situations (variety of pattern shapes). The other method is the model-based method [43]. Model-based OPC starts first to calculate image intensity for all the patterns to find out where a pattern is overexposed or underexposed. Then OPC is applied to balance the image intensity over each pattern by adding or subtracting tentative features to the pattern. The process of analyzing and correcting is repeated. After each iteration the OPC features are modified, until the satisfactory image intensity distribution is achieved. The model-based method is obviously more accurate but very time consuming. With the advances in computer technology, simulation power has increased exponentially in recent years. Now the mainstream OPC methodologies are all model-based, especially for below 90-nm optical lithography. Optical proximity correction software has taken advantage of the latest advances in computer technology by adopting scalable and distributed computer simulation. For example, the Proteus™ OPC software from

Synopsys can simulate and modify a complete IC design containing over 500 million transistors by using more than 100 Intel Xeon processors. Proteus^{TM} can reduce the time to obtain OPC results from days to hours on 90-nm and 65-nm IC designs [44]. Further reduction of mask data volume and OPC simulation run time can be achieved by balancing the OPC complexity and minimum performance requirement, to avoid those unnecessary OPC implementations [45]. The models for OPC are no longer for aerial image simulation alone but incorporating lithography process modelling. To speed up the OPC process, the lithography models tend to be phenomenological models which do not model the actual physical and chemical processes but fitted from experimental parameters. As long as they are calibrated by test patterns through real wafer processing, they are safe to use with considerable efficiency because of their simplicity [46].

The ultimate solution for achieving the best optical lithography result is to design the photomask by working out how an ideal image is generated. Such technology is currently being developed and called “Inverse Lithography Technology” (ILT) [47]. Inverse lithography technology is to analyze a perfect optical lithography image and working backwards to find the “perfect” mask which can generate such an image. It is a rigorous mathematical approach to the design of photomasks for optimal lithography. The concept is not new [48]. However, as mask pattern generation is pixel-based, the final mask design which generates the “perfect” lithography image can be very complicated and very different from conventional design. It was then considered impractical because the mask was too complex to be manufactured. With IC manufacturing moving toward 45 nm and below, every possible way of improving optical lithography resolution has been explored. Inverse lithography technology would not be an exception despite its very complex mask structures. Inverse lithography technology can be viewed as an ultimate solution for OPC, or replacement of OPC, because the mask design has been optimized pixel by pixel to achieve the perfect balance of light intensity distribution on image plane. To reduce the difficulty in mask fabrication, the algorithm of finding the mask design can be tuned to different degree of aggressiveness. Less aggressive optimization can have a mask design which is easier to fabricate. Figure 2.30a shows two mask patterns from ILT design. The top one is an aggressive design and the bottom one is a nonaggressive design. The nonaggressive design lacks fine details; therefore, the mask is easier to make. Figure 2.30b shows the relevant resist images printed on wafer with the ILT masks.

All OPC features are sub-resolution, which means they are much smaller than the normal mask patterns. An OPC feature should be at most half the size of a normal feature. For 65-nm-generation IC, the mask feature will be 265 nm for 4:1 reduction optical lithography. The OPC feature on mask should then be less than 130 nm. Making such small patterns on a mask plate can be a considerable challenge. The 45-nm-generation IC manufacturing puts even tougher demand on mask manufacturing. Currently there are two mainstream technologies for mask manufacturing: laser direct write and electron beam direct write. Laser tools are fast and have high throughput, while electron beam technique has the high resolution but is slow. Figure 2.31 shows part of

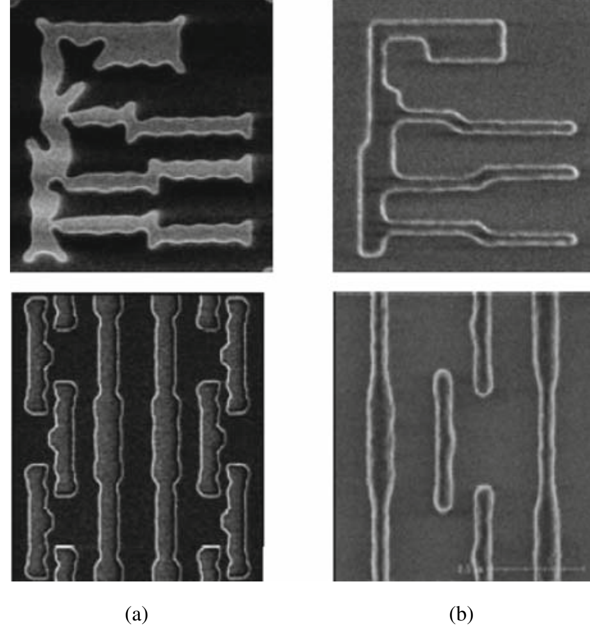


Fig. 2.30 (**a**) Aggressive (*top*) and nonaggressive (*bottom*) ILT masks (**b**) resist images printed from aggressive design (*top*) and nonaggressive design (*bottom*) (Reprint from [49] with permission)

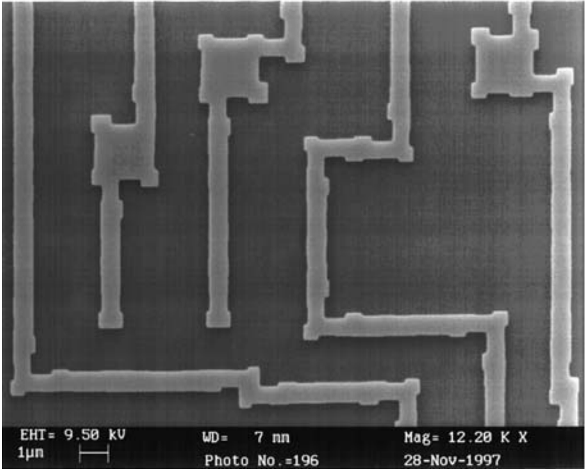


Fig. 2.31 Scanning electron microscope (SEM) image of an OPC mask showing OPC features added to normal mask features

an OPC mask made by electron beam lithography, where the smallest features (serifs) are only 100 nm. Any distortions introduced during mask-making can potentially degrade the OPC performance [50]. Other challenges are mask inspection and repair for these very small OPC features, and data handling of OPC design. The data file for a full mask OPC design at 90-nm technology is as large as 263 GB and increases to 729 GB at 65-nm technology [51]. With such a large data volume, not only computer processing time but also hard disk

accessing time has become limiting factors to the mask-making throughput. All have contributed to the escalating cost of masks for the sub-100 nm-generation IC manufacturing.

2.5.4 Photoresists

All the techniques described in previous sections are merely forming a light image on the surface of a photoresist layer. The mask pattern that represents an IC design will have to transform into a photoresist pattern so that it can be further transferred into the substrate material. The description of basic techniques for processing photoresists, such as exposure, baking, development, etc., can be found in many literatures [3, 52, 53]. Although the basic process steps are more or less the same whether for patterning of sub-micrometer features or for patterning of sub-100 nm features, many technical difficulties and challenges are encountered when the photoresist patterns are less than 100 nm. The four key parameters for photoresists and their processes at sub-100 nm-resolution regime are sensitivity, contrast, LER, and etch resistance.

2.5.4.1 Sensitivity

Since the optical lithography moved to DUV wavelength at 248 nm, CA resists, which are distinctively of high sensitivity, have become the working resists. The CA resists technology was pioneered by IBM in the 1980s in order to accommodate early DUV light sources [54]. At the time, the choice of exposure light sources was still mercury lamps which were widely used for G-line and I-line optical lithography. Although excimer lasers emerged as new sources of DUV-wavelength radiation, they were in a very early development phase and there was concern over their ability to operate reliably. Mercury lamps, already proven in manufacturing, were deemed a more reasonable approach. However, the DUV spectrum in a mercury lamp was rather weak, hence required very fast resists, leading to the development of CA resists. The high-sensitivity characteristics of CA resists rely on a substance called "photoacid generator" (PAG). In these resists, the lithographically important chemistry is not driven directly by photons absorbed during the exposure step. Instead, an acid formed during exposure of a PAG catalyzes chemical changes in the resist during the baking step immediately following exposure, which is called post-exposure baking (PEB). Since the reaction is catalytic, the acid is regenerated after each chemical reaction; the same acid molecule can participate in further reactions, hence is named chemical amplification. Each PAG can induce 800–1200 more photoacid molecules after the PEB step [55]. These photoacid molecules convert the exposed region of the CA resist into soluble or insoluble, depending on whether it is a positive or a negative resist. Because only a small amount of photon energy is needed to generate the initial photoacid "seeds" and the major chemical process occurs during PEB, they are therefore of very high sensitivity. While

conventional I-line resists need more than 100 mJ cm^{-2} photon energy, CA resists need only 10–20 mJ cm^{-2} photon energy.

The very high catalytic chain length in CA resists, which provides the high sensitivity, are also problematic because the generated photoacid molecules can be deactivated by airborne amines diffusion into the resist, which results in "T-topping" or swelling in resist pattern. Amine concentrations as small as 5 ppb in air are enough to induce the T-topping, depending on the length of time the resist is exposed to ambient air having trace of amine. For example, Fig. 2.32 compares the CA resist patterns when the exposed resist was developed immediately without delay (Fig. 2.32a) and when there was a delay of 45 min in cleanroom ambient before development (Fig. 2.32b). T-topping has been formed because of the exposure to ambient air. If the wafer was delayed for 45 min but kept in the exposure tool where the environment is enclosed and well filtrated, no T-topping was formed as shown in Fig. 2.32c. Improvement in resist chemistry to make it less sensitive to amines is a possible solution. But the most straightforward approach is to control airborne contamination with robust filters in the exposure tool or in the cleanroom. Chemical filters, in some cases, have moved from traditional carbon-based media to new formulations, including highly porous polymeric materials that can provide improved removal efficiency and capacity. High rate of chemical amplification during the PEB step also increases process sensitivity to resist baking cycles, necessitating extremely tight control of baking temperatures and times. From the process control point of view, too high sensitivity of a CA resist may not be beneficial. A compromise in sensitivity could render a CA resist more stable in a production process environment.

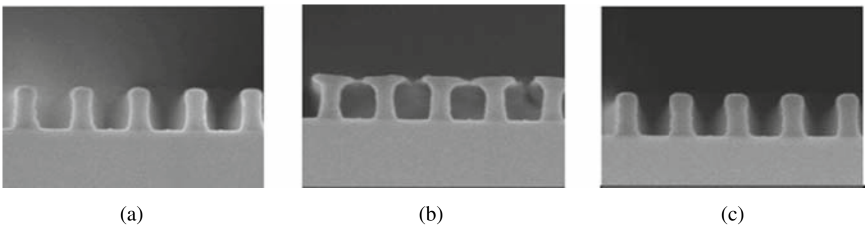


Fig. 2.32 SEM images of CA patterns (**a**) without delay, (**b**) with 45-min delay in cleanroom ambient, (**c**) with 45-mins delay but kept inside exposure tool (150 nm CD and ArF lithography) (Reprint from [56] with permission)

2.5.4.2 Contrast

High contrast is the key to achieve high resolution in a photoresist. There are two different contrasts, one is the contrast of optical imaging and the other is the contrast of photoresist, each of which has different definition [3]. However, one can tell the high-contrast optical imaging or high-contrast resist structure simply by comparing the aerial images (Fig. 2.33) or the resist profiles (Fig. 2.34a and b). It must be pointed out that high contrast of optical imaging does not

Fig. 2.33 Low- and high-contrast optical imaging

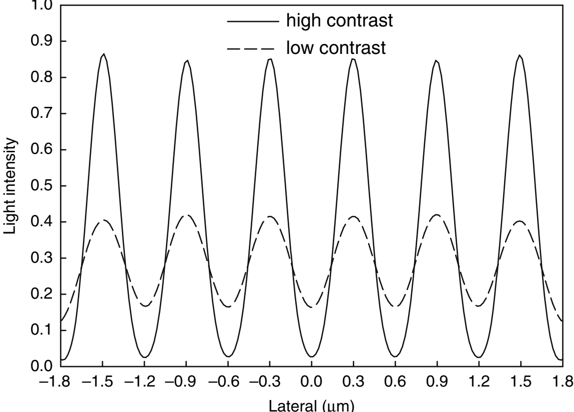
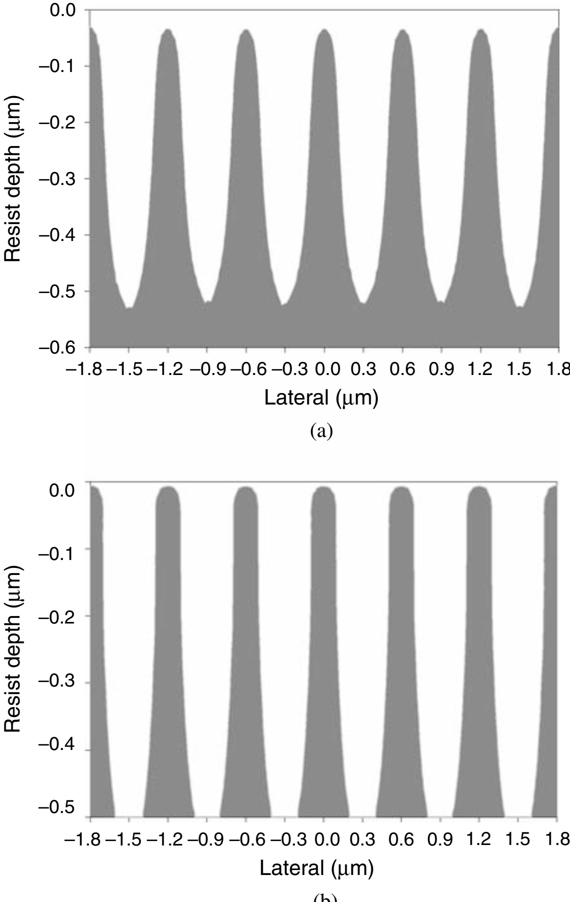


Fig. 2.34 (**a**) Low-contrast resist pattern (**b**) high-contrast resist pattern



necessarily lead to high-contrast resist profile, whereas a high-contrast photoresist can, to some extent, compensate a low-contrast optical imaging to produce high-resolution resist patterns; that is why the photoresist is a vital part of resolution enhancement technology.

To achieve high contrast, the resist material should be less absorptive to UV radiation. With low UV absorption, it is easier to achieve uniform distribution of photon energy through the depth of photoresist layer. Too absorptive will result in more photon energy being absorbed at the top and less at the bottom. As a consequence, the development of resist will be fast at the top and slow down through the depth, resulting in slopped pattern profile instead of a sharp and vertical resist profile. Unfortunately, materials are always more absorptive at shorter-wavelength radiation. Unlike I-line resists which work at 1-μm thickness, CA resists for DUV have to work at less than 0.5-μm thickness in order to compensate for the high absorption of photons. Alternatively, a bilayer resist system can be used [57]. Bilayer resist systems consist of a photosensitive silicon-containing top layer and an organic underlayer with good optical characteristics and planarization properties. The top imaging layer with a typical thickness of 100–200 nm contains 5–20% silicon. The bottom layer will provide antireflective properties and act as good etch-resistant mask for the substrate etch. Typical bottom layer thicknesses range from 400 nm to 1000 nm and allow for excellent planarization over local topography and provide high-aspect-ratio etch masks. Bilayer resist technique is also called top surface imaging (TSI) technique. It can effectively decouple the functions of imaging and etch masking, which are required for a normal single-layer photoresist. With the top thin imaging layer it is easier to achieve high-resolution pattern structure, because it is thin and the silicon-containing material is much more transparent to DUV wavelength than a CA resist. Once the imaging layer is exposed and developed, it serves as a etch mask to transfer the pattern into the bottom layer by oxygen plasma etching of the bottom organic material. The contrast of the final resist structure no longer depends on the photon energy distribution in the resist but on the RIE of the bottom resist layer. While bilayer resist systems offer many advantages, single-layer CA resists are still being developed and used because of its simplicity in process [58].

As the IC industry has widely accepted 193-nm immersion photolithography as a viable solution for half-pitch nodes down to 45 nm and possibly beyond, a photoresist has to be exposed under a fluid (currently water). Conventional 193-nm photoresists that are optimized for "dry" exposures may not be suitable for immersion exposure since resist components show some solubility in water and thus leach out, thereby contaminating the water and creating defects. The contaminated water can further degrade the front lens and wafer stage of the scanner. Water can also permeate into the resist film, causing resist swelling and changing its photochemical properties. An additional topcoat layer is an effective solution to block both the leaching of resist components into the water and the water permeation into the photoresist. The topcoat is spin-coated on the resist and is transparent at 193-nm wavelength. The topcoat layer is removed

after exposure and post-exposure bake (PEB) but before pattern development. There are two types of topcoats: those only soluble in some solvent and those that dissolve in resist developer (aqueous TMAH). The trend is toward developer-soluble topcoats, because the additional process complexity associated with a solvent-removed topcoat is too high. Because 193-nm immersion lithography has just been introduced to volume production of ICs, many technical issues associated with topcoat are still being investigated. Whether a topcoat is needed is still debatable [59]. In the meantime, immersion-capable photoresists are also being developed, since the additional topcoat will inevitably complicate the resist process and increase the cost of manufacturing [60].

2.5.4.3 Line Edge Roughness (LER)

For very small feature dimension, the edge roughness of developed resist lines can be a significant fraction of the linewidth. A measure of the LER is the 3-sigma standard deviation of the actual line edge relative to an average line edge. The roughness can be quantitatively measured by a scanning electron microscope (SEM) or an atomic force microscope (AFM). Figure 2.35 shows a typical LER measurement result with an AFM. Line edge roughness was included in the International Technology Roadmap for Semiconductors (ITRS) since 2001, as one of the key requirements on photoresists. International Technology Roadmap for Semiconductors indicates that the LER should be below 5% of the CD. For the 90-nm IC technology node, the LER should be less than 4.5 nm. While the functional lines patterned for ICs are shrinking toward 45 nm and below, LER does not shrink accordingly. This is particularly true for CA resists. Figure 2.36 shows some CA resist lines with average linewidth of 111 nm. The roughness, however, is 37 nm as measured by SEM. The LER for CA resists at sub-100 nm feature dimension has become a serious issue and could be the most significant source of linewidth control problem from now on in IC manufacturing.

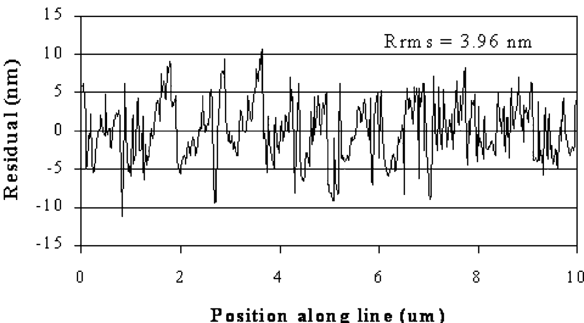


Fig. 2.35 Typical LER measured by AFM (Reprint from [61] with permission)

It is still unclear what are the causes of LER. A number of statistically fluctuating effects at these small dimensions may contribute to the LER, such as shot noise (photon flux variations), statistical distributions of chemical species

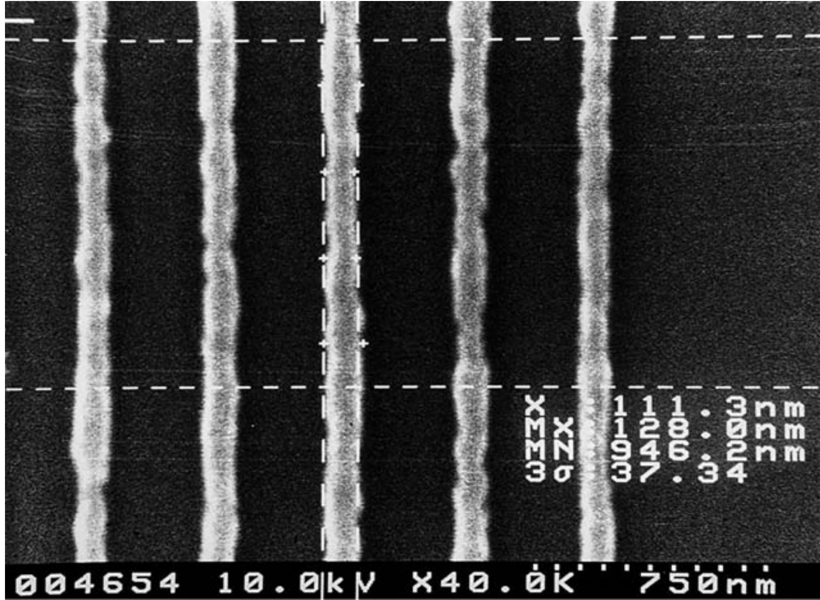


Fig. 2.36 Scanning electron microscope image of CA resist lines with average linewidth of 111 nm and LER of 37 nm

in the resist (such as PAGs), the random walk nature of acid diffusion during chemical amplification, and the non-zero size of resist polymers being dissolved during development [62]. Careful measurements and evaluation of LER data have ruled out the correlation between polymer molecule size and LER [63]. There have been attempts of using statistical modeling to explain the LER measurement data. The general trend is that the higher the resist sensitivity the rougher the resulting line edge. As demonstrated in Fig. 2.11, CA resists have not been able to meet the requirements of both high sensitivity and low LER. By analytical modeling of three key resist processing steps, the exposure, PEB, and development, it was demonstrated that the observed trade-off between resist sensitivity and LER is fundamental, which implies the basic limitation of achieving simultaneously low roughness, low dose, and high resolution in any standard CA resist process [64]. How the LER affects the final device performance is even more difficult to analyze, as there are so many process steps between the developed resist pattern and the final device. Any or all of these process steps can increase, decrease, or at least alter the character of the roughness, making any direct correlation of resist LER with final device performance very difficult. We are entering an arena in which microelectronics has become microstructural electronics. As Alexander E. Braun the Senior Editor of *Semiconductor International* puts it: “We are no longer just miniaturizing things, but dealing with changes on the microstructure, starting with surface roughness. Physics is really in play — there are no longer any smooth surfaces” [65].

2.5.4.4 Etch Resistance

Traditionally, etch resistance means the etch rate of resist polymer in plasma relative to the etch rate of substrate which the resist film is protecting. It is also called etch selectivity. The transition from 248 nm to 193 nm radiation wavelength results in reduced DOF because of the reduced wavelength and increased NA, and therefore requires a reduction in resist thickness from 450–600 nm to 300–350 nm. For two of the most demanding processes in IC manufacturing, the high-aspect-ratio contacts and dual damascene patterning, where the aspect ratios of deep holes can be as high as 20:1 and 6:1, respectively, the thin resist film has to endure longer etching in plasma in order to form these deep structures in substrate. Much higher etch resistance is required on 193-nm resists because of the reduced resist film thickness. On the other hand, the CA resists which used to work well at 248-nm wavelength have to go through radical change in polymer composition at 193-nm wavelength, in order to reduce the absorption of UV radiation in resist film. Compared to previous generations of resists, 193-nm resists are displaying more pronounced and more rapid morphological changes in plasma etch environments, because of the changes in polymer composition. Figure 2.37 shows the blank films of 193-nm resist and 248-nm resist after plasma etch under the same conditions. While the 248-nm resist film is intact, the 193-nm resist film has suffered considerable morphological change on its surface. The LER, which is created at the resist development stage, can only get worse because of the resist surface roughening in plasma etch. This can have serious impact on any pattern structures being transferred into a substrate.

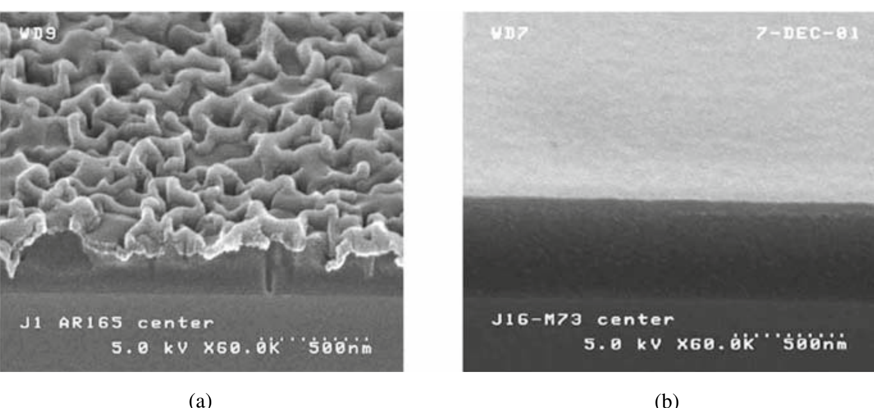


Fig. 2.37 Morphological changes in (**a**) 193-nm resist as compared to (**b**) 248-nm resist (Reprint from [66] with permission)

While the 193-nm resists are being developed and improved, the most direct way to improve the etching performance of 193-nm resists is to modify the etch conditions. One approach is to find process regimes which are gentler to the

resist but the etching of the substrate can still be accomplished. This would include hardware modifications and using less energetic process conditions by controlling gas, power, temperature, and other parameters. Another method is to passivate the resist surface, for example, by increasing the polymerization properties of the etch plasma. The resist is protected from the influence of the plasma and may become more resistant to etch. It is also possible to pretreat the resist pattern structure after development and prior to etch. Post-development hard baking is one of such methods. Resist structures are hardened at relatively high baking temperature, which helps to endure plasma etching better. This approach is desirable from an etch standpoint because less modification of etching process conditions will be needed, though it would add cost and complication to the manufacturing processes.

To elevate the burden of 193-nm resist as the pattern transfer mask, an alternative hard masking (HM) layer can be used. With either a metal HM or a bottom layer resist, an unstable 193-nm resist mask is quickly transferred into another layer that is more robust in the plasma. The metal HM is opened in a metal etch system and the subsequent etch steps are performed as usual, or to perform both the metal and substrate etches in the same etching system. However, there are concerns about the impact of metal residues and by-products in the etching chamber after long-term usage. While the metal HM technique is high on the option list to replace 193-nm resists for deep trench etching, opening the metal mask with minimum contamination is a top priority in the future development. The multilayer resist systems, as described earlier, are also an option. A bilayer or trilayer resist system can separate the imaging and etch layers, the uppermost layer designed for optimum imaging properties and the underlying layer(s) for dry etch resistance. Despite its long established maturity, the multilayer resist technology has seen little commercial success because of the complexity of the process, defect levels, and cost [66].

In addition to the four performance gauges mentioned above, there are a few other factors that can critically influence the production worthiness of 193-nm resists. They are the linearity of translating mask pattern to wafer, line slimming due to electron radiation during SEM inspection, and pattern collapse during wet development. They are briefly described in the following.

Usually, the mask features are linearly imaged onto a wafer. A variation of mask CD will cause variation of wafer CD, except four times smaller if it is a 4:1 reduction optics. However, at sub-100 nm-resolution regime with low-k_1 factor, small variations of CD on the mask can cause disproportionately large changes in wafer CDs. This amplification of mask errors is called the mask error enhancement factor (MEEF). For a linear imaging, MEEF = 1. If MEEF > 1, the errors in mask CDs are nonlinearly transferred to the wafer even if the mask CD itself gets geometrically reduced by the proper ratio (4:1). For the dense line/space pattern, both theory and experiment indicate that mask errors are amplified by 2.5 times at $k_1 = 0.35$ [67]. Fundamentally, the nonlinear imaging originates from the decrease of aerial image contrast because of the loss of high spatial frequencies in projection imaging. An easy solution to counter the large MEEF can be simply

biasing the mask feature. In the meantime, mask errors and defects have to be reduced at low-k_1 optical lithography, to avoid the detrimental effect of MEEF.

With IC manufacturing entering sub-100 nm regime, wafer inspections have widely adopted CD-SEM, a SEM-based CD inspection and measurement, because of its much higher resolution capability than optical inspection methods. However, it has been observed that patterns of 193-nm resist shrinks under electron irradiation during CD-SEM inspection, although very low electron beam energy (below 1 keV) and beam current (below 10 pA) are used in CD-SEM. The shrinkage due to electron irradiation ranges from 0.2 nm to 0.6 nm per measurement scan for different groups of 193-nm resists [68]. The longer the electron beam dwells on the resist pattern, the more significant the feature shrinkage becomes. Among the two common polymer systems of 193-nm CA resists, acrylate and methacrylate, the methacrylate system seems suffers more severe shrinkages. After 30 scan measurements at 800-eV electron energy, the methacrylate-based resist shrinks up to 30 nm for contact hole patterns [69]. The exact mechanism of e-beam-induced shrinkage is not clear. While the 193-nm resists will still be improved, a reasonable approach to this problem is to readjust the CD measurement procedure to minimize resist deformation.

Collapse of high-aspect-ratio resist patterns is a well-known phenomenon. After development and rinsing, capillary forces are present between closely packed resist structures. Unbalanced capillary forces lead to a net force on the pattern during drying process. Pattern collapse occurs if the capillary forces exceed the elastic restoring forces of the polymer. Figure 2.38 demonstrates the typical resist pattern collapse phenomenon. A figure of merit for the collapse behavior is the critical aspect ratio of collapse (CARC), which is defined as the aspect ratio for a given width and pitch with 0% of line collapse [70]. The worst line collapsing happens for asymmetric structures consisting of an asymmetric pitch. On the other hand pattern collapse occurs only for unbalanced forces. Therefore, the collapsing effect is more probable in the middle part of a line than

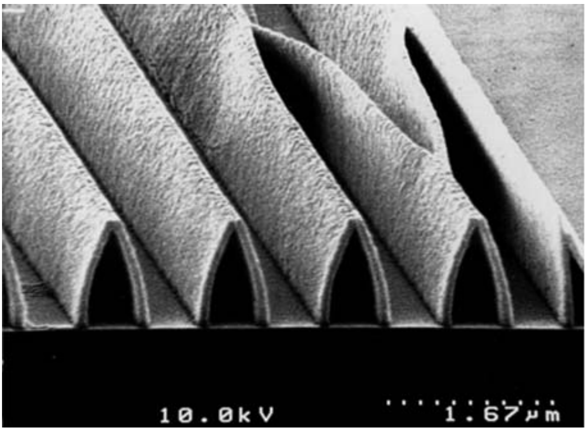


Fig. 2.38 Resist pattern collapse

at their end, because water level is easier equalized at the line end. For the same reason, the probability of line collapse is higher for longer lines in comparison to shorter lines. Because the capillary force is proportional to the surface tension of the rinsing liquid, using a supercritical fluid, such as liquid CO_2, instead of water for rinsing can eliminate the capillary force effect. The simplest approach to reducing pattern collapse is to reduce the resist film thickness. However, for single-layer 193-nm resists the lower limit for the resist thickness has to be in the range of 300–500 nm because of the etch resistance consideration. In this case, using multilayer resist systems described earlier can be an option.

2.5.5 Design for Manufacturing (DFM)

As described above, to lower the k_1 factor many RETs have to be used. The original mask layout for an IC design may have changed considerably after applying various phase-shifting, OPC, SRAF, and other RET schemes. The printed photoresist image on a wafer can be vastly different from the pattern on a mask because most of the SRAFs are not printable. In the earlier above-100 nm process technologies, IC designers do not have to know the lithography processes. They only need to follow a set of design rules which defines how small a feature should be and how close two features to each other should be, etc. In fact IC engineering is probably the only engineering discipline in which designer and manufacturer do not know each other's trade. This traditional method has been successful and the semiconductor industry has until now considered that the business benefits of hiding the complexity of lithography corrections outweighed the potential benefits of involving IC designers in lithographic optimization. Coming to post-90 nm technologies, however, that consideration has been reversed. In fact, the traditional goal of implementing RETs without involving the IC designers prevents use of more sophisticated RETs that cannot otherwise be applied to an existing design. Integrated circuit layout designers have seen design rules exploding and, for a 90-nm process, it has been a 4–5 times rise in the number of rules, many of which are never explained to a designer whether they are related to lithography, etch, CMP, other manufacturing needs, or all of the above. On the other hand, complicated RET measures and addition of RET features to the original design have put considerable pressure on IC process engineers. Some of the RET modifications may be sound for the optimization of optical imaging, but may be completely unmanufacturable, or with very low manufacturing yield. At 90 nm, IC manufacturers find that only less than half of IC designs operate as expected and more than half of them need a complete mask "re-spin" for acceptable yield and performance. With the design costs approaching $10 million or more and mask costs alone nearly $1 million, design for manufacturability (DFM) has become a key issue and IC designers must consider manufacturing yield and DFM issues throughout the design cycle [71].

The key to DFM is to identify the manufacturability of a particular design. Traditional DFM relays on design rule checking, parasitic extraction and physical analysis, while manufacturers try to tune the manufacturing processes in order to produce shapes and structures defined by a design. There are very little communications between the designers and manufacturers. The new methodology for DFM is to start with identification of manufacturability, then goes back to designers to tune the design in order to achieve the maximum manufacturing yield. One of the DFM approaches favoured by IC manufacturers is to predict, during the design stage by model-based computer simulation, how the ideal design layout will print as contours in silicon. These contours are equivalent to 3D structure profiles produced through the flow of manufacturing processes, which are accurate representation of the final manufacturing results. From that, designers can predict hot spots which are marginless spots where the circuits tend to fail. Design for manufacturability involves predicting how those ideal design layouts will change into contours and how those contours will impact the chip performance. The models used for the prediction are carefully calibrated by a particular manufacturer's processes, so that they are correct to each individual IC manufacturing processes. Once the hot spots are identified, the design layouts will be manipulated to make sure those potential failure spots are eliminated [72]. This model-based prediction is different from previously described OPC simulation. It has incorporated not only optical imaging but also manufacturing processes. A new set of lithography friendly design rules can be established by the manufacturing process-related computer simulation [73, 74]. Such a large-scale simulation, which would have not been possible in the past, is now feasible with the latest computer simulation power. A full-chip prediction of silicon contours across a range of processing parameters can be done within 8 hours for an 18 mm $\times$ 18 mm chip [75]. From 90-nm technology onwards, DFM has become a necessity to keep manufacturing yield up and cost down.

2.5.6 Double Processing

The most active and creative researches in optical lithography in the last 20 years have been various ways of lowering the k_1 factor. The typical RETs that help lowering the k_1 factor have been introduced in previous sections. As demonstrated in Fig. 2.20, k_1 factor has been driven down continuously from 0.8 in 1980s to below 0.5 in 2000. Since year 2000, k_1 factor has been further reduced. While the wavelength in optical lithography comes to an end at 193 nm, before the introduction of EUV, reduction of feature size in IC manufacturing has been largely attributed to the reduction of k_1 factor at 90-nm technology downwards, except the recent introduction of high NA with immersion imaging. Figure 2.39 shows the trend of reduction of k_1 factor since 2001 [76]. A slight jump up in 2004 represents the advent of immersion

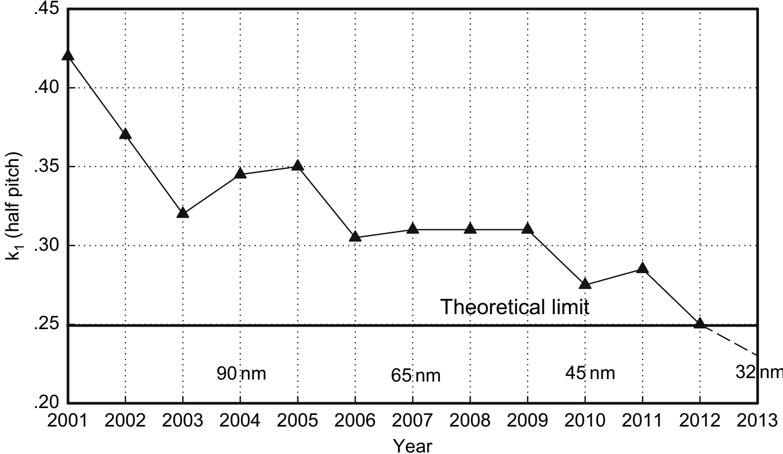


Fig. 2.39 Trend of k_1 factor reduction since 2001 (Reprint from [76] with permission)

lithography at 193-nm wavelength, whereby higher NA relaxed a bit on the demand of lower k_1 factor. However, with the IC features going from 65 nm to 45 nm to 32 nm, k_1 factor again has to be reduced, even though hyper NA imaging has become a norm in 193-nm lithography. It will soon approach the theoretical limit of 0.25.

If optical lithography still sticks to 193-nm wavelength, the k_1 factor has to be further reduced at 32-nm technology node, even with higher NA. The only technique which can help to achieve k_1 factor below the theoretical limit is double processing which includes either double exposure or double patterning. Double processing is illustrated and compared with conventional single processing in Fig. 2.40. It is known that high resolution is much easier to achieve in isolated features than in dense features, which is the reason why the IC manufacturing technology is benchmarked by minimum half-pitch instead of minimum feature size. If traditional single processing, as illustrated in Fig. 2.40a, cannot pattern very dense high-resolution features, why not split the high-density patterns into two sets of low-density patterns. By separately processing the low-density patterns and overlaying them together, the final result is high-density patterns.

The double processing can be double exposure, as shown in Fig. 2.40b, which involves single coating of photoresist, two sequential exposures, then single development and single etch. Double patterning, as shown in Fig. 2.40c, involves much more processing steps. It needs two coating of photoresists, two development and etching steps, and a hard mask to separate the imaging layer from the underlayer. Both double exposure and double patterning require two sets of masks. The two exposures have to be accurately aligned to each other within 1.5–3 nm. As seen in Fig. 2.40, the double exposure approach is much simpler than double patterning. It requires fewer processing steps.

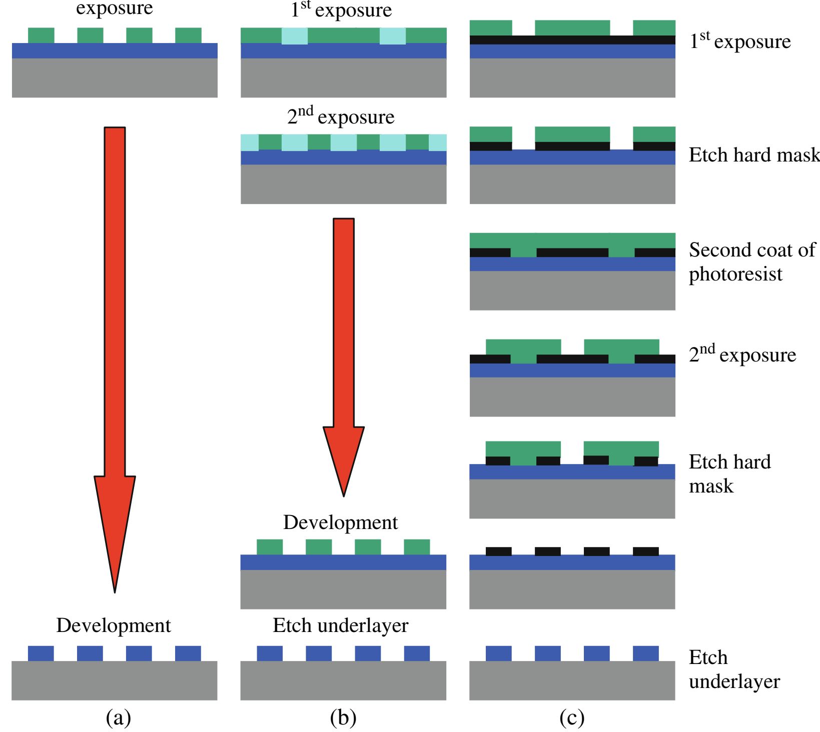


Fig. 2.40 Comparison of single processing (**a**) with double exposure (**b**) and double patterning (**c**)

However, this approach requires a nonlinear photoresist. Because the two exposures take place in the same resist layer, propagation of photon energy and PAGs distribution tend to overlap each other for the two consecutive exposures in the same photoresist, unless the photoresist is nonlinear or "memory-free", which confines the photon energy and PAG within the exposure region. Several potential solutions are being explored, but no final material solutions are available at present. Alternatively, the photoresist can be developed after the first exposure. A fresh layer of photoresist is then coated onto the developed resist pattern for the second exposure. It saves one additional step of etching but complicates the second spin-coating of resist, because the photoresist is no longer coated on a flat surface but a topographical surface.

Double processing can break the theoretical limit of k_1 factor because it has no longer to do with optical imaging. For optical imaging of each mask set, the k_1 factor is still above the theoretical limit. With the combination of two exposures the theoretical limit is no longer applicable. It is already possible to produce 32-nm half-pitch patterns in a laboratory environment using double processing technology, with the equivalent k_1 factor of only 0.14 [77]. Unlike

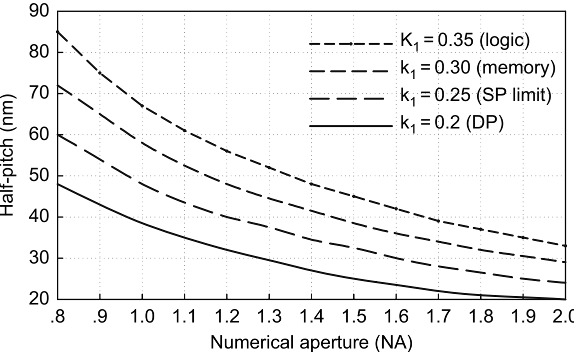


Fig. 2.41 Shrink of half-pitch size by single processing (SP) and double processing (DP) at higher NAs (Reprint from [79] with permission)

EUV lithography, double processing technique still utilizes the 193-nm wavelength and all the existing optical lithography infrastructures. It can be applied to either dry or wet (immersion) lithography. Therefore, it has been greatly favored by IC industry. Figure 2.41 predicts how double processing (DP) can surpass current technologies to bring IC manufacturing to well below 30-nm half-pitch, with the introduction of high-NA immersion lithography. Double processing has been added to the IRTS in 2006 [78], and has been considered the most likely candidate for 32-nm IC manufacturing. There were even talks about whether EUVLwould ever be needed in IC manufacturing at the 2007 SPIE Advanced Lithography Conference.

2.6 Near-Field Optical Lithography

Projection optical lithography has been well developed over the course of more than 30 years. It has become the mainstream lithography technology for high-volume IC manufacturing. However, one should not forget that optical lithography originated from contact printing. Contact printing relies on a maskaligner to print a mask pattern onto a photoresist in 1:1 ratio. Although maskaligner is no longer used in modern IC manufacturing for many reasons such as low throughput, no image reduction, and easy damage to mask plate, it still has a role to play in large-feature-size lithography, particularly in manufacturing of MEMS (micro-electro-mechanical-system) devices. In recent years, contact optical lithography has been revitalized as an alternative low-cost technique for sub-100 nm patterning technique.

It is known that projection optical lithography is a far-field optical imaging process, which has a fundamental resolution limit of $\lambda/2$, where λ is the wavelength. Contact optical lithography is by nature a near-field optical imaging process. There is effectively no space for light wave to travel between the mask-opening and photoresist layer, except within the photoresist layer. At such an intimate contact, light waves no longer propagate as sinusoidal wave but as evanescent wave, which decay exponentially with distance. Therefore, near-

field optical lithography is also called evanescent near-field optical lithography (ENFOL) [80]. Near-field optical lithography evolved from the contact printing but is not the same as a conventional contact printing. Instead of using a hard glass mask, a thin conformable mask is used. Figure 2.42 shows the setup of near-field optical lithography, where the mask is made of 2-μm thick silicon nitride membrane with chrome patterns. The membrane mask is supported by a silicon frame and mounted onto a vacuum chuck. The intimate and conformable contact is realized by vacuum between the membrane mask and photoresist-coated substrate. Though not a transparent mask, the thin nitride membrane can transmit 61% of G-line wavelength (436 nm) and 29% of H-line wavelength (405 nm) irradiation from a broadband mercury arc lamp [81], which is adequate to expose commercial G-line photoresists.

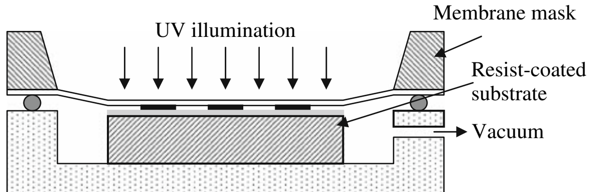


Fig. 2.42 Setup of near-field optical lithography (Reprint from [81] with permission)

Because the evanescent wave can only travel very short distance, the ENFOL can only be performed on very thin photoresist, normally less than 100-nm thickness. Such a characteristic is best explained by the computer simulation of electromagnetic field distribution, as shown in Fig. 2.43. It shows contour plots of the squared electric-field intensity for a cross-section through a 140-nm-period and 40-nm-thick Cr mask illuminated from above with 436-nm light. The exposure is into a 60-nm-thick photoresist layer on silicon. The black and white regions correspond to the ratios of $|E|^2/|E_0|^2$ from <0.1 to >1.3, where $|E_0|^2$ is the incident intensity.

Two properties are revealed for the evanescent wave propagation. The first is that the penetration of evanescent wave is only about the top 40 nm of the resist layer, as seen in the bright region from the Cr mask opening into the resist layer, which implies that complete resist clearance of exposed patterns will be difficult to achieve. In this case, a short O_2 plasma etching will have to be performed to remove this residual resist layer prior to transferring the resist pattern into the underlying silicon. The second is the large field enhancements seen at the edges of the mask pattern made of metal conductor, where the bright regions represent $|E|^2/|E_0|^2>1$, indicating that the local electrical-field intensity can be greater than the incident intensity. The local enhancement is strongly dependent on the sharpness of the edges, which is a possible explanation that patterns produced by evanescent near-field lithography are always of high LER.

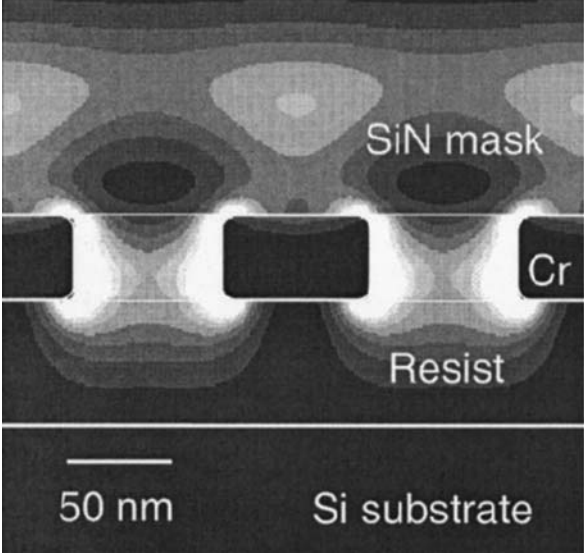


Fig. 2.43 Computer-simulated contours of squared electric-field intensity $|E|^2/|E_0|^2$, where $|E_0|^2$ is the incident intensity (Reprint from [81] with permission)

The fact that electrical field of incident light has been enhanced or amplified indicates there is something else happening in the evanescent near-field imaging process. This is the surface plasmon polariton (SPP) phenomena. Surface plasmon polariton is a charge density wave on the surface of material which has plenty of free electrons, such as silver or gold, with silver being the best material for exciting the SPP. The SPP wave can be excited by an incident light and it can propagate along the interface of the conducting material and a dielectric material. The SPP wavelength depends on the wavelength of incident light and the permittivity of conducting material [82]. Surface plasmon polariton is a form of electromagnetic radiation. Therefore, a photoresist can be exposed by the SPP radiation. Since SPP is a propagating wave, interference can occur, resulting in standing waves with periodically distributed maxima and minima. Figure 2.44a shows the computer-simulated near-field distribution for two well-separated mask-opening apertures illuminated by G-line ($\lambda = 436$ nm) light wave [83]. The mask material is silver. The separation of the two apertures is 300 nm. The simulation reveals clearly that the electrical-field maxima exist not only at the mask apertures, but also between the apertures where there are no opening to the incident light. This is the result of interference of SPP waves from both sides of the aperture. The computer simulation has been proved by experiment. The line features on the mask are 60-nm wide with pitch of 300 nm. Exposure is carried out using a G-line illumination source. The photoresist layer is 50-nm thick. A vacuum chuck is used to pull the mask onto the resist-coated substrate to ensure intimate contact. Near-field lithography using the widely separated mask has created dense lines in photoresist, as shown in

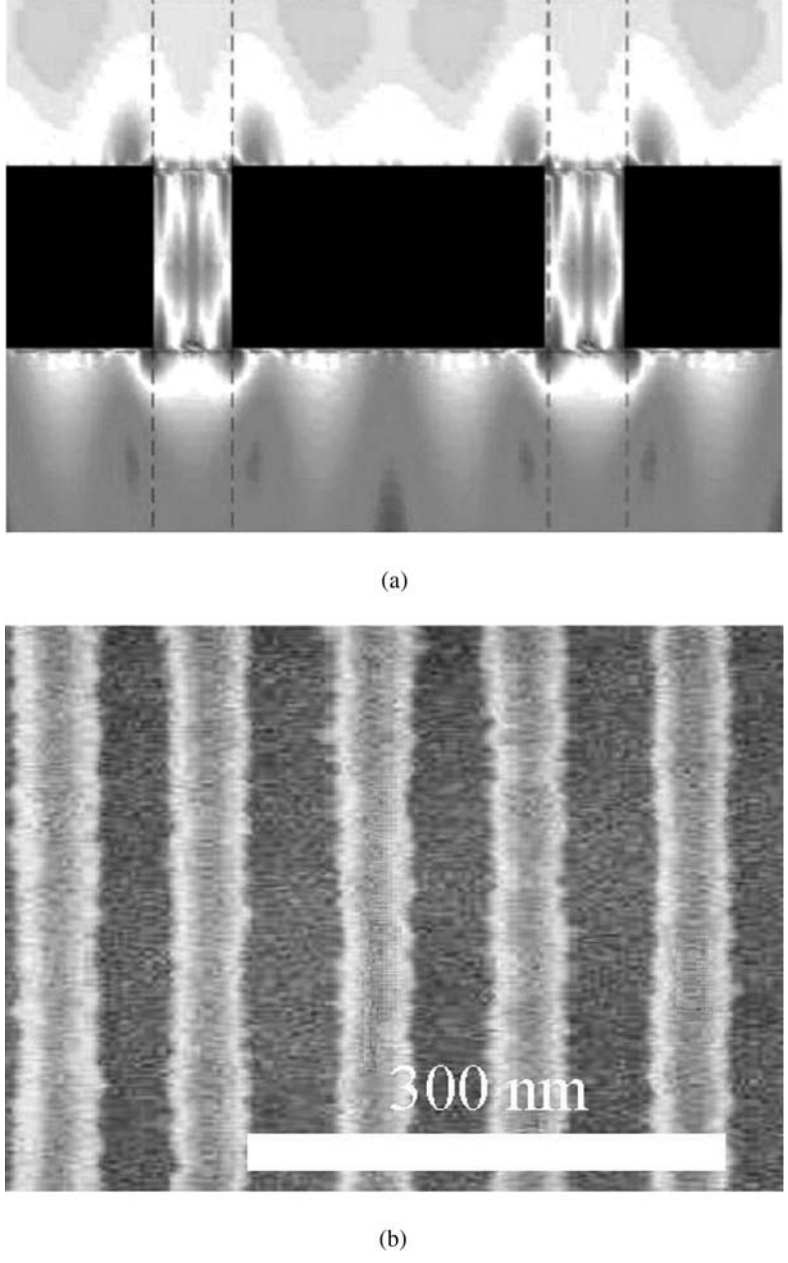


Fig. 2.44 (**a**) Computer simulation of near-field distribution (**b**) experimental lithography result (Reprint from [83] with permission)

Fig. 2.44b. Although the pitch of line grating on mask is 300 nm, the actual photoresist pattern has a pitch of only 100 nm. The extra lines between the mask features are created by the interference standing wave of SPP. The above interference occurs in the opaque space between two mask feature openings. Similar SPP interference can also happen within the opening of a mask feature, which creates smaller features in photoresist than the mask feature itself [84]. It is also noticed that the LER is quite significant as shown in Fig. 2.44b. This is caused by local enhancement of electrical field, as indicated in the simulation in Fig. 2.43a. Such an enhancement is due to the high concentration of charge at the rugged edge of the mask feature.

Because photoresist can be exposed by SPP, not necessarily by incident light, then it is not necessary to have the mask in intimate contact with photoresist for patterning. A pattern can be printed onto photoresist, as long as there is a mechanism to generate the SPP pattern. This is the concept of silver superlens [85]. The superlens is simply a thin film of silver sandwiched between a photomask and a photoresist layer, as schematically illustrated in Fig. 2.45. A conventional photomask (not the membrane-type conformable mask) is coated with 40-nm-thick polymethylmethacrylate (PMMA) as a spacer, then 35-nm-thick silver is coated as the superlens, and finally 120-nm photoresist is coated on top of the silver film. There is no optical path from the mask to the photoresist. Yet the mask pattern can be transferred to the photoresist through the superlens layer. It is the SPP resonance in the silver layer that causes the exposure of photoresist. The SPP resonance is initiated by the evanescent wave generated through the optical illumination of photomask. Forty-nm line features have been produced although the optical illumination wavelength is 365 nm. The key condition for the superlens effect is the thickness of silver layer. Above 40-nm thickness of silver, the resonance enhancement is dampened due to material absorption, which destroys the superlens effect [85]. Because of the resonance, the usually weak evanescent wave can be substantially enhanced. With the enhanced amplitude of field, a thin SiO_2 layer ($\sim$10 nm) can be placed between the silver superlens layer and photoresist to protect the superlens layer [86].

As described above, near-field optical lithography can indeed realize sub-100 nm patterning of photoresist using much longer illumination wavelength. However, there are a number of inherent drawbacks in this technology:

– *Intimate contact is required*. Because the near-field optical lithography relies on evanescent wave which decay exponentially with distance, the mask has to be intimately in contact with photoresist. This is the reason that a membrane-type conformable mask has to be used. A membrane mask is fragile. The requirement of conformable contact with photoresist-coated substrate makes it even more vulnerable to break. Membrane mask is also difficult to be made into large size. The silver superlens approach can alleviate the awkwardness of using a membrane mask. It still requires the silver superlens layer to be in intimate contact with the photomask. To meet this

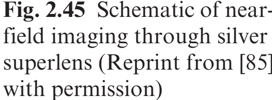
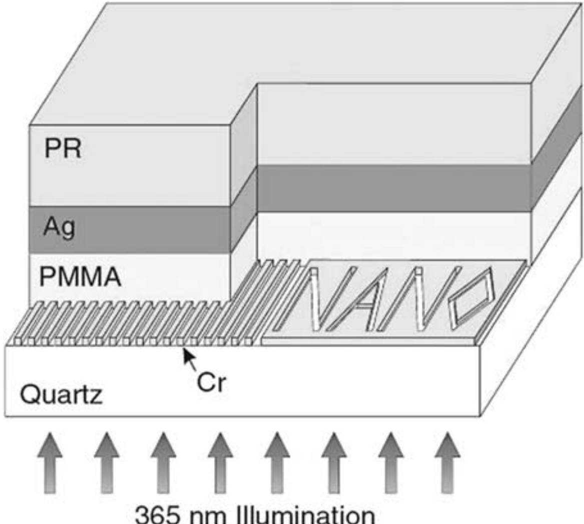


Fig. 2.45 Schematic of near-field imaging through silver superlens (Reprint from [85] with permission)



requirement, the silver superlens, together with a spacer layer, has to be directly coated onto the photomask surface, as illustrated in Fig. 2.45. Such a process can make the reusability of the photomask greatly reduced.

– *Exposure depth is very limited.* Both evanescent wave and SPP wave are absorbed very rapidly in photoresist, resulting in a very shallow exposure depth. The thickness of photoresist for near-field optical lithography should be no more than 100 nm. Such a thin photoresist layer makes subsequent pattern transfer by etching very difficult. Multilayer resist structure is commonly used, whereby the near-field optical lithography only patterns the top imaging layer and transfers to underlayer by RIE. Fortunately, multilayer photoresist process is a very mature technology as described previously.
– *LER is significant*. Because the charge density induced in the metal mask by incident light is highly concentrated at any sharp edges of a mask feature opening, the locally enhanced intensity can cause significant LER in the exposed photoresist features. There are two approaches that can reduce the LER: one is to improve the mask-making process to reduce the LER of mask features, the second is to improve the photoresist material to make it less sensitive to the speckles of image intensity.

Nevertheless, near-field optical lithography is a simple and low-cost technology for making sub-100 nm patterns. Even though it may never become viable in mass manufacturing of ICs, this does not stop the technology to be used in laboratory for research wherever there are requirements for making sub-100 nm structures. Its similarity to conventional contact printing using a maskaligner makes the technology much more accessible than the state-of-the-art optical projection lithography.

2.7 Interferometric Optical Lithography

Another simple, but capable of sub-100 nm resolution, photon-based lithography is interferometric optical lithography. Optical interference refers to the interaction of light waves which are correlated or coherent with each other, either because they come from the same source or because they have the same or nearly the same wavelengths. The most famous example of optical interference is the Young's double-slit experiment. The two beams from the double-slit are both of same wavelength and same phase because they emit from the same light source. The inference of the two beams causes distribution of light intensity with periodical maxima and minima. Interferometric lithography (IL) is the use of a small number of coherent optical beams (minimum of two) incident from different directions on a 2D (thin) or 3D (thick) photosensitive layer to produce an interference pattern whose intensity distribution is recorded in the photosensitive layer. Interferometric lithography is a well-established lithographic technique. The use of IL for photoresist patterning goes back to 1970s when it was mainly used to produce hologram plates [87, 88]. The alternative names are interference lithography or holographic lithography. Figure 2.46 illustrates the process of a simple two-beam interferometric lithography.

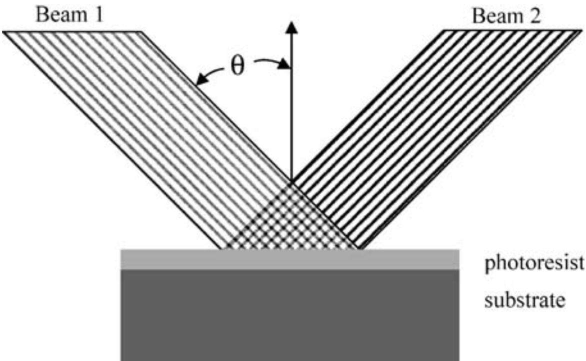


Fig. 2.46 Schematic of interferometric lithography

The intensity of interference pattern at the photoresist surface can be expressed as Eq. 2.9: [89]

$$I(x) = 2|E|^2\left[1 + \sigma_{\text{pol}}\cos(2kx\sin\theta)\right], \quad k = \frac{2n\pi}{\lambda}, \tag{2.9}$$

where $\sigma_{\text{pol}} = 1$ for TE (transverse-electric) polarization or $\sigma_{\text{pol}} = \cos(2\theta)$ for TM (transverse-magnetic) polarization. λ is the illumination wavelength, θ is the angle of incident beam relative to normal as indicated in Fig. 2.46, and n is the refractive index. For light passing through air into photoresist, $n = 1$. The period of the interference pattern along the x direction is $\lambda/2n\sin(\theta)$. For a fixed

wavelength, periodical lines with different pitches can be obtained by varying the angle of incident beams. At a fixed angle of incidence, further increase of resolution can be achieved by using short wavelength of illumination. It is apparent that at the maximum oblique angle of incidence ($\theta = 90°$) the minimum possible pitch is $\lambda/2$, which represents the resolution limit of IL. For half-pitch feature, the limit is $\lambda/4$.

The basic setup for IL is very simple. It only needs a laser, some lenses and mirrors. Figure 2.47 shows schematically a setup of IL system. A laser beam is focused through a pin hole (to improve coherence) and collimated through a collimator lens, then split into two beams which are reflected by two mirrors onto a photoresist-coated substrate. An interference light intensity distribution is formed and the photoresist is exposed by the interference light. For the lasers, early work was based on continuous-wave (cw) single-mode lasers such as the Ar ion laser at 488 nm (Ar^{+}) and the He–Cd laser at 442 nm. With the commercial development of I-line lithography (365 nm), attention shifted to the cw Ar^{++} lines at 364 nm and 352 nm and to the 355-nm third-harmonic of a YAG laser to take advantage of availability of commercial I-line resists. Further pursuing of short wavelength has seen the use of the intracavity doubled Ar-ion line at 244 nm and fifth-harmonic YAG at 213 nm. At wavelengths shorter than 257 nm it is difficult to obtain sources with temporal and spatial coherence and output power sufficient for conventional IL. ArF excimer laser has the necessary output power, but poor in temporal and spatial coherence. Hence, to expose gratings and grids of sub-100 nm period, achromatic interferometric lithography (AIL) is used. Achromatic interferometric lithography utilizes a group of phase grating plates to modulate the incoming laser light, so that the spatial period of the printed grating is dependent only on the period of the parent gratings used in the interferometer regardless of the optical path or the wavelength and coherence of the source [90].

Interferometric lithography has two distinctive advantages; one is its simplicity. In fact, any laboratory of optics can install an IL system on an optical

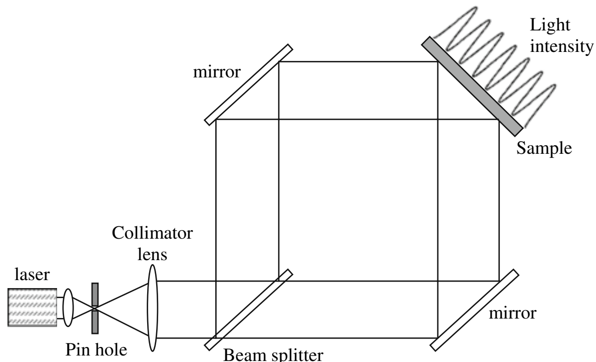


Fig. 2.47 Schematic of interferometric lithography setup

bench. Another is lithography without a mask. The interference pattern is generated directly on a photoresist-coated substrate. The downside is that the interference pattern can only be in the form of straight lines from a single exposure, or dots if two exposures are carried out on the same substrate with the second exposure rotating 90° to the first one. Whether straight lines or dots, they are periodical patterns with the pitch determined by the beam incident angle or wavelength of the laser.

Although the pitch of exposure pattern is dictated by illumination wavelength and beam incident angle in IL, the actual line or dot feature within the periodical pattern can be controlled by working with the sensitivity and contrast of a photoresist. The typical sensitivity curves of positive photoresist are shown in Fig. 2.48, where the initial thickness of a positive photoresist reduces with the increase in exposure dose until it is completely removed. Higher-sensitivity photoresist requires less exposure dose to clear up. At the slope section of the curve, the remaining thickness depends linearly on the exposure dose. Therefore, it is possible to control the exposure feature dimension by selecting appropriate exposure doses. Figure 2.49 shows a typical interference intensity distribution which is responsible for creating the exposure patterns in a photoresist. The light intensity distribution is narrow at the top section. Then the line image in photoresist can be wide or narrow, depending on where to set the threshold intensity, as shown in Fig. 2.49. High threshold intensity implies less exposure dose needed. Using a low-sensitivity photoresist or a low exposure dose, line features smaller than the half-pitch of interference period can be achieved, or vice versa to obtain line features larger than the half-pitch.

Apart from tuning photoresist process, the other two routes to realize sub-100 nm by IL use immersion technique, as described previously, and using short

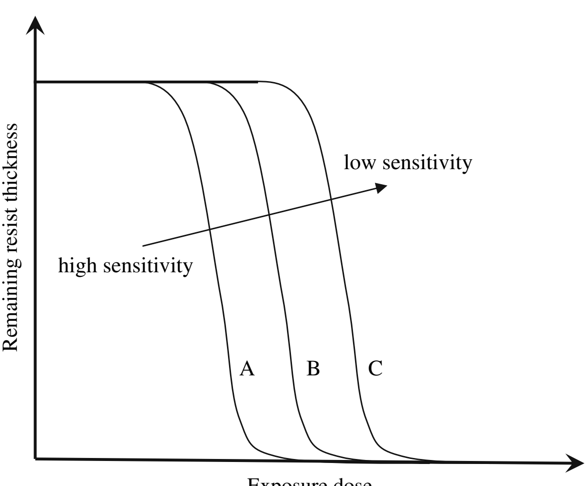


Fig. 2.48 Typical sensitivity curve of photoresist (sensitivity A>B>C)

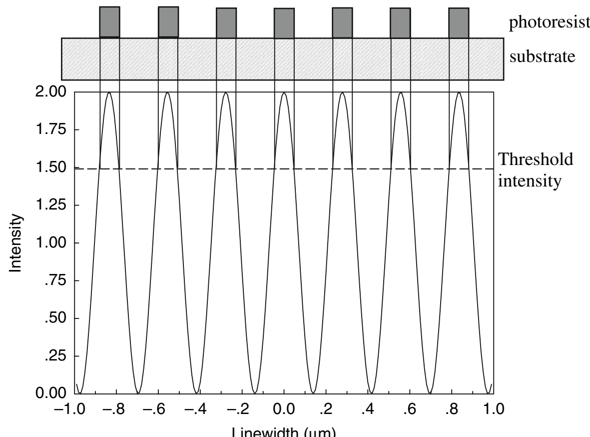


Fig. 2.49 Control of exposure feature dimension by tuning exposure dose

wavelength. As seen in Eq. 2.9, increase of n or decrease of λ can both reduce the pitch, therefore, increase the resolution. The refractive index can be increased by immersion lithography. There has been demonstration of 29.9-nm half-pitch lines by IL with the combination of 193-nm wavelength and high-index fluid immersion imaging [30]. Interferometric lithography with EUV wavelength has demonstrated 12.5-nm half-pitch features [91]. Without using a mask and its associated complications, such as phase-shifting and OPC, the IL certainly provides a very simple solution to push down further the resolution of optical lithography.

Although it is maskless, the IL does not have the flexibility as projection optical lithography because it can only pattern periodical features. However, there are many applications which require periodical features, for example, 2D and 3D photonic crystal structures [92], magnetic storage devices [93], etc. There was a report using seven laser beams shining from all directions through a thick lump of SU-8 photoresist, in order to form complex 3D interference structures as a photonic band gap material [94].

Interferometric lithography is not completely useless in IC manufacturing. It can be adapted to patterning IC design layouts. One of the ideas developed recently is hybrid maskless optical lithography, which uses IL to produce periodical patterns first, then uses a conventional photomask to trim the interference patterns into desired arbitrary pattern shapes. Figure 2.50 gives an example of the hybrid lithography, where a line/space pattern has been trimmed in the middle with a coarse dark field mask. The trim mask is usually low resolution. One can work out many ways of designing a trim mask to tailor the interference pattern into a desired IC layout. The key advantage of the hybrid approach is that the most difficult part of lithography, that is, the highest resolution patterning, is done by the maskless IL, which saves many troubles from using those RET masks mentioned earlier. Although the hybrid

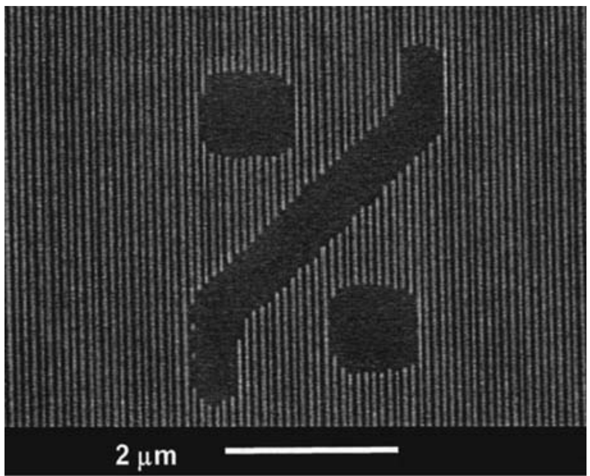


Fig. 2.50 Photoresist pattern produced by interference and projection hybrid lithography (Reprint from [95] with permission)

lithography requires double exposure, the IC industry has already accepted double processing as a necessary step toward 32-nm technology as described in previous sections.

The IL should, however, not be confused with imaging IL (IIL) technique. Imaging IL is basically a projection optical lithography with a photomask, but combined with interference from a reference beam to enhance the resolution of projection lithography [96].

2.8 Maskless Optical Lithography

"The devil is in the mask". This was the message delivered by Burn Lin, the world renowned optical lithography expert, at the SPIE 20th Annual Conference on Optical Lithography in March 2007 [97]. Optical lithography, since it started from contact printing, has always been relying on photomasks to deliver pattern designs to substrates. With IC manufacturing going into sub-100 nm feature dimension, photomasks have increasingly become a bottleneck as more and more RET features are incorporated into the photomasks, such as phase-shifting, SRAF, OPC, and masks for inverse lithography, as have been described in previous sections. Not only has the mask-making become increasingly complicated, but also the control of defect level, defect inspection, and repair have become more and more difficult. All add to the cost of photomasks for IC manufacturing. It was estimated that the cost of mask set could double for each IC technology node and the mask set for state-of-the-art 65-nm technology cost was already at £3–£4 millions by 2005 [98]. The escalating mask cost can have serious impact on the overall cost of low-volume IC production such as foundry runs and fabrication of ASICs. For small-volume

IC manufacturing, maskless lithography (ML2) has long been favored and it could potentially become a mainstream lithography technology for manufacturing below-32-nm generation of ICs. ML2 includes a number of optical and nonoptical technologies. The nonoptical ML2 techniques use electron beam or ion beam, which will be described in Chapter 3. For optical ML2, a number of techniques exist. The interferometric lithography (IL) introduced in the previous section is a ML2 technique. However, as mentioned above, IL does not offer the total flexibility because only periodical patterns can be generated by IL. In this section, some other optical ML2 techniques are introduced.

The technology of optical lithography without a mask has been around for two decades. It was started in 1987 by the introduction of first laser direct write tool [99]. It used a laser beam to scan to perform direct exposure of photoresist. The technology deemed to be slow and low throughput, therefore, was not able to compete with conventional projection optical lithography. However, it was highly suitable for lithography in making photomasks. At the time, mask-making was dominantly using electron beam (e-beam) raster scan system. The first laser tool CORE-2000 offered multi laser beams with raster scan, which greatly increased the throughput of exposure. Since then, laser direct write tools have evolved from the first generation of CORE 2000, CORE 2564 to ALTA 3000, and the latest ALTA 4300 series, manufactured by Etec Systems (an Applied Materials company). The first laser tool was used for making masks for the 0.7-μm-generation ICs. The latest ALTA 4300 tool can be used for making masks for the 90-nm-generation ICs. While the ALTA 4300 has comparable resolution as a 50 keV shaped e-beam mask writer, the writing time is less than half of that of an e-beam writer [100]. Over 75% of masks are now produced by laser tools.

A simple schematic drawing of direct laser scan is shown in Fig. 2.51. A single laser beam from a frequency-doubled cw Ar-ion laser at wavelength of 257 nm is split into 32 individually addressable beams, like a light brush. The brush is scanned across the plate by a 24-facet polygonal mirror. The pattern design data is "painted" onto the reticle plate in the form of pixels which lie on a rectangular grid. A single facet of polygon mirror scans the brush over an area of 4096 pixels wide $\times$ 32 pixels high. While the brush is scanned across the reticle by the facet of mirrors, the reticle stage is driven by a linear motor stage moving at constant velocity in a direction perpendicular to the scan. A set of brush strokes made during one stage move is called a stripe. A stripe is always 4096 pixels tall, but its length varies up to the length of the reticle depending on the pattern data. After a stripe is written, the stage performs a rapid retrace and moves to the starting position of the next stripe. The velocity of the stage is 32 pixels per facet. This stage movement lines up each successive "brush stroke" to provide complete area coverage. For ALTA 4300, the pixel size is 96 nm in the x-axis and 160 nm in the y-axis. One should remember that a mask feature is four times larger than the feature on a wafer, because of the 4:1 reduction of projection optical lithography system. Therefore, such a pixel size is adequate for patterning the

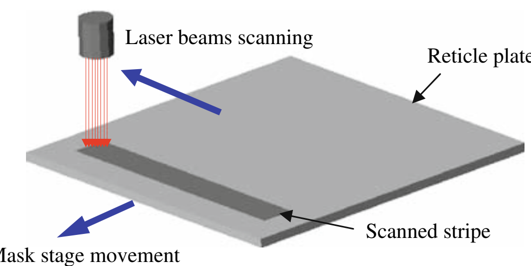


Fig. 2.51 Schematic of laser scan in making reticles

masks for 90-nm generation of IC manufacturing, including all the subresolution features used in RET.

In parallel to the ALTA series of laser tools, a new class of optical maskless lithography was developed in 2001, which was based on micromirror array to reflect or deflect light [101]. This new technology benefited from the development of MEMS (micro-electro-mechanical system). A micromirror is a MEMS device in which the mirror surface is tilted by applying electrostatic field between the mirror plate and an underlying electrode, as schematically shown in Fig. 2.52a. The areal dimension of a mirror is only a few tenths of square micrometers. A large array of closely packed mirror array can be fabricated by silicon surface micromachining process. Digital Micromirror Devices (DMDs), mass produced by Texas Instruments (TI), has been widely used in video projectors. Because the deflection of each micromirror is individually controlled by a pair of underlying electrodes, it acts as a pixel. Light shining onto the mirror surface can be reflected to different angles, depending on the deflection of mirror surface.

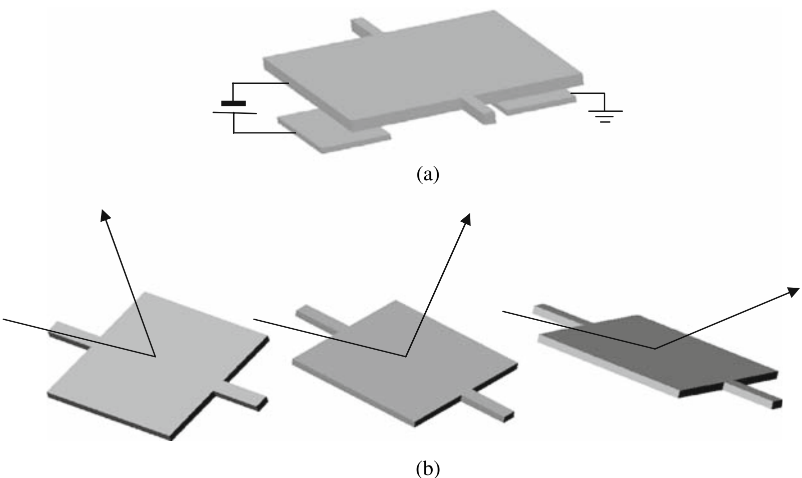


Fig. 2.52 (**a**) Micromirror structure. (**b**) White, grey, and black pixels, represented by different deflections of mirror surface

An aperture in an optical imaging system can control which angle of reflected light can pass through. A flat mirror represents a "white" pixel, and a mirror deflected by one-fourth of wavelength at the edge of mirror represents a "black" pixel. Deflection in between represents a "grey" pixel, as schematically illustrated in Fig. 2.52b. Therefore, the micromirror array is commonly called spatial light modulator (SLM). There are other types of SLM that have been adopted for optical maskless lithography, including grating light valve (GLV) for phase modulation [102] and liquid crystal light valve for intensity modulation [103].

The SLM-based laser lithography tool works like an optical stepper. A SLM chip, consisting of an array of 2048×512 micromirrors with individual mirror size of 16×16 μm, is controlled by a pattern generator in which pattern data is converted into electrical signal to control the deflection of each micromirror to produce a reflected light pattern when a laser beam is shining onto it. The SLM is uniformly illuminated by light from a pulsed excimer laser which has been processed through a beam homogenizer. The reflected light pattern is imaged through a demagnifying optical system onto the photoresist-coated substrate. An interferometer-controlled air-bearing stage moves continuously in the x-direction and the laser is triggered at the exact time to stitch the exposed field with the previous exposure. In the time between two sequential exposures, the SLM is reloaded with new data corresponding to the pattern to be written. After a stripe of fields has been exposed over the full length of the substrate, the stage moves in the y-direction and a new stripe is exposed next to the previous one.

There are three important differences between the multibeam-based laser tool and the SLM-based laser tool.

1. Multibeam laser scanning system relies on a cw as a light source. The wavelength of a cw laser, currently the frequency-doubling Ar ion laser, can only get down to 257 nm. The SLM-based laser tool works in pulse mode. It can use an excimer laser as the light source, and therefore can utilize the 193-nm wavelength. Short wavelength provides potential for high resolution.
2. The SLM-based tool works in stepping manner with a group of images, represented by the light reflection patterns of the SLM chip, projecting onto a substrate, which is faster than scanning sequentially over the pattern area by a multibeam tool.
3. The SLM-based tool uses high reduction imaging optics, which can provide much small pixel size. The first generation of SLM laser tool has reduction of $\times160$ [101]. The new generation SLM laser tool can offer reduction of $\times267$. With a micromirror size of 8×8 μm, the achievable minimum pixel size is 30 nm [98].

The SLM-based laser tools, represented by the Sigma-7500 series from Micronic Laser System AB, have become a serious competing technology to ALTA tools (Etec Systems) for making photomasks. The Sigma-7500 tool is now capable of making masks for the 45-nm generation of IC manufacturing [104]. At 30 nm pixel size, it is also possible to use the technology for direct

maskless lithography of ICs at sub-100 nm resolution. The most important advantage of SLM-based maskless lithography is that most of the RETs, such as phase-shifting, OPC, SRAF, or CPL, can be easily realized by a SLM, because each micromirror pixel can be programmed to change the intensity as well as the phase of incoming light. Isolated lines of 35 nm in size have been produced by SLM-based maskless lithography when combining with the SLM-enabled RETs. There have been sufficient evidence to believe that optical maskless lithography is a viable technology, capable of competing with mask-based optical lithography in terms of printing performance at sub-100 nm resolution. Although the throughput is low, estimated about 5 wafers (300 mm diameter) per hour, optical maskless lithography can still provide an economical alternative for low-volume runs and in cases where throughput is less critical. In addition, it can also be an alternative for printing the most difficult layers, on a wafer-to-wafer basis with the same tool, due to its flexibility to print in all imaging modes utilizing phase-shifting and strong OPC [98].

While the above optical ML2 technologies are well developed by industry and very much geared toward volume production of photomasks or ICs, there is another laser-based lithography technique developed in research laboratories, which can directly form 3D structures without using a photomask. This is the multi-photon lithography (MPL). Multi-photon lithography is based on photon-induced polymerization process to solidify a liquid polymer upon irradiation by a laser beam. It is very similar to a microscale photon-polymerization technique, called microstereolithography [3]. While microstereolithography uses a single laser beam to scan a surface layer of liquid polymer to form a 2D pattern, MPL uses two or more laser beams to form a focus point inside a pool of liquid polymer. The MPL process is illustrated in Fig. 2.53. When two laser beams focused on the same spot, the focusing point has the highest photon energy, which can solidify the liquid polymer if the peak energy in the focus point is over the threshold of photon-polymerization. By moving the laser focus point around

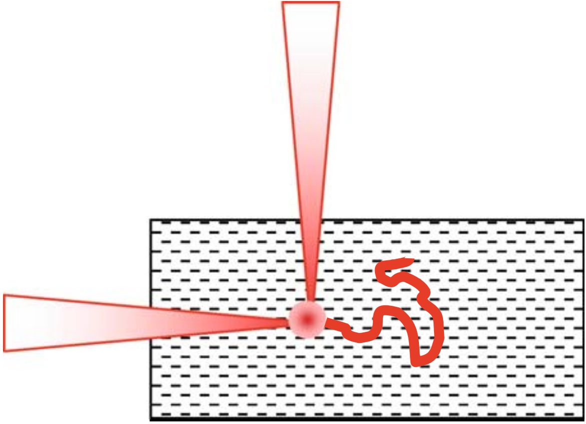


Fig. 2.53 Schematic of two-photon lithography process

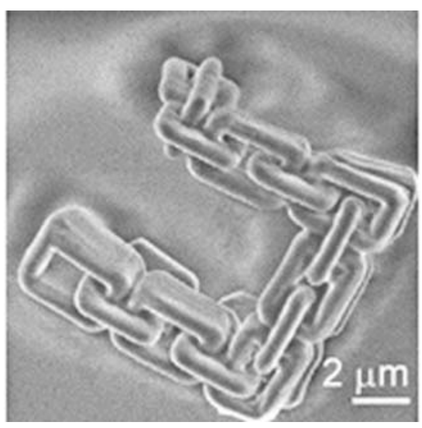


Fig. 2.54 A chain-like 3D structure formed by two-photon polymerization of liquid resin (Reprint from [105] with permission)

in the liquid polymer, a 3D solid object can be formed. Figure 2.54 shows a chain-like 3D structure formed by two-photon polymerization of a liquid resin. The smallest structure dimension is only 120 nm.

The key condition to realize MPL is that the laser beams do not cause significant absorption of photon energy in the liquid polymer except at the focus point of two laser beams. In the past infrared (IR) lasers were used for MPL because only IR wavelength has the minimum absorption in photon-polymerizable liquid resins. Recently, new photon-sensitive polymers have been discovered, which can have peak absorption of photon energy at shorter wavelength. Visible femtosecond lasers at wavelengths of 520 nm and 730 nm were used. By controlling the laser exposure dose, 3D polymer “woodpile” structures have been formed with the minimum line diameters around 60–70 nm, as is shown in Fig. 2.55. Such structure can be useful as a photonic crystal material.

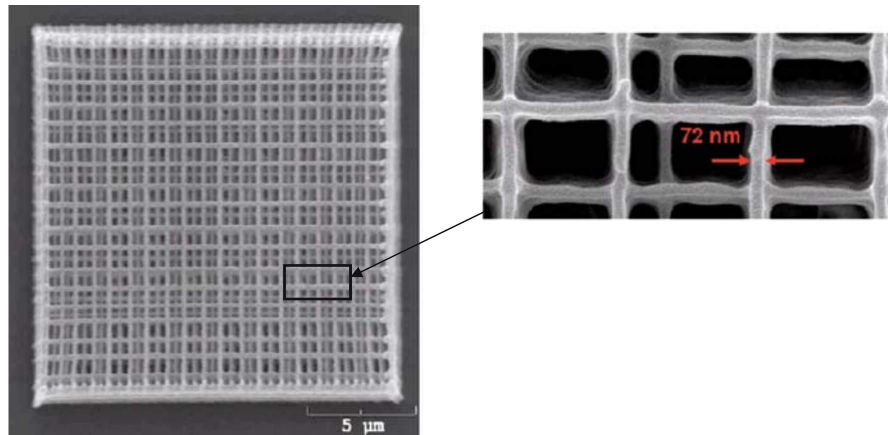


Fig. 2.55 “Woodpile” structure formed by two-photon exposure of liquid polymer (Reprint from [106] with permission)

References

1. Lin, B., *The ending of optical lithography and the prospects of its successors*. Microelectronic Engineering, 2006. **83**: pp. 604–613.
2. Attwood, D., *Soft X-Rays and Extreme Ultraviolet Radiation: Principles and Applications*. 2000, Cambridge University Press.
3. Cui, Z., *Micro-Nanofabrication Technologies and Applications*. 2006, Springer.
4. Wong, A.K., *Optical Imaging in Projection Microlithography*. 2005, SPIE Press.
5. Cui, Z., *Computer simulation of optical lithography in VLSI fabrication*, in *Symposium of Chinese Young Scholars on Optical Application Technologies*. 1994.
6. Burnett, J.H., Z.H. Levine, and E.L. Shirley, *Intrinsic birefringence in calcium fluoride and barium fluoride*. Phys. Rev. B, 2001. **64**: p. 241102.
7. Hatzakis, M., B.J. Canavello, and I.M. Shaw, *Single-step optical lift-off process*. IBM J. Res. Develop., 1980. **24**(4): pp. 452–460.
8. Bakshi, V., (Editor), *EUV Sources for Lithography*. 2006, SPIE Press.
9. Banine, V. and R. Moors, Plasma sources for *EUV* lithography exposure tools. J. Phys. D: Appl. Phys., 2004. **37**: p. 3207.
10. Gargini, P., et al., *International EUV Initiative (IEUVI) overview: challenges and collaborative efforts*. Future Fab Intl., 2006. **21**.
11. Jonkers, J., *High power extreme ultra-violet (EUV) light sources for future lithography*. Plasma Sources Sci. Technol., 2006. **15**: pp. S8–S16.
12. Tomofuji, T., et al., *Mo/Si multilayer (ML) mirror deposited with ion beam sputtering using Kr gas*, in *3rd international EUVL symposium*. 2004.
13. Stivers, A., et al., *EUV mask defects strategy*, in *EUV mask development seminar*. 2002.
14. Naulleau, P., et al., *Lithographic characterization of the printability of programmed extreme ultraviolet substrate defects*. J. Vac. Sci. Technol., 2003. **B21**(4): p. 1286.
15. Liang, T., et al., *Demonstration of damage-free mask repair using electron beam-induced processes*. Proc. SPIE, 2004. **5446**: pp. 291–300.
16. Hand, A., *EUVL Results Show Promise, But Still Many Challenges*. Semiconductor International, 2007 (September).
17. Gwyn, C.W. and P.J. Silverman, *EUV lithography: transition from research to commercialization*. Proc. SPIE, 2003. **5130**: pp. 990–1004.
18. Silverman, P.J., *Extreme ultraviolet lithography: overview and development status*. J. Microlith. Microfab. Microsyst., 2005. **4**(1): p. 011006.
19. Smith, H.I. and F. Cerrina, *X-ray lithography for ULSI manufacturing*. Microlithography World, 1997. **6**(1): p. 10.
20. Wind, S., et al., *Lithography and fabrication processes for sub-100 nm scale complementary metal-oxide semiconductor*. J. Vac. Sci. Technol., 1995. **B13**(6): p. 2688.
21. Krasnoperva, A.A., et al., *Imaging capability of proximity X-ray lithography at 70 nm ground rule*. Proc. SPIE, 1999. **3676**: p. 24.
22. K. Early, M.L. Schattenberg, and H.I. Smith, *Absence of resolution degradation in x-ray lithography*. Microelectronic Engineering, 1990. **11**: pp. 317–321.
23. Vladimirsky, Y., et al., *Demagnification in proximity X-ray lithography and extensibility to 25 nm by optimizing Fresnel diffraction*. J. Phys. D: Appl. Phys., 1999. **32**: p. L114.
24. Ayon, A.A., R.L. Bayt, and K.S. Breuer, *Deep reactive ion etching: a promising technology for micro- and nanosatellites*. Smart Mater. Struct., 2001. **10**: pp. 1135–1144.
25. Lin, B.J., The k3 coefficient in nonparaxial lambda/NA scaling equations for resolution, depth of focus, and immersion lithography J. Microlitho, Microfab. Microsyst., 2002. **1**: p. 7.
26. Fay, B., *Advanced optical lithography development, from UV to EUV*. Microelectronic Engineering, 2002. **61–62**: pp. 11–24.
27. *Technology backgrounder: Immersion Lithography* published by *IC Knowledge*. LLC 2003.
28. Smith, B.W. and J. Cashmore, *Challenges in high NA, polarization, and photoresists*. Proc. SPIE, 2002. **4691**.

29. McCallum, M., G. Fuller, and S. Owa, *From hyper NA to low NA*. Microelectronic Engineering, 2006. **83**: pp. 667–671.
30. Hewett, J., *Immersion ideas extend optical lithography*. optics.org., 2006 (April).
31. Tounai, K., et al., *Resolution improvement with annular illumination*. Proc. SPIE, 1992. **1674**: p. 1753.
32. Shiraishi, N., et al., *New image technique for 64 M-DRAM*. Proc. SPIE, 1992. **1674**: p. 1741.
33. Levenson, M.D., N.S. Viswanathan, and R.A. Simpson, *Improving resolution in photolithography with a phase-shifting mask*. IEEE Trans. Electron. Devices, 1982. **ED-29**: p. 1828.
34. Pelka, J. and W. Henke, *Simulation and optimization of phase-shift masks for printing of contact holes*. Microelectronic Engineering, 1994. **26**: pp. 1–26.
35. Cui, Z., P.D. Prewett, and S. Johnson, *Transmission and side-lobe effect in attenuated phase shift masks*. Microelectronic Engineering, 1995. **27**: p. 259.
36. Hand, A., *Advanced Masks Help Keep Photolithography Alive*. Semiconductor International, 2007 (September).
37. Van Den Broeke, D., et al., *Complex 2D pattern lithography at $\lambda/4$ resolution using chromeless phase lithography (CPL)*. Proc. SPIE, 2002. **4691**: p. 196.
38. Chen, J.F., et al., *Manufacturing at $k_1 = 0.2$ with chromeless phase lithography*. Solid State Technology, 2002 (June).
39. Levenson, M.D., *Extending the lifetime of optical lithography by wave-front engineering*. Jpn. J. Appl. Phys., 1994. **33**(12B): p. 6765.
40. Cui, Z., et al., *Optical proximity correction by grey tone photolithography*. Microelectronic Engineering, 2000. **53**: p. 153.
41. Mack, C.A., *Scattering bars*. Solid State Technology, 2003 (November).
42. Otto, O.W., et al., *Automatic optical proximity correction: a rule-based approach*. Proc. SPIE, 1994. **2197**: p. 278.
43. Rieger, M.L. and J.P. Stirniman, *Using behavior modelling for proximity correction*. Proc. SPIE, 1994. **2197**: p. 371.
44. Information available from Synopsys Inc. (web link: http://www.synopsys.com/).
45. Gupta, P., et al., *Performance-driven optical proximity correction for mask cost reduction*. J. Micro/Nanolith. MEMS MOEMS, 2007. **6**: p. 031005.
46. Yu, P., S.X. Shi, and D.Z. Pan, *True process variation aware optical proximity correction with variational lithography modeling and model calibration*. J. Micro/Nanolith. MEMS MOEMS, 2007. **6**: p. 031004.
47. Information available from Luminescent Inc. (web link: http://www.luminescent.com/).
48. Saleh, B.E.A. and S.I. Sayegh, *Reductions of errors of microphotographic reproductions by optical corrections of original masks*. Optical Engg., 1981. **20**: pp. 781–784.
49. Pang, L., Y. Liu, and D. Abrams, *Inverse lithography technology (ILT): What is the impact to the photomask industry?* Proc. SPIE, 2006. **6283**.
50. Cui, Z. and J. Du, *Investigation of OPC mask distortion effect*. Proc. SPIE, 2001. **4404**.
51. Hector, S., *Behind the photomask*, in *Global Semiconductor Forum*. 2005.
52. Rai-choudhury, P. (Editor), *Handbook of Microlithograhy, Micromachining and Microfabrication*, Vol. 1. 1997: SPIE Press.
53. Zant, P.V., *Microchip Fabrication, 5th Edition*. 2004: McGraw-Hill.
54. Ito, H. and C.G. Willson, *Polymers in Electronics*, in *Symposium Series 242*, T. Davidson, Editor. 1984, American Chemical Society: Washington, D.C. p. 11.
55. Reichmanis, E. and L.F. Thompson, *Chemistry and processes for deep-UV resists*. Microelectronic Engineering, 1991. **13**: p. 3.
56. Ruede, D., M. Ercken, and T. Borgers, *The impact of airborne molecular bases on DUV photoresists*. Solid State Technology, 2001 (August).
57. Slezak, M., *Multilayer resist strategies*. Solid State Technology, 2003 (July).
58. Nalamasu, O., et al., *Single-layer resist design for 193 nm lithography*. Solid State Technology, 1999 (May).

59. Peters, L., *Photoresists Meet the 193 nm Milestone*. Semiconductor International, 2005 (February).
60. Wei, Y., K. Petrillo, and P.A. Benson, *Evaluating Topcoat Options for Immersion Litho Resists*. Semiconductor International, 2006 (July).
61. Information available from Advanced Surface Microscopy, Inc. (web link: http://www.asmicro.com/).
62. Mack, C.A., *The Lithography Expert: Line-edge Roughness, Part 1*. Microlithography World, 2007 (February).
63. Cutler, C.A., J.F. Mackevich, and J. Li, *Effect of polymer molecular weight on AFM polymer aggregate size and LER of EUV resists*. Proc. SPIE, 2003. **5037**: p. 406.
64. Gallatin, G.M., *Resist blur and line edge roughness*. Proc. SPIE, 2005. **5754**: p. 38.
65. Braun, A.E., *Line Edge Roughness is Here to Stay*. Semiconductor International, 2005 (Feburary).
66. Lassig, S. and E. Hudson, *Integrating dielectric etching with 193 nm resists*. Solid State Technology, 2002 (October).
67. Kim, B.-G., et al., *Beneath the MEEF*. Solid State Technology, 2000 (August).
68. Su, B., et al., *Analyzing and characterizing 193 nm resist shrinkage*. Solid State Technology, 2001 (May).
69. Czech, G., E. Richter, and O. Wunnicke, *193 nm Resists: A Status Report (Part One)*. Future Fab Intl., 2002. **12**.
70. Cao, H.B. and P.F. Nealey, *Comparison of resist collapse properties for deep ultraviolet and 193 nm resist platforms*. J. Vac. Sci. Technol., 2000. **B18**: p. 3303.
71. Miller, M., *Yield-aware designs target production issues*. Solid State Technology, 2005 (February).
72. Kobayashi, S., et al., *Automated hot-spot fixing system applied to the metal layers of 65-nm logic devices*. J. Micro/Nanolith. MEMS MOEMS, 2007. **6**: p. 031010.
73. Ho, J., et al., *Lithography-simulation-based design for manufacturability rule development: an integrated circuit design house's approach*. J. Micro/Nanolith. MEMS MOEMS, 2007. **6**: p. 031008.
74. Mansfield, S., G. Han, and L. Liebmann, *Through-process modeling for design-for-manufacturability applications*. J. Micro/Nanolith. MEMS MOEMS, 2007. **6**: p. 031007.
75. Information available from Clear Shape Technologies, Inc. (web link: http://www.clearshape.com/).
76. Staud, W., et al., *Subwavelength Imaging at* $k_1 < 0.3$. Semiconductor International, 2005 (September).
77. Levenson, M.D., *Double, double, toil and trouble*! Solid State Technology, 2007 (April).
78. ITRS 2006 Update, published by ITRS, 2006 (web link: http://www.itrs.net/).
79. Hand, A., *Double Patterning Wrings More from Immersion Lithography*. Semiconductor International, 2007 (February).
80. Blaikie, R.J., et al., *Nanolithography using optical contact exposure in the evanescent near field*. Microelectronic Engineering, 1999. **46**: p. 85.
81. Alkaisi, M.M., et al., *Sub-diffraction-limited patterning using evanescent near-field optical lithography*. Appl. Phys. Lett., 1999. **75**(22): p. 3560.
82. Barnes, W.L., *Surface plasmon–polariton length scales: a route to sub-wavelength optics*. J. Opt. A: Pure Appl. Opt., 2006. **8**: pp. S87–S93.
83. Luo, X. and T. Ishihara, *Surface plasmon resonant interference nanolithography technique*. Appl. Phys. Lett., 2004. **84**(23): p. 4780.
84. Goodberlet, J.G. and H. Kavak, *Patterning sub-50 nm features with near-field embedded-amplitude masks*. Appl. Phys. Lett., 2002. **81**(7): p. 1315.
85. Fang, N., et al., *Sub-diffraction-limited optical imaging with a silver superlens*. Science, 2005. **308**: p. 534.
86. Blaikie, R.J., D.O.S. Melville, and M.M. Alkaisi, *Super-resolution near-field lithography using planar silver lenses: a review of recent developments*. Microelectronic Engineering, 2006. **83**: p. 723–729.

87. Beesley, M.J. and J.G. Castledine, *The use of photoresist as a holographic recording medium*. Appl. Opt., 1970. **9**: p. 2720.
88. Zaidi, S.H. and S.R.J. Brueck, *High aspect-ratio holographic photoresist gratings*. Appl. Opt., 1988. **27**: pp. 2999–3002.
89. Brueck, S.R.J., *Optical and interferometric lithography – Nanotechnology enablers*. Proc. IEEE, 2005. **93**(10): p.1074.
90. Savas, T.A., et al., *Achromatic interferometric lithography for 100-nm-period gratings and grids*. J. Vac. Sci. Technol., 1995. **B13**(6): p. 2732.
91. Solak, H.H., et al., *Photon-beam lithography reaches 12.5 nm half-pitch resolution*. J. Vac. Sci. Technol., 2007. **B25**: p. 91.
92. Campbell, M., et al., *Fabrication of photonic crystals for the visible spectrum by holographic lithography*. Nature, 2000. **404**: pp. 53–56.
93. Ross, C.A., et al., *Magnetic behavior of lithographically patterned particle arrays*. J. Appl. Phys., 2002. **91**: pp. 6848–6853.
94. Pang, Y.K., et al., *Chiral microstructures (spirals) fabrication by holographic lithography*. Opt. Express, 2005. **13**(19): p. 7615.
95. Fritze, M., et al., *Hybrid optical maskless lithography: Scaling beyond the 45 nm node*. J. Vac. Sci. Technol., 2005. **23**(6): p. 2743.
96. Chen, X. and S.R.J. Brueck, *Imaging interferometric lithography: a wavelength division multiplex approach to extending optical lithography*. J. Vac. Sci. Technol., 1998. **B16** (6): p. 3392.
97. Resor, G., *Burn Lin's wish list for optical lithography: eliminate the mask*. Solid State Technology, 2007 (March).
98. Martinsson, H., et al., *Current status of optical maskless lithography*. J. Microlith. Microfab. Microsyst., 2005. **4**(1): p. 011003.
99. Burns, G.A. and J.A. Schoeffel, *Performance evaluation of the ATEQ CORE-2000 scanning laser reticle writer*. Proc. SPIE, 1987. **772**: pp. 55–64.
100. Jackson, C.A., et al., *DUV laser lithography for photomask fabrication*. Proc. SPIE, 2004. **5377**: pp. 1005–1016.
101. Ljungblad, U., et al., *New laser pattern generator for DUV using a spatial light modulator*. Microelectronic Engineering, 2001. **57–58**: pp. 23–29.
102. Gil, D., et al., *Parallel maskless optical lithography for prototyping, low-volume production, and research*. J. Vac. Sci. Technol., 2002. **B20**(6): p. 2597.
103. Klosner, M. and K. Jain, *Massively parallel, large-area maskless lithography*. Appl. Phys. Lett., 2004. **84**(15): p. 2880.
104. Yao, M.-J., et al., *Application of Sigma7500 pattern generator to X architecture and 45-nm generation mask making*. Proc. SPIE, 2007. **6607**.
105. Tanaka, T., H.-B. Sun, and S. Kawata, *Rapid sub-diffraction-limit laser micro nano processing in a threshold material system*. Appl. Phys. Lett., 2002. **80**(2): p. 312.
106. Haske, W., et al., *65 nm feature sizes using visible wavelength 3-D multiphoton lithography*. Opt. Express, 2007. **15**(6): p. 3426.

Chapter 3
Nanofabrication by Charged Beams

3.1 Introduction

Nanofabrication by charged beams, including electron beam (e-beam) and ion beam, is very much like what photons do. They carry energy and transfer the energy into a substrate material to perform structuring either by exposure of energy-sensitive polymer materials such as photoresists or by removing material directly. Charged beams have an important role in nanofabrication because they can be focused into extremely small probes using sophisticated charged particle optics.

Electron beam (e-beam) lithography appeared in early 1960s almost around the same time as the optical lithography. Electron beam lithography evolved from scanning electron microscope (SEM). An e-beam lithography system is very much like a SEM. The SEM could evolve to e-beam lithography only because of the discovery of an electron-sensitive polymer material, called polymethylmethacrylate (PMMA) [1]. Exposure of PMMA resist by e-beam is very much like exposure of photoresists by photons. Nanofabrication realized by e-beam lithography was much earlier than photons. Because of the fine beam enabled by electron optics, e-beam lithography using PMMA resist immediately demonstrated much higher resolution capability than optical lithography could dream of around the time. In the early 1970s, e-beam lithography was already able to pattern features as small as 60 nm [2]. Nowadays, the state-of-the-art e-beam lithography system, combined with special e-beam resist materials and processes, can fabricate structures less than 10 nm [3].

Ions are different from electrons because of their heavy mass. A hydrogen ion is 1840 times heavier than an electron. The heavy mass of ions makes them particularly suitable for nanofabrication by direct structuring of material, rather than by exposure of a polymer resist as e-beam lithography does. Ions were discovered and made use of in many applications, other than nanofabrication, many years earlier than electrons. In the early days, nanofabrication utilizing ion beam was hampered because a fine size of ion beam was difficult to achieve from the early gas discharge ion sources. Focused ion beam (FIB) really became a nanofabrication tool following the introduction of liquid metal ion sources (LMISs) [4]. The LMIS made it possible to focus an ion beam to less

Z. Cui, *Nanofabrication*, DOI: 10.1007/978-0-387-75577-9_3,

© Springer Science+Business Media, LLC 2008

than 5 nm in beam size. With such a fine beam and its heavy mass, FIB can directly remove materials from surface, as well as deposit materials with the assistance of some chemical compound gases, with extremely high precision. Focused ion beam is a truly flexible and versatile nanofabrication tool.

Although both e-beam and ion beam can form a fine probe and are able to pattern sub-100 nm features whether by irradiation exposure or by direct machining, they are fundamentally handicapped when applied to volume manufacturing of integrated circuits (ICs) because of their low throughput. The serial scanning manner of e-beam or FIB cannot compete with the parallel projection of photons in optical lithography. The industry has been trying hard to make parallel projection versions of electrons and ions over the last decade. Electron or ion projection lithography (IPL) has been vigorously pursued because they can potentially offer higher resolution than optical lithography, based on the fact that they do not have the diffraction limitation associated with the optical projection lithography. Limited success has been achieved, by which it means the throughput has been much improved but still not on par with optical projection lithography. There are other technical difficulties associated with electron or ion project lithography technology, notably the fabrication and robustness of the masks required for projection of electrons or ions. Another route toward parallel projection of electrons of ions is to use multibeams, whereby thousands of beams are projected instead of a single beam to boost the lithography throughput.

The basics of e-beam lithography and FIB technology have been introduced in the previous book [5]. This chapter goes beyond the basics and focuses on those issues associated with patterning at sub-100 nm scale when e-beam lithography and FIB are used. For example, to achieve high-resolution e-beam lithography, the following conditions are generally preferred:

- high electron energy,
- small scanning field,
- low beam current,
- low-sensitivity resist,
- thin resist layer,
- optimized resist process,
- low pattern density,
- light and conductive substrate material, and
- stable environment.

Why these are desirable conditions is of most interest to the users of e-beam lithography. To answer the question, one has to look at how the e-beam lithography works and what are the factors that can influence the ultimate resolution of e-beam lithography. These factors are related to the performance of e-beam lithography systems, the e-beam resist processes, and the nature of electron scattering in resist and substrate materials. The same applies to FIB technology. This chapter will review the e-beam and FIB patterning techniques

developed so far and emphasize on their capabilities and the limiting factors toward achieving fabrication of sub-100 nm feature dimension.

3.2 Focusing Charged Particle Beam

3.2.1 Charged Particle Optics

To achieve high-resolution patterning using either an e-beam or an ion beam, the beam itself has to be small. The fine beam size can be obtained from an electron optical or ion optical system. Though the complete theories of electron optics and ion optics are complex, it can be conceptually understood from the Lorentz force equation as shown in Eq. 3.1. The equation expresses the force exerted to a particle with charge q and speed v in a combined electrostatic field $\mathbf{E}$ and magnetic field $\mathbf{B}$:

$$F = q(\mathbf{E} + \nu \times \mathbf{B}). \tag{3.1}$$

For simplification, assuming the electrostatic field and magnetic field are rotationally symmetric, which is true in most of the charged particle optical systems, the movement of charged particles can be described by the following equations [6].

In a rotationally symmetric electrostatic field $E(z,r)$

$$m\frac{d^2z}{dt^2} = qE_z = q\frac{\partial V(r,z)}{\partial z}, \tag{3.2}$$

$$m\frac{d^2r}{dt^2} = qE_r = q\frac{\partial V(r,z)}{\partial r} \tag{3.3}$$

The charged particles experience a radial force in r direction to pull them toward the axis, as well as an acceleration force in z direction. It is the radial force that focuses charged particles into a fine beam. The radial movement of charged particles as a function of z can be derived from Eqs. (3.2) and (3.3) as

$$r'' = \frac{(1 + r'^2)}{2V(r,z)}\left[\frac{\partial}{\partial r}V(r,z) - r'\frac{\partial}{\partial z}V(r,z)\right], \text{ where } r'' = \frac{\mathrm{d}^2 r}{\mathrm{d}z^2};\ r' = \frac{\mathrm{d}r}{\mathrm{d}z} \tag{3.4}$$

The Eq. 3.4 reveals three characteristics of focusing charged particles in an electrostatic field.

1. The focusing does not depend on charge or mass of a particle. Whether electrons or ions of different mass they will experience the same focusing action.
2. The focusing is directly proportional to the variation of potential distribution along z direction. Figure 3.1a shows a typical electrostatic lens which

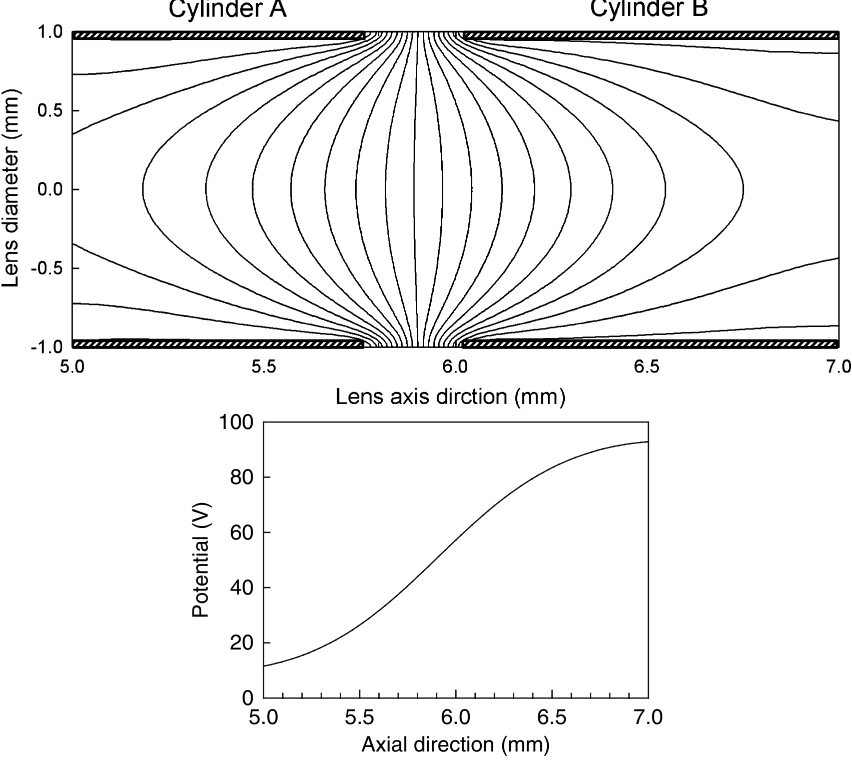


Fig. 3.1 (**a**) Potential contours inside two cylindrical electrodes with different voltages (*Cylinder A*: 0 V; *Cylinder B*: 100 V). **(b)** Potential distribution along z-axis ($r = 0$)

consists of two cylinders with different potentials. The contours of potential distribution inside the cylinders are plotted. Figure 3.1b is the potential variation along z-axis. It is this variation that gives the twin cylinders its focusing lens function.

3. High potential $V(r,z)$ reduces the focusing action of a lens because the charged particles pass through the lens in high speed.

In the case of a rotationally symmetric magnetic field $B(z, r)$

$$m\frac{\mathrm{d}^2 z}{\mathrm{d}t^2} = qr\frac{\mathrm{d}\phi}{\mathrm{d}t}B_r, \tag{3.5}$$

$$m\frac{\mathrm{d}^2 \mathrm{r}}{\mathrm{d}t^2} = -qr\frac{\mathrm{d}\phi}{\mathrm{d}t}B_z, \tag{3.6}$$

$$F_\phi = -q\frac{\mathrm{d}z}{\mathrm{d}t}B_r + q\frac{\mathrm{d}r}{\mathrm{d}t}B_z, \text{ where } F_\phi = \frac{m}{r}\frac{\mathrm{d}}{\mathrm{d}t}(r^2\frac{\mathrm{d}\phi}{\mathrm{d}t}). \tag{3.7}$$

These equations show that charged particles are not only focused toward the axis, but also rotate around the axis (rotational angleϕ). $\mathrm{d}\phi/\mathrm{d}t$ is the rotational speed of electrons. The trajectory of charged particles in the rotationally symmetric magnetic field can be derived from the above equations.

$$\frac{\mathrm{d}\phi}{\mathrm{d}t} = \frac{q}{2m} B_z, \tag{3.8}$$

$$\frac{\mathrm{d}\phi}{dz} = \sqrt{\frac{q}{8mV}} B_z, \tag{3.9}$$

$$\frac{d^2 r}{dz^2} = \frac{q}{8mV} r B_z^2. \tag{3.10}$$

Compared with the movement of charged particles in an electrostatic field, their movement in a magnetic field depends on their charge and mass. This will have serious effect on focusing ions in a magnetic lens. While moving toward the axis, the charged particles also rotate around the axis. The electrical potential V, which is considered to be constant in a pure magnetic lens, will reduce the lens power because it increases the speed of charged particles passing through a lens.

For electrostatic lenses, the focusing power depends on how much is the difference of the voltage applied to the lens elements and how rapid is the change of potential. This requires high voltage and very narrow gap between the lens elements, for example, the two cylinders in Fig. 3.1a. Then there is the risk of electrical breakdown if the voltage difference is too great. Therefore, electrostatic lens is in general limited in focusing power. On the other hand, high magnetic field can be generated by high current passing through coils. The heat generated by the high current can be cooled by forced circulation of water or air. In this way, high focusing power can be achieved with a magnetic lens. This is the reason that all the e-beam lithography systems, as well as electron microscopes, use magnetic lenses to focus e-beams. However, magnetic lens cannot be used for focusing an ion beam because of its dependence on charge and mass of ions. A magnetic lens will focus ions of different charges or masses to different positions on an image plane. Such a property has been employed for separation of ions in mass spectrometers. For focusing a fine ion beam, electrostatic lenses are generally used.

3.2.2 Sources

3.2.2.1 Electron Sources

A fine beam spot focused by an electron or ion optical system is the image of an electron source or ion source. To achieve a very small size of beam, the size of

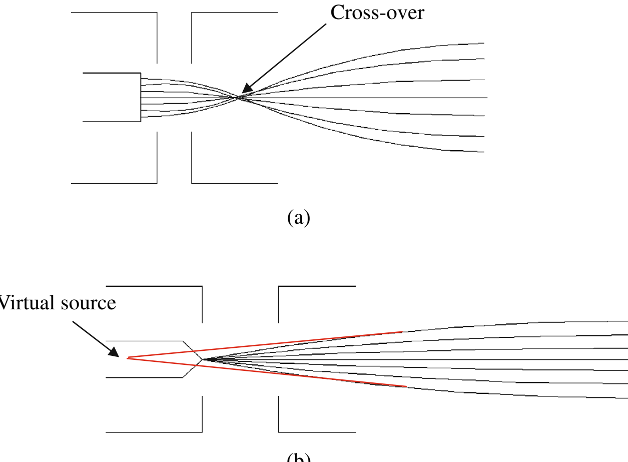


Fig. 3.2 (**a**) Virtual source (the cross-over) in a thermionic cathode. (**b**) Virtual source in a FE cathode

its source has to be small. Before the advent of field emission (FE) electron sources, tungsten thermal cathodes were the commonly used electron emission source. Lately, LaB_6 compound was employed as thermal cathode material which performs better than tungsten cathodes. The fundamental drawback of thermionic cathode is their large emission area. In order to reduce the source size, a strong lens in front of the cathode has to be used to focus the electrons emitted from the cathode into a much small cross-over, as illustrated in Fig. 3.2a. This cross-over acts as a virtual source seen by the downstream lenses, so that a fine spot size can be obtained through the focusing system. The problem with this method is that it is difficult to achieve a very small diameter of cross-over because of the Coulomb repulsion of electrons arising from the high concentration of electrons at the cross-over. Field emission sources have completely overcome this problem. A FE source does not rely on high-temperature heating of the source to generate electron emission. The electrons are directly extracted from the material surface by extremely high electrical field ($>10^8$ V m^{-1}). A FE source is a fine needle. Very high electrical field can be created at the sharp end of needle with relatively low applied voltage [7]. Because electrons only emit at the points where the electrical field exceeds threshold strength, the emission area can be finely controlled and very small. With a FE source, there is no need to have a cross-over. The virtual source extends to the back of real source, as illustrated in Fig. 3.2b. The virtual source is only a few nanometers. Very high resolution can be achieved using such small virtual source as the object of electron optical imaging system.

Table 3.1 gives a comprehensive list of technical parameters for the four main electron sources which are widely used in various e-beam systems [8]. First, it can be seen that FE sources have much small sizes of both the real cathode geometry

Table 3.1 Comparison of electron emission sources

Emitter type	Thermionic	Thermionic	Cold FE	Schottky FE
Cathode materials	W	LaB_6	W	ZrO/W
Operating temperature (K)	2800	1900	300	1800
Cathode radius (μm)	60	10	<0.1	<1
Virtual source radius (nm)	15,000	5000	2.5	15
Emission current density (A cm^{-2})	3	30	17,000	5300
Total emission current (μA)	200	80	5	200
Brightness	10^4	10^5	2×10^7	10^7
Maximum probe current (nA)	1000	1000	0.2	10
Energy spread at cathode (eV)	0.59	0.40	0.26	0.31
Energy spread at gun exit (eV)	1.5–2.5	1.3–2.5	0.3–0.7	0.35–0.7
Beam noise (%)	1	1	5–10	1
Emission current drift (% hr^{-1})	0.1	0.2	5	<0.5
Vacuum requirement (Torr)	$\leq10^{-5}$	$\leq10^{-6}$	$\leq10^{-10}$	$\leq10^{-8}$
Cathode life (h)	200	1000	2000	2000
Cathode regeneration	Not required	Not required	Every 6–8 h	Not required
Sensitivity to external influence	Minimal	Minimal	High	Low

and the virtual source. Second, their emission current density is several orders of magnitude higher than the thermionic sources. The cold FE source appears to be the best among the four types of sources. All high-grade electron microscopes nowadays are using cold FE sources. However, it is rarely used as the electron emission source in an e-beam lithography system. This is because the emission surface of a cold cathode can easily absorb gaseous atoms, which can affect the electron emission, causing noise and drift of emission current. As seen in Table 3.1, the noise and emission current drift of a cold FE source is the highest among the four types of sources. These effects are not serious problems for electron microscope because a flickering of image does not matter very much. But they are unacceptable to e-beam lithography application. E-beam lithography relies on stable emission current to deliver consistent exposure dose. Any emission current drift or noise can seriously deteriorate the quality of resist exposure. This has been the reason to develop thermal FE cathodes for e-beam lithography systems. The thermal FE cathode, also called Schottky cathode, is to heat up the conventional FE cathode. At high temperature, absorption of gaseous atoms at cathode emission surface is greatly reduced, resulting in stable working of the cathode. As seen in the table, Schottky cathode has most characteristics of a FE source, but is much lower in beam noise and emission current drift.

3.2.2.2 Liquid Metal Ion Sources

Ions can be focused into a fine beam size and find its place in nanofabrication only because of the LMISs. The basic structure of a LMIS is illustrated in Fig. 3.3. A tungsten wire of about 0.5-mm diameter is electrochemically etched

into a needle with the tip radius of only 5–10 μm. The tungsten needle is then wetted with molten metal. Gallium is the most used liquid metal material because of its low melting temperature (29.8°C). A small amount of liquid metal adheres to the needle. When a high voltage is applied, an electrostatic force exerts on the liquid metal, pulling it into an extremely small apex [9]. The electrical field strength at the liquid apex can be as high as 10^{10} V m^{-1}. At such extremely high field, metal atoms at the liquid apex become ionized and escape from the liquid metal surface in the form of field evaporation, resulting in ion emission. Although the total emission current may be only a few microamperes, the current density can be as high as $\sim 10^6$ A cm^{-2} because of the extremely small emission area.

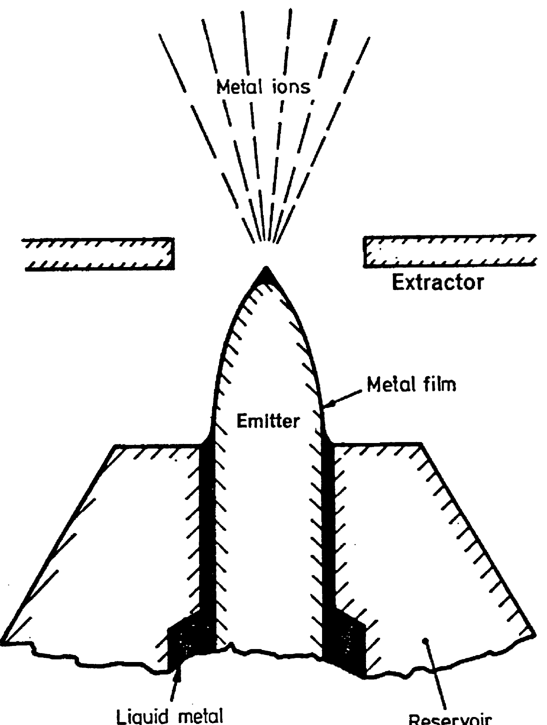


Fig. 3.3 Schematic of LMIS (Reprint from [10] with permission)

Liquid metal ion source is a FE ion source. Like the FE electron sources mentioned earlier, LMIS has high emission current density (high brightness) and extremely small size of emission area. These are favorable characteristics for achieving fine focus of ion beam. However, the emitted ions from a LMIS have much larger energy spread, ~15 eV compared to less than 1 eV for field electron emission. Liquid metal ion source also has large angular emission, around 30°. These are unfavorable characteristics which makes it much hard to focus an ion beam as small as an e-beam.

3.2.3 Aberrations

A small size of source does not guarantee a final small size of beam. An electromagnetic lens system with minimum aberration is necessary to achieve a fine beam size. Similar to light optics, charged particle optics also has geometrical aberrations, such as spherical aberration, coma, field curvature, astigmatism and distortion, and nongeometrical aberration such as chromatic aberration. A fundamental difference between charged particle optics and light optics is that the refractive index in an optical lens is a discontinuous function of space coordinates; that is, light entering from air to an optical lens undergoes a discontinuous refraction, while the refractive index of an electromagnetic lens is a function of electromagnetic field distribution, which is a continuous function of space coordinates. The geometrical aberrations in light optics can be eliminated by changing arbitrarily the curvature of refractive surface (geometrical shape of an optical lens). However, electromagnetic field in space cannot be arbitrarily changed. An optical imaging system may consist of a hundred pieces of lenses to correct all the aberrations. An electromagnetic imaging system has only a few lenses. A typical e-beam lithography column is schematically illustrated in Fig. 3.4 [3]. The electron gun is at the top, which consists of the electron source and lenses to initially focus the electrons. The condenser lens further focuses the beam, and in conjunction with the beam blanker and selectable aperture, the beam can be switched on or off (pass or not pass the aperture). The deflectors perform the beam scanning function. The final objective lens focuses the beam onto a substrate surface.

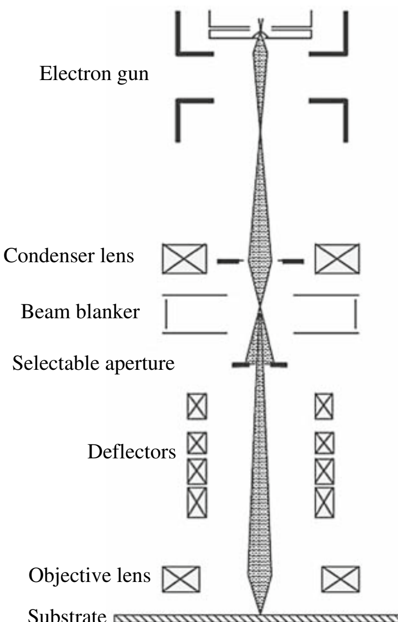


Fig. 3.4 Schematic of complete electron optical system in an e-beam lithography column (Reprint from [3] with permission)

Due to aberrations, an e-beam spot on-axis (off-axis deflection is not considered here) at the substrate (image plane) is not simply a demagnified electron source but includes other aberration components, as expressed in Eq. (3.11) [11].

$$\text{Beam spot diameter} \quad d = \sqrt{d_g^2 + d_s^2 + d_c^2 + d_d^2}, \tag{3.11}$$

where

beam spot without aberration $d_g = \frac{d_\mathrm{v}}{M}$ (d_v, virtual source diameter; M, demagnification),

spherical aberration $d_\mathrm{s} = \frac{1}{2} C_\mathrm{s} \alpha^3$ (C_s, spherical aberration coefficient; α beam convergence angle),

chromatic aberration $d_\mathrm{c} = C_\mathrm{c} \alpha \frac{\Delta V}{V}$ (C_c, chromatic aberration coefficient; ΔV, beam energy spread; V, beam energy), and

diffraction limit $d_d = 0.6 \ \frac{\lambda}{\alpha}$ (λ (nm), electron wavelength $\lambda = \frac{1.2}{\sqrt{V}}$ nm).

Spherical aberration is caused by large convergent angle of beam. Although an electron optical system can be designed to minimize the spherical aberration coefficient C_s, the most effective way of reducing the spherical aberration is to reduce the angle α, as the aberration is proportional to the cubic power of α. The α can be reduced by using an aperture to prevent wide angular electrons from entering the objective lens. The selectable aperture in Fig. 3.4 performs exactly this function. For the objective lens, the aperture is the object to be imaged. Its size has great influence on the final spot size on the substrate. An e-beam lithography system is normally equipped with a series of different sizes of apertures. They are selectable and exchangeable to suite different requirements of beam current and resolution. It is obvious that for ultimate high-resolution e-beam lithography the smallest aperture should be selected, albeit the usable beam current is also greatly reduced. Unfortunately, reduction of convergent angle will cause increase of electron diffraction aberration. Although the wavelength of high-energy electrons is extremely small, as indicated in Eq. (3.11), the diffraction can be noticeable if the aperture is also extremely small. Obviously, increasing beam energy can significantly reduce the diffraction effect.

Chromatic aberration is caused by energy spread of emitted electrons or ions, and by instability of applied voltage in the case of electrostatic lenses or instability of current in the case of magnetic lenses. While the fluctuation of external power supply to an electromagnetic lens system can be well controlled to minimum, the energy spread of an emission source is, however, beyond anyone's control. Therefore, the chromatic aberration is always there. The energy spread can be exasperated if there is a beam cross-over. At the cross-over, high concentration of electrons or ions creates significant

repulsion among the charged particles, causing increase of the energy spread. The evidence is displayed in Table 3.1, where it indicates that electrons leaving a thermionic cathode have energy spread of 0.59 eV but increase significantly to 1.5 –2.5 eV at the electron gun exit, because of the cross-over formed in front of the cathode. Apparently, the cross-over has made the energy spread much worse. Not only the cross-over at the electron gun side, any cross-over along the optical axis of an electromagnetic lens system can create additional energy spread, hence the chromatic aberration. Therefore, an advanced electron optical system will try to avoid any cross-over. For example, the GEMINI column developed at Zeiss does not have any cross-over along the path from a FE source. The only cross-over is the focus point at the substrate surface where the beam lands [8].

For a FIB system using a LMIS, the chromatic aberration is much worse, because of its much higher energy spread (~15 eV) than a field electron emission source (<1 eV). The spherical aberration in a FIB system is also higher because it is proportional to the cubic power of emission angle. Liquid metal ion source is known to have large angular emission. An electrostatic lens system generally has higher spherical aberration coefficient and higher chromatic aberration coefficient than a magnetic lens system. All the combined factors make it much hard to achieve a fine probe size in a FIB system. The size of an ion beam in the best FIB system can never be expected to be as small as that of an e-beam in the best e-beam lithography system.

Another aberration which is unique in charged particle optics is caused by space charge effect. As mentioned earlier, electrons or ions at a cross-over will experience Coulomb repulsion force. The repulsions between charged particles not only increase the energy spread, but also increase the diameter of beam, resulting in defocus and radial blur of the beam [12]. The higher the emission current, the greater the space charge effect, hence the more difficult to achieve small beam diameter in an electromagnetic imaging system. Electron beam lithography has much greater demand on beam current than a SEM, because a resist can only be exposed above a threshold exposure dose, whereas the exposure dose is the product of beam current and beam dwelling time. High beam current can reduce the beam dwelling time, resulting in high throughput in e-beam lithography. However, high beam current makes the focusing of e-beam more difficult. That is partly the reason that a SEM can have much small size of e-beam than can an e-beam lithography system. The other reason is that SEM can use a cold FE source while e-beam lithography system cannot, as explained previously. A modest scanning transmission microscope (STEM) in 1970s had already been able to provide a 0.5-nm diameter e-beam. The current at this beam size was 1 pA (10^{-12} A) [13]. In contrast, an e-beam lithography system can rarely operate below 100 pA beam current. The state-of-the-art e-beam lithography system reported in 2004 was trying hard to get the beam diameter below 3 nm at beam current of 200 pA [3].

The only way to reduce the space charge effect without compromising the beam current is to increase the beam energy – in other words, operating at high acceleration voltage in an e-beam lithography system. High beam energy can also minimize the chromatic aberration caused by beam energy spread, because both the space charge-induced beam enlargement and the chromatic aberration are inversely proportional to the beam energy. This is partly the reason why the most advanced e-beam lithography systems always operate at 100-keV beam energy. The relation between beam current and beam diameter is best demonstrated in Fig. 3.5, where beam diameter versus beam current at 20-keV and 50-keV beam energies are compared [14]. The plots in Fig. 3.5 are only for illustration purpose, though for high-resolution e-beam lithography such high beam currents are rarely used. Therefore, either smaller beam current or higher beam energy, or both, is preferred when high-resolution e-beam lithography is required.

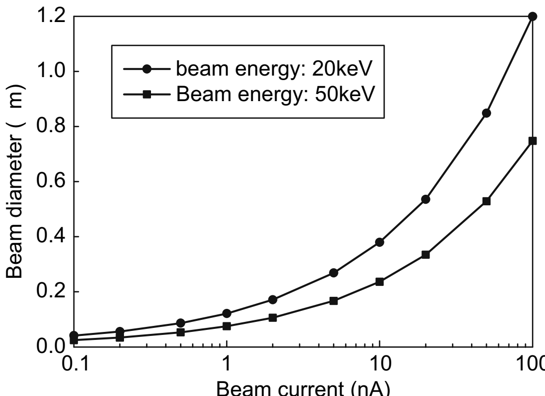


Fig. 3.5 Beam diameter versus beam current

A FIB system has even more stringent restriction on ion beam current. Higher emission current from a LMIS results in greater energy spread and angular emission, both of which are extremely detrimental to fine focusing of the beam. Besides, with increase of ion emission, the total ion current no longer consists of singly charged ions, but includes multicharge ions, charged molecules, ion clusters, and even charged liquid droplets. Only below 10 μA, the beam current is nearly 100% of singly charged ions. So there is no point in having high emission current. A FIB system with gallium LMIS normally operates at below 2-μA emission current, whereby the usable ion current is no less than for operating at high emission regime.

Besides the abovementioned aberrations which are associated with imaging, there are also aberrations related to beam deflection. They are mainly distortion and astigmatism which increase in cubic power proportion with the distance of deflection from the center axis. The beam spot is no longer small and circular at the corners of deflected field. They become bigger and of elliptical shape, and a square scanning field becomes distorted [15]. For example, a previously

reported prototype e-beam lithography system had achieved beam spot diameter of 2.7 nm on-axis, but the beam spot deteriorated to 6.5 nm at the edge of 500-μm deflection field [3]. An e-beam lithography system normally has scan fields from a few hundreds of micrometers to a few millimeters.

The above analysis has more or less answered the questions why high-resolution e-beam lithography and FIB application require high beam energy, small scanning field, low beam current, and a stable environment. These are the prerequisites to low aberrations and to fine focusing of a beam. With a fine charged particle beam, sub-100 nm resolution patterning then becomes possible.

3.3 Scattering and Proximity Effect

To realize nanoscale lithography or direct nanofabrication with charged beams, a small beam size is necessary but it is only half of the story. Charged particles have to interact with a substrate material, in the case of e-beam lithography it is a polymer resist or in the case of FIB patterning it can be any solid materials, so that the energy carried by the charged particles are transferred to the substrate material, causing changes of molecular chains in a polymer or sputtering off atoms from the material surface. The energy transfer process takes place during collisions between charged particles and the atoms in the substrate material. The process is called scattering, which can lead to a much larger area of energy deposition than the original size of beam landing area, creating the so-called proximity effect. The proximity effect is the major obstacle in the process of achieving high-resolution nanofabrication by charged beams.

3.3.1 Electron Scattering

High-energy electrons penetrating into a solid material will collide with atoms and the electrons surrounding an atom. As we have learnt from general physics, two things can happen when two rigid balls collide with each other. If a small ball collides with a massive ball, the small ball will only change the direction of movement without losing any energy it carries. If a small ball collides with another small ball of same mass, the small ball will transfer half its energy to the other ball. The similar scenario happens in electron scattering in solid. The massive balls are the atoms and the small balls are electrons. There are two types of scattering taking place in the solid: elastic and inelastic scatterings. Collision of incident electrons with atoms in the material is an elastic scattering. As the mass of an electron is at least less than 1/1840 of atomic mass, the collision is an extremely small ball with an enormous ball. An electron undergoing elastic scattering does not lose its energy but simply changes its

direction. Collision of incident electrons with the electrons surrounding the atoms of solid material is inelastic scattering, during which the incident electron will pass part of its energy to the electrons of solid material. It is this part of transferred energy that causes chain reaction of resist polymer material, resulting in the exposure of resist. Therefore, the electron scattering process in a solid material can be pictured as this: electrons from an incident e-beam undergo continuous elastic scattering by material atoms and continuously change their flying directions. In between two elastic scatterings there are inelastic scatterings with material electrons whereby the beam energy is gradually deposited into the material.

The rigid ball collision is only an analogy to describe the scattering process taking place during e-beam exposure of resist. Electrons and atoms are not rigid balls. Their collisions are through the interaction of energy field around the atoms and electrons, which can be modeled by Monte Carlo simulation [16, 17]. Monte Carlo method is a numerical technique for solving mathematic equations by a random process. Here it is not to use the method to solve mathematic equations but to borrow its name to reflect the random nature of electron scattering process. There are mathematic descriptions of elastic and inelastic scattering cross-sections and electron energy loss in a solid material, which have long been established [15]. The computer simulation is to calculate how a primary electron from an e-beam is scattered by atoms inside a solid material and how the primary electron loses its energy. The electron scattering is a random process because which atom the primary electron will collide with is nondeterministic; which direction the primary electron will go after the collision is random and how far the electron will travel before hitting next atom is random. Figure 3.6 shows two electron trajectories (primary electron A and B), when they travel inside a resist layer on top of substrate material. The resist is PMMA of 0.1-μm thickness and the substrate is silicon. It was simulated by a Monte Carlo simulation program MOCASEL [18]. Each dot along a trajectory represents a scattering event. One can see that the position, direction, and free travel distance of each scattering event are totally random. Yet the probability of a scattering event taking place and the energy loss of a primary electron are governed by the physical law of electron–atom collisions. Although each individual scattering is random, simulation of a large number of electron scatterings will generate the general picture of the extent of energy deposition due to the collision and loss of energy from the primary electrons. Figure 3.7 plots out 100 electron trajectories. The electron energy is 20 keV with 1-μm resist layer on top of a silicon substrate. This plot gives a rough idea how deep and how wide the electrons travel inside resist and substrate. Most of electrons scatter forward into substrate but some electrons scatter back into resist layer. These forward scatterings and backward scatterings are the cause of broadening of exposure area in resist. To generate statistically correct and reliable results, the number of electrons to be simulated has to be between 50,000 to 100,000.

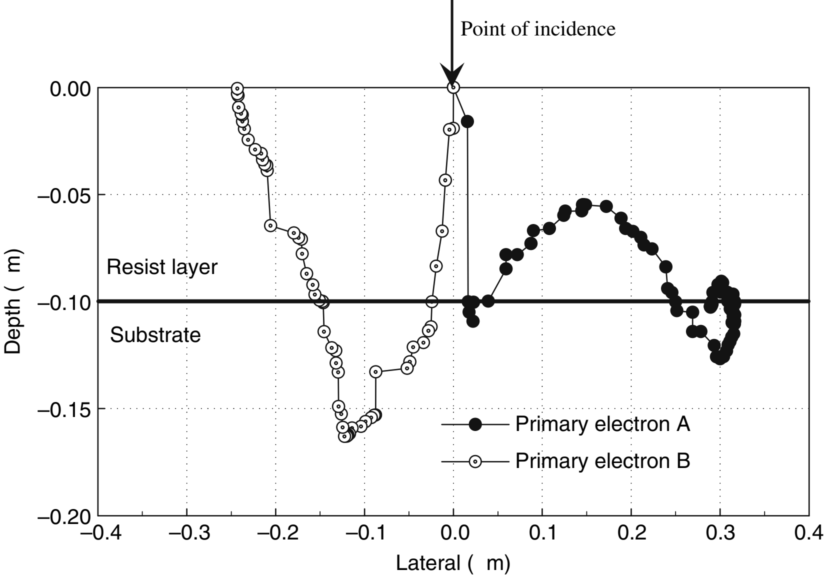


Fig. 3.6 Two electron trajectories showing random scattering events

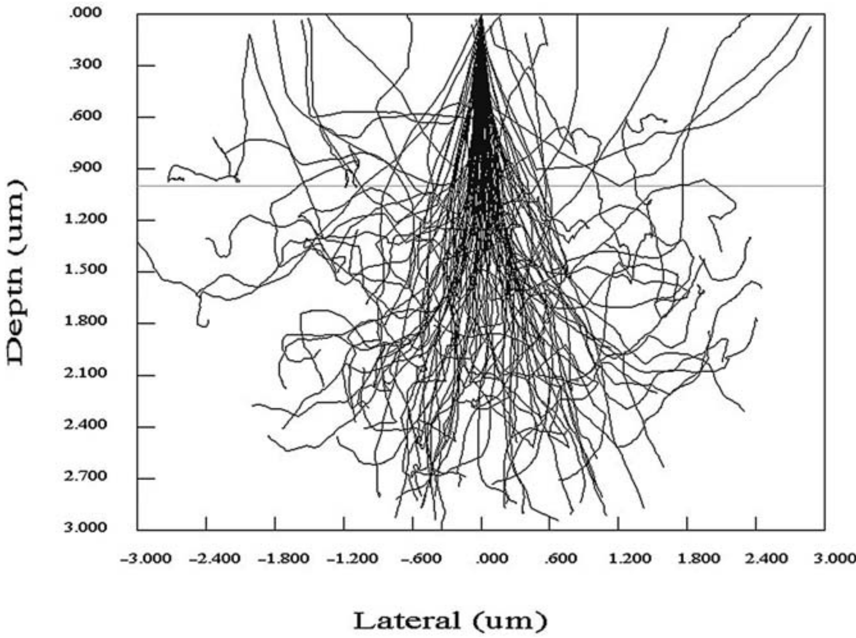


Fig. 3.7 Hundred electron trajectories showing the extent of electron scattering

Instead of counting scattering electrons, forward and backward scattering can also be expressed by a point spread function (PSF) which represents deposited energy distribution in resist from single point of incidence. The typical PSF is shown in Fig. 3.8, where forward scattering part and back-scattering part of energy spread are indicated. The figure also compares the PSFs at 10% and 90% depth of resist. It clearly displays a significant increase of broadening in deposited energy by forward scattering, as well as accumulated energy by backscattering, as electrons going deeper.

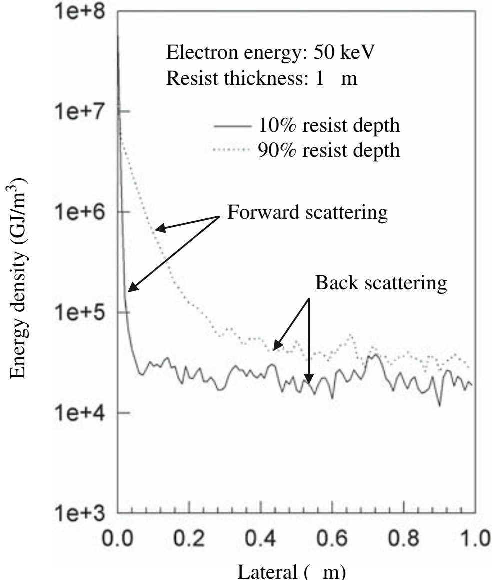


Fig. 3.8 Point spread function of electron energy distribution in resist

There are three factors that strongly influence the electron scatterings.

1. *Electron energy*. Higher energy of primary electrons have much smaller scattering angle, therefore, less broadening of the beam due to small-angle forward scattering. Figure 3.9 compares the PSFs from 50 keV to 100 keV electron energies. The higher the electron energy the narrower the PSF. This is another proof that high-energy e-beam is preferred in e-beam lithography, in addition to previously mentioned benefits in lowering chromatic and space charge aberrations.
2. *Density of substrate material*. More electrons will scatter back from substrate if the substrate material is of higher density. Figure 3.10 compares the PSFs for silicon and GaAs substrates. Electron energies are both at 50 keV and the resist thickness is the same as 0.5 μm. However, the deposited energy level due to backscattering electrons is significantly higher for GaAs substrate than for silicon substrate, because the density of GaAs is 2.3 times higher than that of silicon. Even though the substrate is a light material, such as

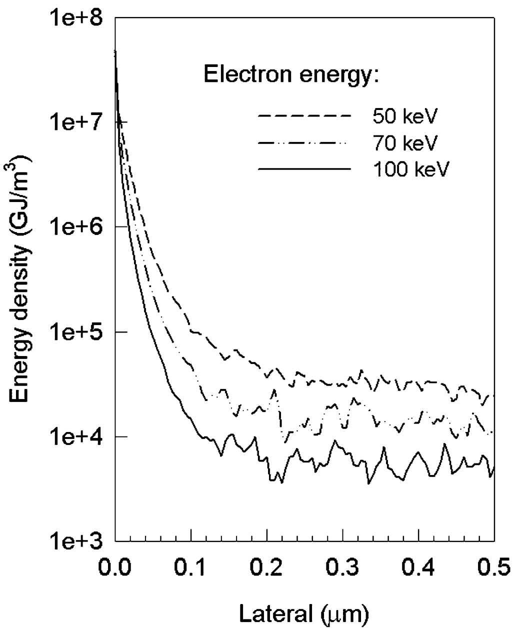


Fig. 3.9 Comparison of PSF at different electron energies (silicon substrate, 0.5-μm resist thickness)

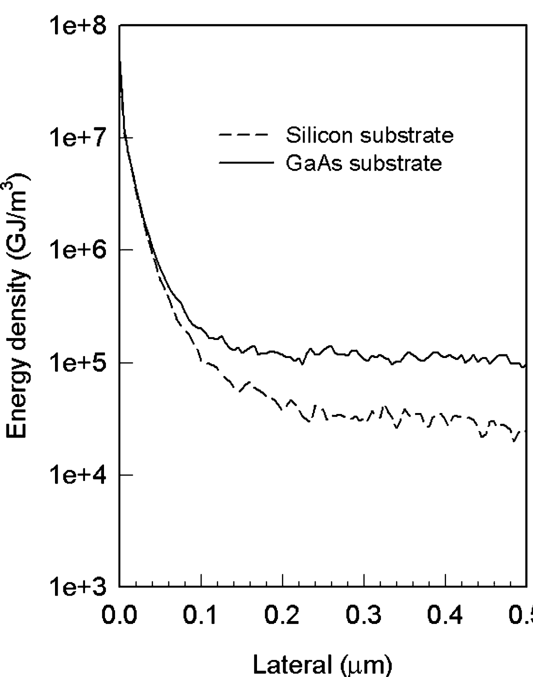


Fig. 3.10 Comparison of PSF for different substrate materials (50-keV beam energy, 0.5-μm resist thickness)

silicon, the backscattering can still be significant if there is a thin film of heavy metal, such as gold or tungsten, coated at the interface between silicon and resist layer. Therefore, it is more difficult to achieve high-resolution e-beam lithography on high-density substrate.

3. *Resist thickness*. This has been apparent from Figs. 3.7 and 3.8 that electrons scatter wider forwardly as they go deeper. Therefore, for high-resolution e-beam lithography thin resist layer is preferred.

3.3.2 Proximity Effect and Correction

Like the optical proximity effect which is caused by light diffraction, the direct consequence of electron scattering in e-beam lithography is the so-called electron proximity effect. Because the deposited energy from scattering electrons can extend to adjacent areas, dense features will become overexposed and very isolated features will be underexposed, creating the so-called intershape and intrashape proximity effect. An example of proximity effect from computer simulation is shown in Fig. 3.11. It shows the cross-section profile of PMMA

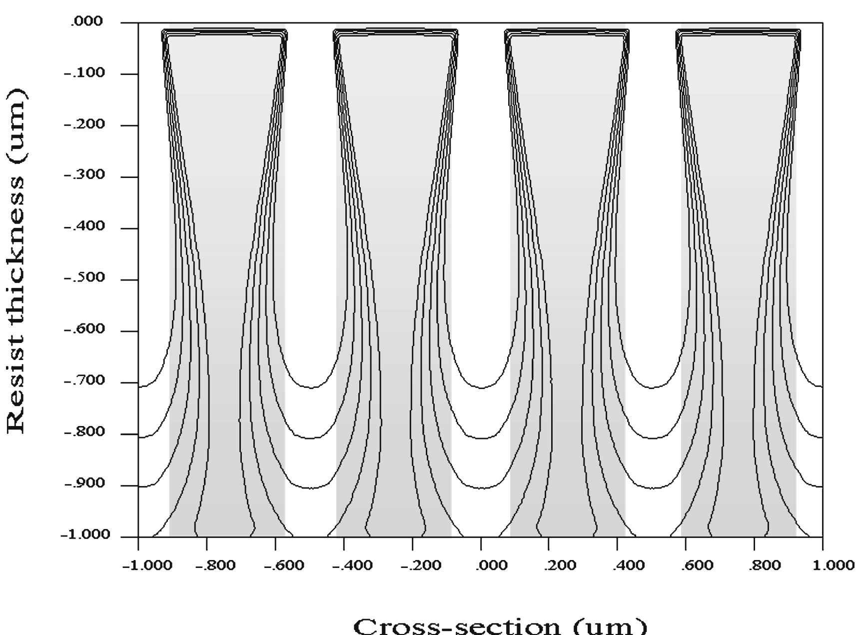
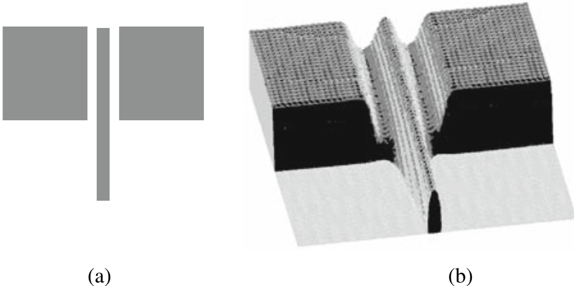


Fig. 3.11 Simulated resist cross-section profiles for 20-keV e-beam exposure of 1-μm PMMA resist layer on silicon substrate; the *shaded area* represents the ideal resist profile and the *contours* represent the actual resist profiles at exposure doses from 80 μC cm^{-2} to 120 μC cm^{-2} with 10 μC cm^{-2} increment

resist of 1-μm thickness, exposed by 20-keV e-beam. The pattern is 100-nm line grating with 500-nm pitch. The ideal profile is shown by the shaded area which is unexposed resist with 100-nm line openings. The actual resist profile is shown by the contours which represent the evolution of resist cross-section at different exposure doses. The higher the exposure dose, the deeper it goes until it reaches the full depth of 1 μm. It is apparent that the actual resist profile is significantly deviated from the ideal profile, because of the proximity effect. The result manifests the effect of electron forward scattering, which becomes wider as the electrons go deeper in the resist.

For high-resolution e-beam lithography, particularly when dealing with irregular pattern size and density, correction of proximity effect is absolutely necessary. Proximity correction has always been part of e-beam lithography development and can be dated back to 30 years ago [19]. Many research results have been published on this subject. In this section, the three principal correction methods are briefly introduced. They are dose correction method, pattern size compensation method, and background exposure compensation method.

Dose correction is the most widely used correction technique and also a technique giving the best correction results. It is already shown that electron scatterings create nonuniform distribution of electron energy in resist, which extends beyond the exposed pattern area and is the root cause of proximity effect. The principle of dose correction is to adjust the exposure dose at different parts of a pattern so that all patterns receive the same exposure energy; for example, Fig. 3.12a shows a group of patterns. They all receive the same exposure dose, represented by the same shading color. Because of the electron scattering and proximity effect, the line between the two big pads cannot be resolved, as the simulated three-dimensional (3D) resist profile shown in Fig. 3.12b. If the patterns are given different exposure doses, as shown in Fig. 3.13a where the light color shading represents lower exposure dose, the line in the middle can be fully resolved, as shown by the 3D resist profile in Fig. 3.13b. In this case, the big pads receive less exposure dose and the line receives different exposure doses at different sections depending on whether it is between the pads or outside the pads.



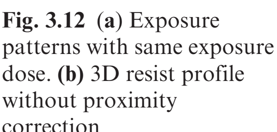


Fig. 3.12 (**a**) Exposure patterns with same exposure dose. **(b)** 3D resist profile without proximity correction

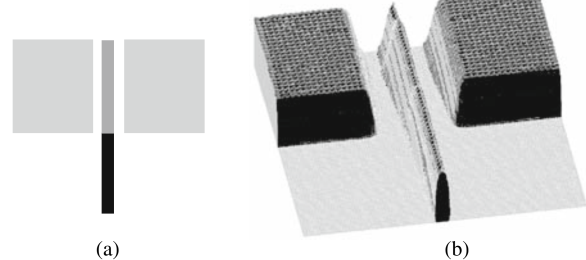


Fig. 3.13 (**a**) Exposure patterns with different exposure doses at different locations. (**b**) 3D resist profile with proximity correction

The above example is a very simple case, for which one can intuitively alter the exposure dose on different parts of the patterns. For much complicated pattern designs, proximity correction has to be done by intensive computation. A self-consistent method has been developed to find the correct dose allocation for a design pattern [19]. The method first assumes an initial exposure dose applying to all patterns. Energy distribution of all the patterns, including contributions from neighboring patterns due to electron scattering, is calculated. Naturally, the distribution will not be uniform. The next step is to adjust the exposure dose on various regions of the exposure area and various parts of each pattern based on the known energy deposition. Now the exposure doses are no longer uniform. Calculation of energy is carried out again according to the new exposure doses. The uniformity of energy distribution should be improved, but may not be good enough. Another round of dose adjustment and energy calculation may be needed. The cycle is repeated until the deviation of energy distribution between two calculations is within an error limit (tolerance). It is then considered that a self-consistent solution has been reached and the iteration can be terminated. The final result will be a map of exposure dose which have different values at different parts of exposure patterns. Theoretically, the self-consistent method requires calculating the electron scattering effect for all the patterns. If the pattern number is N, there will be N sets of equation to solve. The amount of computing is enormous and may not be necessary. As we know that the electron scattering range is limited around the point of incidence, the self-consistent method can be simplified by zoning the exposure area; that is, the calculation of scattering and energy deposition is only carried out in certain zones where electron scattering effect is significant [20].

The procedure of getting correct exposure doses by self-consistent solution is of direct physical meaning, hence called a physics-based correction method. Its solution is also relatively accurate. However, to find self-consistent solutions for an IC design where there are millions of pattern shapes, the simulation procedure becomes very complicated and the computation is very time-consuming. There is another simple geometry fraction method, which gives much faster proximity correction, though the correction result is relatively coarse. The difference of the two methods is that the physics-based correction fractures a pattern according to the self-consistent calculation of electron energy at every

point of the pattern. So the fractured patterns are irregular. Geometry fraction method first divides a pattern into regular sub-shapes according to the radius of electron scattering range, then calculates the electron energy deposition from which appropriate doses are assigned to each sub-shape [21]. It is a faster proximity correction method because there are no iterative computations. The electron energy in all patterns is only calculated once.

The above proximity correction requires complicated computer algorithms. There are a number of commercial proximity correction softwares available. Usually the proximity correction software comes as part of software tool sets when an e-beam lithography system is purchased. These commercial software tools can process any pattern designs with reasonably accurate correction results. However, in some cases the proximity correction can be done by resizing the pattern geometry. The resizing, also called biasing, is best suitable for simple and highly repeated patterns, such as lines and squares. For a group of repeated lines, the lines in the middle of pattern area can be resized narrower (negative biased) than nominal design to compensate the higher-than-average energy contributed from neighboring lines, or the lines at the outer rim of pattern area can be resized wider (positive biasing) to compensate the lower-than-average energy due to electron scattering. The overall result is a relatively uniform distribution of energy across the whole pattern area. The amount of biasing can be determined empirically or from simulation of electron energy distribution. An example is shown in Fig. 3.14. The figure shows deposited energy for a group of seven lines with linewidth of 150 nm and pitch of 450 nm and exposed by 20-keV e-beam into 0.5-μm thick resist on silicon. If all lines are exposed at equal width, the lines in the middle receive higher energy (solid line plot) because electron scatterings from both sides have contributed energy to the middle lines. If the middle lines are negatively biased by 10–25 nm, all the lines receive more or less the same energy (dash line plot).

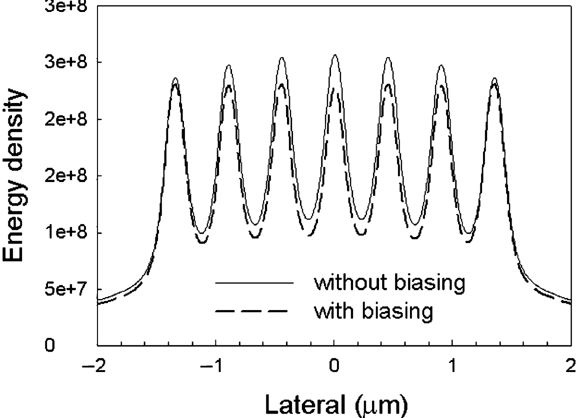


Fig. 3.14 Comparison of pattern biasing for proximity effect correction (150-nm lines and 450-nm pitch, exposed with 20-keV beam energy into 0.5-μm resist on silicon)

The third method for proximity correction uses background exposure to compensate the nonuniformity of electron energy distribution. As aforementioned, electron scattering always causes excess of energy at dense pattern region and less energy at sparse pattern region. If a second exposure is carried out at all non-pattern regions (background), the pattern density for the second exposure is exactly reversed. The originally sparse pattern region now becomes dense pattern region and receives more energy. Meanwhile, the original dense pattern region where there are less blank areas now receives less energy. The superimposition of the two exposures can well balance the energy distribution so that it becomes uniform everywhere. Figure 3.15a shows the energy distribution of a group of seven lines (150-nm width and 450-nm pitch). The first exposure produces an unbalanced energy profile, with the lines in the middle receiving higher energy. If a second background exposure, with energy profile shown by the dash line plot in Fig. 3.15a, is added to the first exposure, the overall energy distribution is well balanced, as the energy profile plot shown in Fig. 3.15b. The background

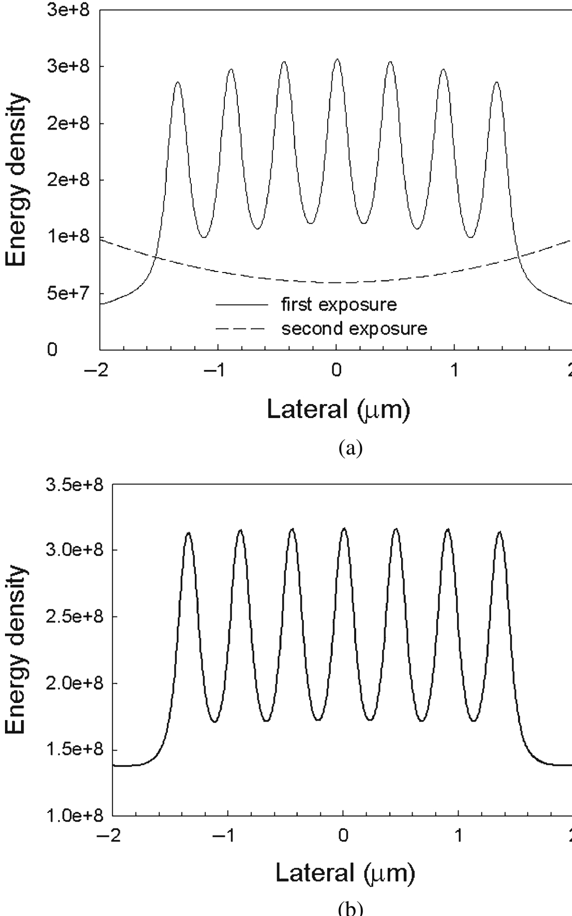


Fig. 3.15 (**a**) Energy profiles of first and second exposures. (**b**) Superimposed energy profile

exposure method has a name called “GHOST” [22]. It is simple and does not need special proximity correction software and does not need to calculate the electron energy distribution. All it needs is to have a second exposure of reverse polarity of the original pattern design. The second exposure can be carried out with a defocused e-beam and at relatively low exposure dose, with only the purpose of compensating the nonuniformity of energy. The drawbacks of this method are increased pattern data volume, increased total exposure time, and deterioration of image contrast due to the background exposure. This technique is only practical in a raster scan e-beam system. A raster scan system scans the whole exposure field all the time. The beam is only switched on where there is a pattern shape to be exposed. Therefore, the background exposure is just another scan with beam switched on in opposite polarity. For a vector scan system, the beam only scans over a pattern area. The background exposure requires another set of design data with all the background patterns, and the exposure time can be exceedingly long.

All the above proximity correction techniques require additional process and special software tools. A simple and effective measure to reduce the proximity effect is to increase beam energy as well as to decrease resist thickness. High beam energy can greatly reduce the beam expansion caused by forward scattering and reduce localized backscattering. As forward scattering is wider with electrons going deeper, a thin resist layer can also minimize the effect of beam expansion. An example is shown in Fig. 3.16, which is a direct comparison with Fig. 3.11. The pattern design and resist thickness are the same in both cases. The only difference is the beam energy for the exposure. While in Fig. 3.11 the resist is exposed at 20 keV, the beam energy in Fig. 3.16 is 100 keV. No proximity effect is seen for 100-keV exposure and resist profile is almost the same as the ideal profile (shaded area).

All the above discussion on electron scattering and proximity effect has favored using high-energy e-beam. High beam energy results in narrower forward scattering and lower backscattering. Although the higher the better, too-high voltage makes the design and construction of electron optical column very difficult. Besides, electron energy above $\sim$120 keV can damage silicon lattice and semiconductor devices. Therefore, it has been concluded that 100 keV is the best suitable energy for e-beam lithography [23]. However, one must be aware that high beam energy requires high exposure doses. The exposure dose increases linearly with the increase of beam energy, because majority of primary electrons have gone deep into substrate and very few of them contribute to the exposure of resist.

3.3.3 Effect of Secondary Electrons

So far the discussion about electron scattering and proximity effect has only been focused on primary electrons. Although the major contribution of proximity effect comes from scattering of primary electrons, secondary electrons

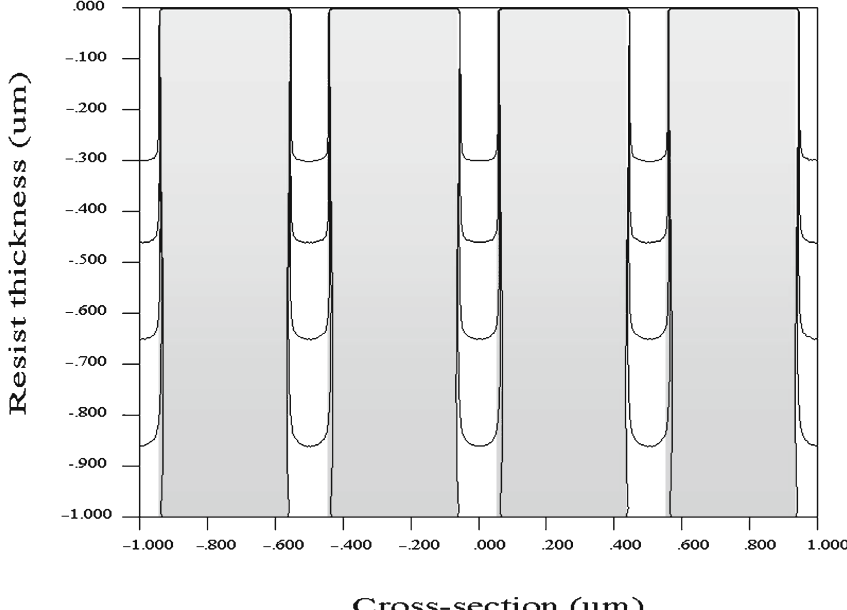


Fig. 3.16 Simulated resist cross-section profiles, to compare with those in Fig. 3.11, for 100-keV e-beam exposure of 1-μm PMMA resist layer on silicon substrate; the *shaded area* represents the ideal resist profile and the *contours* represent the actual resist profiles at exposure doses from 150 μC cm^{-2} to 350 μC cm^{-2} with 50 μC cm^{-2} increment

can play a vital role when e-beam lithography is conducted at close to the beam diameter, namely around 10 nm or below. The secondary electrons are electrons knocked off from the atoms of solid material by high-energy primary electrons. While traveling inside the resist, primary electrons transfer their energy to the electrons surrounding the atoms of materials. These secondary electrons acquiring up to kilovolts of energy become mobile and can travel a short distance. For example, electrons of 400-eV energy can travel about 12 nm and 200-eV electrons can travel about 5-nm distance [24]. They scatter the same way as primary electrons, only in a short distance, but always in directions perpendicular to the path of primary electrons, and leave their energy along the path of scattering. The consequence of the scatterings of these "fast" secondary electrons is a broadened energy deposition in resist. Figure 3.17 gives a comparison of PSFs with and without contributions from secondary electrons. It is apparent that secondary electrons have made a noticeable contribution to the low-energy part of PSF.

Detailed Monte Carlo simulations have concluded that on average the secondary electrons scatter about 10-nm distance around the main path of a primary electron, which limits the e-beam lithography resolution to no better

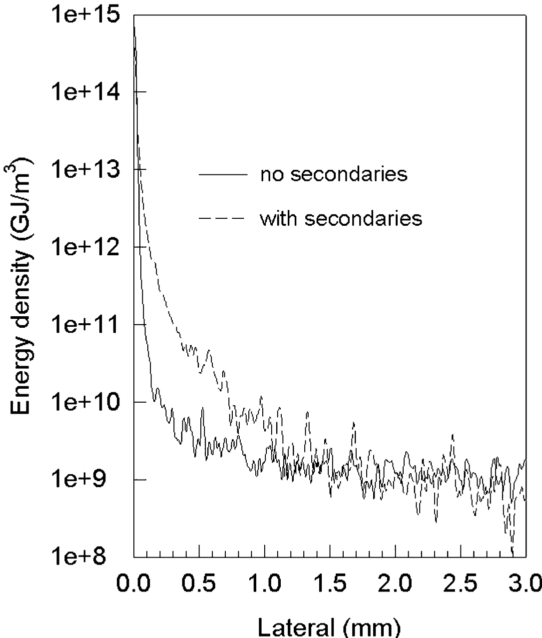


Fig. 3.17 Comparison of PSFs with and without secondary electron contribution

than 10 nm, no matter how small the e-beam spot size is [25]. However, this conclusion is not completely true, because e-beam lithography depends not only on electron energy deposition but also on resist material and process. The later can to some extent suppress the energy broadening effect caused by secondary electrons and enable lithography resolution to go below 10 nm. The e-beam resist process will be discussed in a later section of this chapter.

3.3.4 Low-Energy E-Beam Lithography

Although high beam energy has been favored in e-beam lithography, its rather inefficient use of energy, the need of high exposure dose and possible damage to substrate material or to buried structures, have prompted investigation on very low energy exposure of resists. At very low end of beam energy (0.5–5 keV), primary electrons cannot travel very far. Figure 3.18 compares the electron scattering trajectories at 10-keV energy (Fig. 3.18a) and 3-keV energy (Fig. 3.18b). The resist thickness is 100 nm on a silicon substrate. At 10 keV, most of the electrons have gone into substrate and there are backscattered electrons as well. The higher the beam energy the more electrons are ended up deep in substrate. Two problems can arise: one is substrate heating and another is substrate charging. For nonconductive substrate materials, such as quartz and glass which are commonly used for making photomasks, both heating and charging

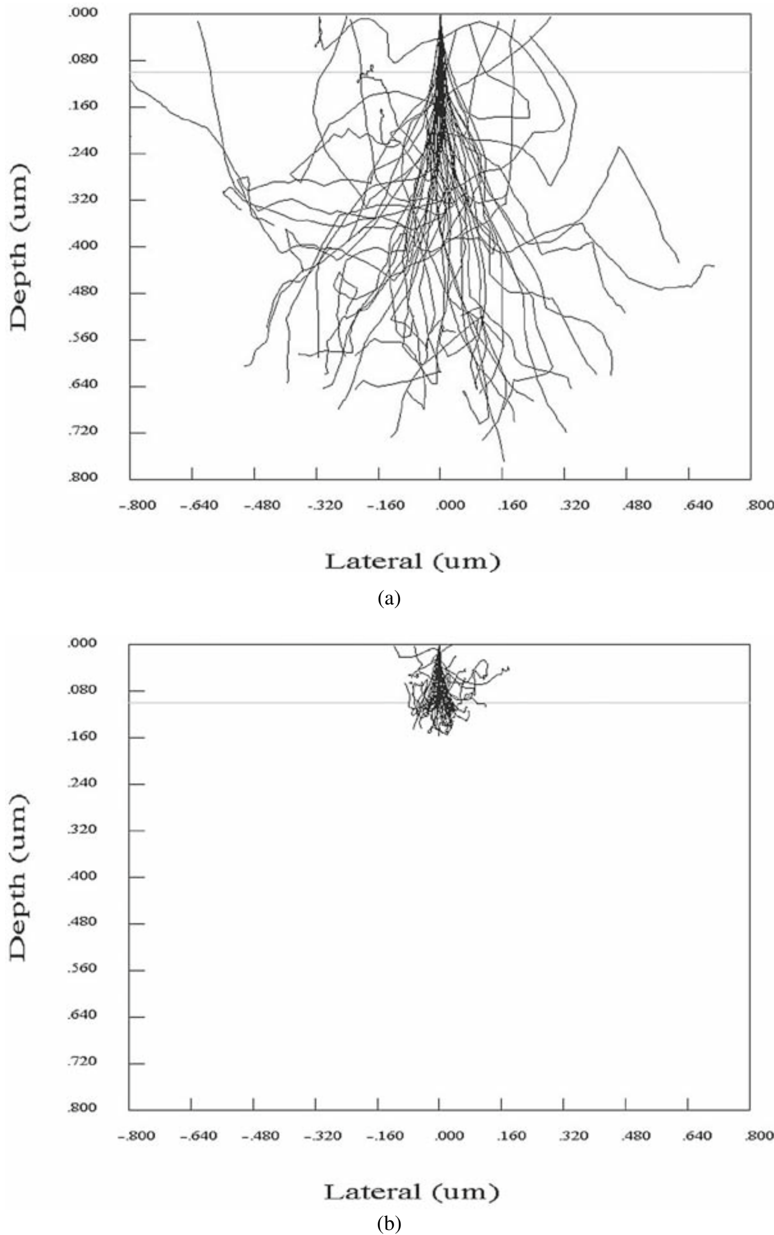


Fig. 3.18 Comparison of electron scattering at 10-keV energy (**a**) and 3-keV energy (**b**)

exist. Charging problem can be avoided by using a metallic underlayer between resist and nonconductive substrate [26]. Resist heating due to deposition of large amount of electron energy in a substrate of poor thermal conductivity can

cause change of resist sensitivity and severe proximity effect [27], particularly in e-beam lithography of photomasks [28], and there is no effective way to overcome the heating problem.

With low-energy electrons, there would be little energy deposited into substrate if a correct combination of resist thickness and beam energy is chosen. Figure 3.19 shows experimentally measured penetration depth in resist at different beam energies. The resist was ZEP-520 positive e-beam resist. A series of large square patterns were exposed by an e-beam with different energies. The resist was then developed and the remaining thickness was measured, from which the penetration depth was worked out [29]. Because the electron penetration is shallow at low energy, it brings two advantages: one is small range of forward scattering and little or no backscattering. If the resist thickness is just about the penetration depth of low-energy electrons, then there will be no backscattered electrons, hence no backscattering-related proximity effect. For example, sub-100 nm lines were produced in 66-nm thick PMMA layer at 2-keV beam energy and the deviation of linewidth between dense pattern region and sparse pattern region was found to be less than 10 nm [30]. Another advantage of shallow penetration of electrons at low-energy e-beam is the high concentration of energy deposition in resist layer so that fewer electrons are need for resist exposure. Resists become much more sensitive to low-energy electrons. Figure 3.20 is the measured resist sensitivity versus electron energy [29]. The resist (ZEP-520) was 180-nm thick and electron energy ranged from 3 keV to 20 keV (the electron penetration depth is over 180 nm at 3 keV energy as shown in Fig. 3.19). The resist sensitivity is defined as the exposure dose needed to enable the 180-nm thick resist layer to be fully developed. The higher the exposure does needed the lower the resist sensitivity. As seen from the plot, the decrease of resist sensitivity is linearly proportional to the increase of beam energy. In fact, low-energy electrons are much more efficient at exposing a resist polymer, because the majority of polymer chain scission or cross-linking happens during interaction with electrons of below 100-eV energy.

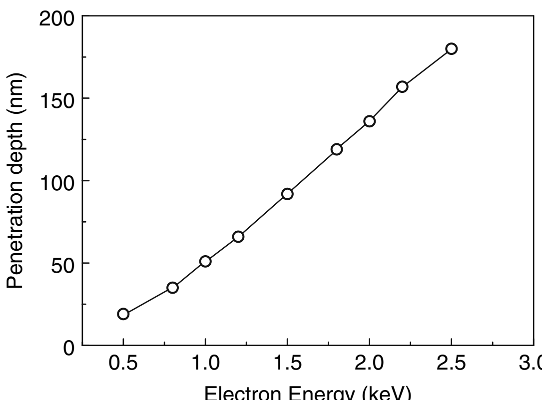


Fig. 3.19 Electron penetration depth at different energies

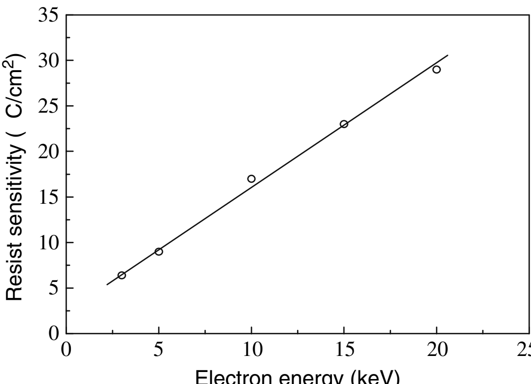


Fig. 3.20 Resist sensitivity vs. electron energy

For high-energy e-beam lithography, it is the low-energy secondary electrons, excited by high-energy primary electrons, which are responsible for the exposure of resist [24].

The shallow penetration of low-energy e-beam lithography is also a drawback, because only very thin resist can be used for lithography. The resist thickness should normally be below 100 nm, depending on the beam energy. As is known from Chapter 2, thin resist is less resistant to reactive-ion etching, making resist pattern transfer to substrate more difficult. In this case, multilayer resist structure has to be used, whereby the top thin layer resist acts as an imaging layer [31]. Another drawback of low-energy e-beam lithography is the difficulty to achieve fine size of beam. As has been described in previous section, both chromatic aberration and space charge effect become severe at low beam energy. The performance of a conventional e-beam lithography system falls rapidly below 6-keV acceleration voltage. The only way to maintain relative focusing quality of the electron optical column is to introduce a retarding field after the objective lens, to slow down high-energy e-beam to a low energy before landing on the substrate [32].

The real interest in low-energy e-beam lithography originated from the development of miniaturized e-beam lithography system, also called microcolumn. The concept of microcolumn was proposed in early 1990s [33]. It was envisioned that the whole electron column in a conventional e-beam lithography system which is meters high can be miniaturized to a few centimeters. An array of such microcolumns can perform e-beam lithography in parallel, so that high throughput can be achieved. The microcolumn array was originally proposed to replace optical lithography when it was run out of resolution capability. Although the dream of replacing optical lithography was never materialized, the development of microcolumn-based e-beam lithography system has never stopped [34]. For such a microcolumn, the voltage applied to the electron lenses cannot be higher than kilovolts. Therefore, low-energy e-beam lithography is inevitable. Although the beam spot size produced by a microcolumn is not as small as that produced

by a high-energy column, which is several tens of nanometers instead of a few nanometers by a high-voltage column, the low proximity effect at low-energy electron exposure may compensate its big beam spot and enable sub-100 nm resolution lithography. A more promising low-energy e-beam lithography technology was proposed in late 1990s, which was based on projection of low-energy electrons using a stencil mask [35]. The resolution of low-energy e-beam projection lithography (LEEPL) is no longer limited by the beam spot size but by the resolution of stencil mask. A feature dimension of 70 nm has been achieved by LEEPL at 2-keV electron energy on 50-nm thick ZEP-520 resist with no proximity effect [36].

3.3.5 Ion Scattering

As charged particles, high-energy ions also scatter in solid material. The difference between electron scattering and ion scattering is that ions are much heavier than electrons, and therefore, scatter very little. Like the modeling of electron scattering, ion scattering can also be modeled by Monte Carlo simulation, though the underlying physics is very different [37]. Figure 3.21 shows the trajectories of gallium ion scattering in PMMA at 50-keV energy [38]. Compared with 20-keV electron scattering shown in Fig. 3.7, the lateral range of 50-keV Ga^+ ion scattering is 10 times smaller and penetration depth is 30 times smaller. The PSF for ion scattering is also very different, as shown in Fig. 3.22. The PSF is very narrow. There is no long tail of backscattering energy, as the PSF for electron scattering shown in Fig. 3.8. The deposited peak energy is over three orders of magnitude higher than that of electrons. All these features imply that ion beam is much better suited for exposure of resist. Much lower proximity effect and much higher sensitivity in resist exposure can be expected. For example, a resist may need 20 μC cm^{-2} exposure

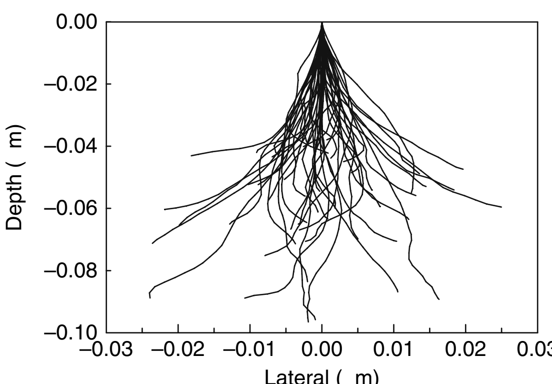
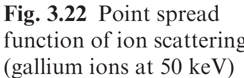


Fig. 3.21 Ion scattering trajectories in PMMA resist (gallium ions at 50 keV)



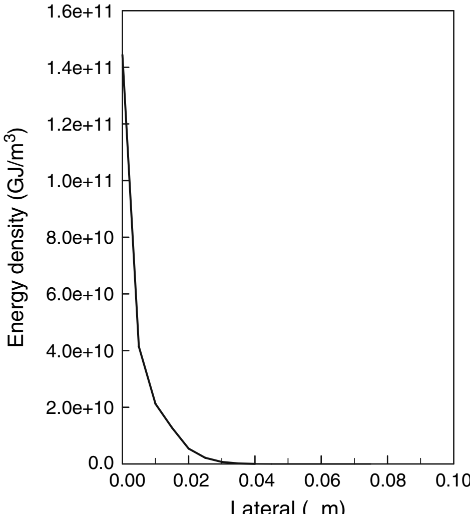


Fig. 3.22 Point spread function of ion scattering (gallium ions at 50 keV)

dose by 20-keV e-beam lithography. The same resist only needs 0.1 μC cm^{-2} exposure dose by gallium ion beam exposure at 100 keV [39], and 40-nm lines had been achieved in mid-1980s by ion proximity lithography through an edge-defined stencil mask [40].

The advantages of using ion beam to expose resist are the negligible proximity effect and extremely high sensitivity. The disadvantage is its very shallow exposure depth. For 50-keV $Ga+$ beam, the exposure depth in PMMA is only about 60 nm. Even at 100 keV, the depth is no more than 110 nm. Lighter ions, such as H^+ or He^+, can penetrate into resist over 500 nm. However, both H^+ and He^+ ion species can only be generated from gaseous ion sources which do not have the small source size as a LMIS, and therefore, impossible to be focused into a fine beam. The more practical approach is ion projection lithography (IPL) whereby a broad H^+ or He^+ beam is projected through a stencil mask for resist exposure. IPL will be introduced in details in a later section of this chapter.

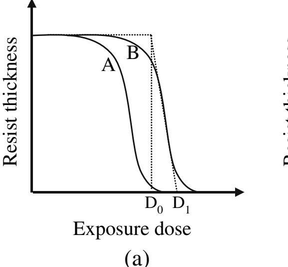
3.4 Resist Materials and Processes

The energy of charged particles deposited in a polymer resist will have to be converted into molecular chain reaction in the resist to eventually form a lithographic pattern. A small beam size and small scattering range are not necessarily translated into a small size of resist features. On the other hand, a poor resist material and non-optimum resist process can certainly result in poor resolution and much enhanced proximity effect. Resist materials and processes

play a vital role in achieving the ultimate resolution capability of lithography by charged particle beams.

3.4.1 Sensitivity of Resist Materials

The most important properties of a resist are sensitivity and contrast, which are directly linked to the resolution capability of the resist. The sensitivity and contrast are derived from a resist development curve which typically looks like the plots in Fig. 3.23. For a positive resist, long molecule chains are broken by energized electrons into short chain, called chain secession process, so that the exposed parts of polymer become soluble in developer. For a negative resist, the initial short chain molecules are joined upon exposure to form long chains, called cross-linking process, so that the exposure parts of polymer become insoluble in developer. The development curves, also called contrast curves, are obtained by exposure and development of a series of large square patterns with different exposure doses. The remaining resist thickness is measured against the exposure dose, which generates the curves shown in Fig. 3.23. For positive resist, there is an exposure dose at which the resist layer is fully developed to bottom. This exposure dose is defined as the resist sensitivity. The smaller this threshold dose, the higher the resist sensitivity. In Fig. 3.23a, resist A is more sensitive than resist B. For negative resist, the development curve looks like the plot C in Fig. 3.23b. With the increase of exposure dose, the degree of cross-linking increases and more resist thickness remains after development, until the whole resist layer is completely cross-linked. The sensitivity of negative resist is defined by the exposure dose which results in more than half the resist thickness remaining due to cross-linking. A practical example of development curves is shown in Fig. 3.24, where the depth of resist development at different exposure doses is plotted for a negative-tone resist, HSQ [41]. The development curves at different beam energies are also shown.



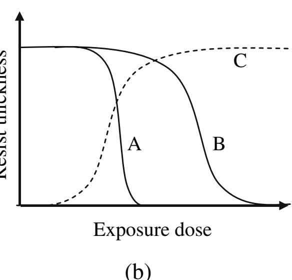
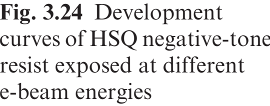


Fig. 3.23 Resist development curves: (**a**) resist A is of higher sensitivity than resist B, (**b**) resist A is of higher contrast than resist B, resist C is negative resist



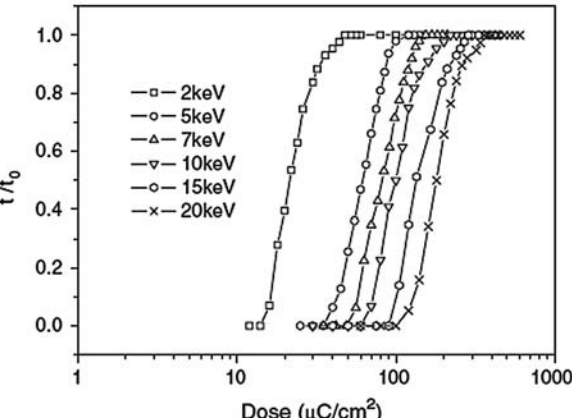


Fig. 3.24 Development curves of HSQ negative-tone resist exposed at different e-beam energies

One must be aware that the above definition of resist sensitivity is only for comparison purpose. It is defined by exposure of much larger pattern area, which is rarely representative in actual e-beam lithography. The actual exposure dose needed for a particular pattern exposure can be quite different. It depends on the following factors:

– *Electron energy*. Higher electron energy requires higher exposure dose, resulting in low resist sensitivity, as is shown in Fig. 3.24.
– *Resist thickness*. The thinner the resist, the lower the resist sensitivity, because fewer electrons deposit their energy within the thin resist layer.
– *Pattern density*. Exposure of dense patterns requires less exposure dose, because electron scattering from neighboring exposure areas contribute to the overall energy deposition. Similarly, isolated patterns require high exposure dose, because a portion of electrons have scattered outside the pattern area.
– *Substrate material*. High-density substrate materials result in more backscattering of electrons into resist layer, contributing to the resist exposure. Therefore, resist coated on a high-density material appears more sensitive than on a low-density material.
– *Process conditions*. The post-exposure baking conditions, in the case of chemically amplified resists, and the strength of developer and the temperature of developer can all have influence on the resist sensitivity, which will be discussed later.

Resist contrast is defined by the slope of a development curve and expressed in Eq. (3.12). $D_1$is the exposure dose at which the resist is fully exposed and D_0 is the exposure dose at which the resist starts to be exposed. They are marked in Fig. 3.23a. The steeper the slope is, the higher the resist contrast. For example, the resist A in Fig. 3.23b has higher contrast than resist B. Resist contrast is an inherent property of a resist material. However, development conditions can have some influence on it.

Table 3.2 Sensitivity and resolution of some e-beam resists

	Resist tone	Sensitivity* ($\mu C\ cm^{-2}$)	Resolution(nm)
PMMA	+	100	10
ZEP-520	+	30	10
ma-N 2400	–	60	80
EBR-9	+	10	200
PBS	+	1	250
COP	–	0.3	1000

*Sensitivity measured at 20-keV beam energy.

$$\gamma = \frac{1}{\log_{10} D_1 - \log_{10} D_0}. \tag{3.12}$$

For high-resolution e-beam lithography, low-sensitivity resists are preferred. Table 3.2 gives a list of some commonly used e-beam resists with their sensitivities and resolution capabilities. Polymethylmethacrylate (PMMA), the resist discovered 40 years ago, is still being used today although it is of very low sensitivity. Polymethylmethacrylate is by far the only positive e-beam resist which has the highest resolution capability. The smallest features size produced by e-beam lithography in PMMA is 4–5 nm [42, 43]. It must be noted that PMMA resist comes with different molecular weights. Typical commercial PMMA resists have 950-K and 495-K molecular weights [44]. High molecular weight is generally of low sensitivity and high-resolution capability.

If lithography throughput is not a critical issue, particularly in R&D where resolution is paramount, then there is a group of inorganic resist materials which is of very low sensitivity but very high resolution capability. Use of inorganic materials as resist in e-beam lithography dated back more than 30 years ago, when 8-nm features were made for the first time by e-beam exposure of metal thin film to form a carbon contamination pattern which acted as an ion sputter etching mask [45]. The exposure dose required to generate the carbon contamination film was, however, as high as 10^7 $\mu C\ cm^{-2}$. A number of other inorganic materials have been tried for e-beam lithography over the years. For example, it was demonstrated that ultra fine e-beam could drill holes directly in a range of inorganic thin films, such as NaCl, MgF_2, Al_2O_3, LiF, and AlF_3, with hole diameter of about 1.5 nm, using a scanning tunneling electron microscope at beam diameter of 0.5 nm [46]. However, the exposure dose needed for the drilling was over 10^5 $\mu C\ cm^{-2}$, and these films could not serve as etching mask material, and therefore, rendered them useless in nanofabrication. There were also attempts to mix organic resists with inorganic materials to create quasi-inorganic resists, for example, fullerene molecules (C60) mixed with ZEP-520 [47], PMMA [48], or NEB-22 [49], and silica nanoparticles mixed with ZEP-520 [50], which have improved the resolution capability and etch resistance of the resists, but at the cost of reduced resist sensitivity. The latest member in the inorganic resist group is the HSQ resist [51]. Hydrogen silsesquioxane (HSQ) is a liquid

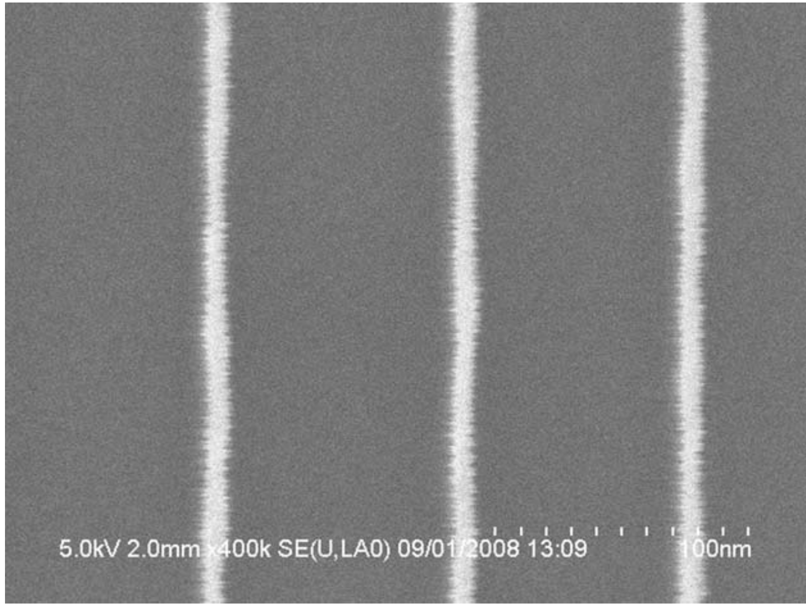


Fig. 3.25 Sub-10 nm lines exposed in HSQ resist (Reprint courtesy of NanoBeam Ltd. [52])

form of silicon dioxide from Dow Corning. The product name is flowable oxide and product code is Fox-12. Although HSQ is not a specially designed e-beam resist, it is sensitive to e-beam exposure and works as a negative-tone resist, with sensitivity and contrast similar to PMMA. Sub-10 nm resist line features can be achieved routinely with HSQ with a high-grade e-beam lithography system. Figure 3.25 shows a group of sub-10 nm HSQ lines obtained by exposure at 80-keV and 5-nA beam current on 58-nm thick HSQ resist [52]. The ruggedness of lines is not the line edge roughness (LER) but due to the noise of SEM when the photo was taken. Isolated line features 5–6 nm in sizehave also been produced in 30-nm thick HSQ by e-beam lithography at 100 keV [3], or sub-10 nm dense line features in ultra thin (10 nm) HSQ layer [53, 54]. Although PMMA has similar or slightly better resolution capability than HSQ, the LER in PMMA at sub-10 nm dimension is very high [42], while the LER in HSQ at sub-10 nm resolution is much lower. The etch resistance of HSQ is much higher than that of PMMA. In these aspects, HSQ is a better e-beam resist for extremely high-resolution e-beam lithography.

It is not difficult to understand why low-sensitivity resists have high resolution capability. As it is known from previous section, electron scattering spreads the energy distribution to a range larger than the exposure feature dimension. The long tail of PSF, as shown in Fig. 3.17, is in low energy level, which will not have significant exposure effect if the resist is low in sensitivity. In other words, these low-energy electrons, including secondary electrons, will

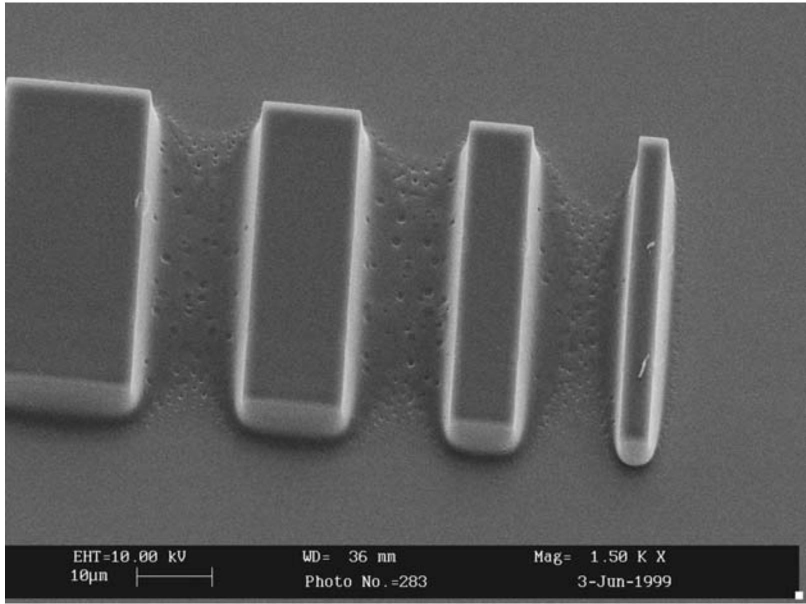


Fig. 3.26 SU-8 resist patterns exposed by e-beam at 50-keV energy

only show up their impact on resist pattern profile if the resist sensitivity is very high. An example is shown in Fig. 3.26, which is the SEM image of exposed SU-8 resist [55]. SU-8 is a photoresist but is also sensitive to electrons. The sensitivity of SU-8 is around 1 μC cm^{-2} at 50-keV beam energy, which is extremely high compared with PMMA. The consequence of such a high sensitivity is the residues of SU-8 resist shown between the exposed resist patterns. The residues are caused by exposure of backscattering electrons. They only exist between the exposure patterns because of the accumulated effect from both sides of electron scattering. If the resist had not been so sensitive, the residues would not have occurred.

Normally a pattern is exposed by e-beam lithography through multi-pass scan to fill up the exposure area, because the beam diameter and scan grid is much smaller than the pattern dimension. For very high resolution lithography of line features ($<$50 nm), a single-pass exposure is often employed, whereby the e-beam scans only a single pass. The line width is then determined by the exposure dose of this single pass. Although a single-pass exposure should theoretically be only the width of beam diameter, the actual line in resist can be much wider if exposed at high dose.

Because the accuracy of resist pattern is so much dependent on exposure dose, it is a common practice that e-beam lithography of any new pattern designs is always started with a dose test, that is, running a series of exposure for the same pattern design at different exposure doses, from which the optimum exposure dose range is determined.

3.4.2 Contrast of Resist Materials

In addition to sensitivity, the performance of a resist is also defined by its contrast. The resist contrast can be quantified by the slope of development curves shown in Fig. 3.23. For example, the contrast of negative-tone resist HSQ is 2.8 [41], as measured from Fig. 3.24. Another negative-tone resist, SU-8, has a contrast of only 0.92, as measured from contrast curve shown in Fig. 3.27 [55, 56] and reported in other experiments. By the appearance of contrast curves of these two resists, one can immediately tell that the HSQ resist has higher contrast. Apart from measuring the slope of resist contrast curves, another direct way of judging the resist contrast is to observe the resist profile after development. An example to compare the effect of resist contrast on resist profile is shown in Fig. 3.28, where computer-simulated resist profiles are shown for 50-keV e-beam exposure of 0.5-μm thick resist. Figure 3.28a is the profile of low-contrast resist while Fig. 3.28b is the profile of high-contrast resist. Although electron scattering and proximity effect are the same in both cases, high-contrast resist can produce the most desirable profile.

Resist contrast has profound effect on lithography resolution and resist pattern usability. High contrast is always desirable because high contrast enables resist profiles after development to have vertical sidewalls and high aspect ratios (feature height to feature width), which brings two benefits: facilitating pattern transfer by reactive-ion etching and enabling lithography of high-density patterns. Whether it is optical lithography or e-beam lithography, high resolution is easier to achieve for sparse or isolated patterns than for dense patterns with small half pitch. One should be aware that whenever talking about the ultimate resolution of lithography, one has to ask the question: at what pattern density? This has been the reason that IC manufacturing industry defines the resolution in half-pitch, not in minimum feature size. In e-beam

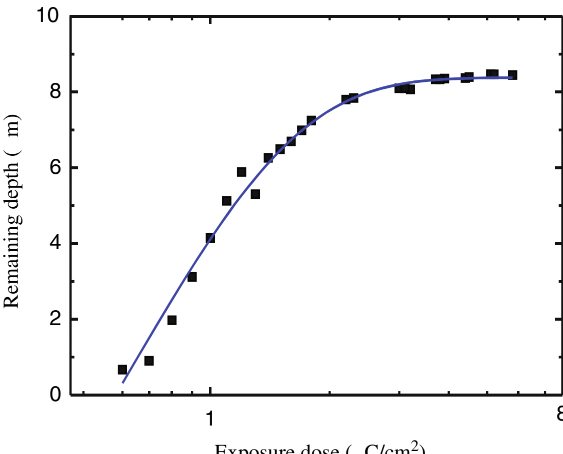


Fig. 3.27 Contrast curve of SU-8 resist

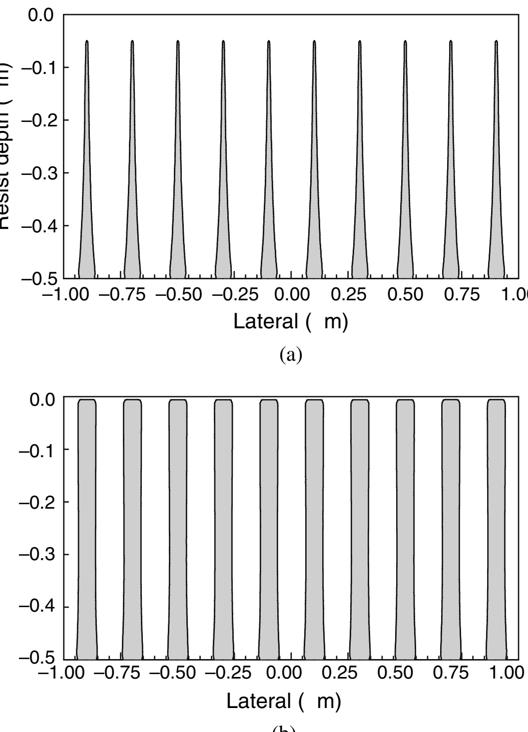
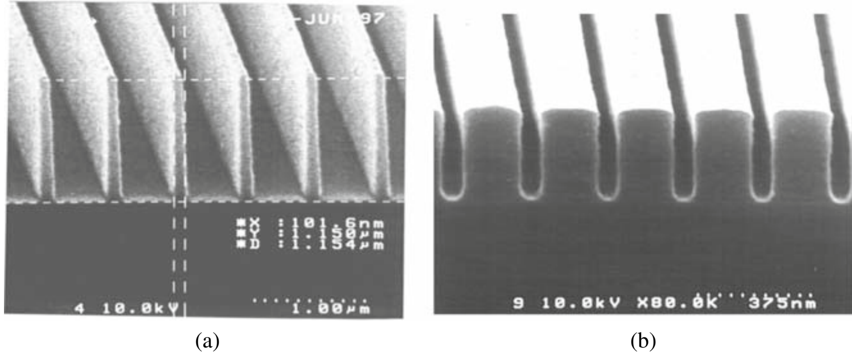


Fig. 3.28 (**a**) Low-contrast resist profile (**b**) High-contrast resist profile

lithography, such a limit in exposure of high-density patterns is caused by electron scattering – in other words, the proximity effect. Proximity effect correction does not work effectively for high-density periodic features. In this case, high-contrast resists can partly offset the proximity effect, making high-density patterning possible. The other route for high-density patterning is to use high beam energy or thin resist layer.

As previously mentioned, low-sensitivity resists are preferred for high-resolution e-beam lithography. However, there is an exception that some chemically amplified (CA) resists are of both high sensitivity and high resolution capability. The reason they are of high resolution capability is because of their high contrast. Figure 3.29 shows high-aspect-ratio patterns in negative-tone CA resist AZPN114 [57] and positive-tone CA resist UV3 [58]. Similar results have been demonstrated for other CA resists, such as SAL601 [57] and NEB-22 [59]. These CA resists are generally capable of resolution below 50 nm. Although their resolution capability is not as high as PMMA or HSQ, their very high sensitivity, two to three orders of magnitude higher, makes them very attractive in high-throughput e-beam lithography. However, one must be aware that resists of too high sensitivity may not be suitable for some slow scanning e-beam systems. If the scanning speed of an e-beam system is slow, for example below 10 MHz, the beam current has to be extremely small



(a) (b)

Fig. 3.29 (**a**) AZPN114 resist profiles (100-nm lines with aspect ratio 11:1). (**b**) UV3 resist profiles (40-nm lines with 8:1 aspect ratio)

in order not to overexpose some very high sensitivity CA resists. Extremely small beam current makes the focus calibration and registration extremely difficult because of low secondary electron image. Some systems may have a lower limit on attainable beam current, which renders the exposure of CA resists by these systems impossible.

3.4.3 Resolution Enhancement Processes

The resist sensitivity and contrast are partly dependent on the resist material property, also partly dependent on the resist process conditions. Although every commercial resist comes with a guideline on process conditions given by resist vendors, the conditions are meant only for general applications and may not be tuned to achieve the highest resolution capability. The process conditions are also tied to pattern density and exposure condition, such as the beam energy and exposure dose. Therefore, to achieve the highest resolution for a particular resist the optimum process conditions should be investigated. In some cases, exotic methods were developed to enable very high resolution of e-beam lithography which has previously been unthinkable. These exotic methods include ultrasonic-assisted resist development and development at lower or higher temperature.

The enhancement effect by ultrasonic-assisted development on e-beam lithography resolution with PMMA resist was discovered in 1993 [60]. Previously it had been generally accepted that e-beam lithography has a resolution limit at around 10 nm because of secondary electron scattering effect [46]. However, one fact was overlooked in considering the resolution limit, that is, the behavior of polymer molecules during the development process. The development is realized by short-chain polymer molecules (after chain scission action by electron exposure) dissolved in a developer solution. When the exposure region becomes very small (<10 nm), the molecules, which for PMMA have a size of around 2 nm, can be trapped in solid PMMA due to intermolecular forces which

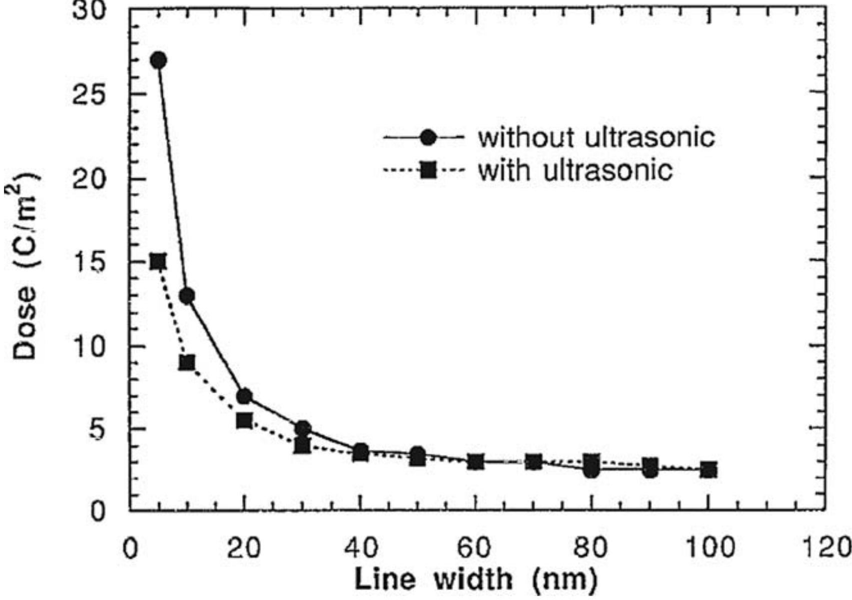


Fig. 3.30 Comparison of exposure dose required for PMMA development with and without ultrasonic assistance (Reprint from [60] with permission)

increase rapidly within the radius of 10 nm. More electron interaction with the polymer is needed to release the trapped polymer molecules, which means higher exposure dose. The exposure region then becomes wider at the increased exposure dose. With ultrasonic-assisted development, the ultrasonic agitation can provide additional energy and help to loosen up the trapped molecules to dissolve into the developer solution. In this way, exposure lines as small as 4 nm in PMMA has been achieved [42]. The ultrasonic agitation can indeed reduce the exposure dose significantly at line feature width below 20 nm, as shown in Fig. 3.30. It becomes no different to conventional development process for wider exposure region because the intermolecular force is no longer dominant at above 10-nm line opening. There were also other explanations as to why the ultrasonic agitation can help enhance the resolution of PMMA, such as reduction of developer viscosity or reduction of resist swelling [61]. However, it must be pointed out that the resolution enhancement by ultrasonic agitation only works for PMMA resist with high molecular weight, which demonstrates that resist development process is very much dependent on individual resist materials.

Another method to enhance resolution is to develop the exposed PMMA in a cold developer. Sub-10 nm lines were achieved at 30-keV beam energy with MIBK + IPA(1:3) developer which was maintained at 4–8°C temperature [43]. Unlike ultrasonic agitation, which promotes the activity of PMMA molecules, cold developer actually suppresses the activity of PMMA molecules. In order to

develop the exposed PMMA resist at low temperature, exposure dose much higher than development in room temperature is needed. High exposure dose is known to result in wider exposure region, as indicated by previous electron scattering analysis. However, energy level deposited in resist is rapidly reduced at the location away from the center of exposure region, as shown by the PSF for very thin resist layer in Fig. 3.8. In a cold developer, those PMMA molecules exposed by low-level electron energy cannot be dissolved because of their low level of kinetic activity at low temperature. Only the very central region where the deposited electron energy is the highest can be developed, resulting in high-resolution exposure features. The similar resolution enhancement at low-temperature development was also found in e-beam lithography of ZEP-520 resist. With the developer temperature as low as –17°C, 13-nm lines in ZEP-520 exposed at 30-keV beam energy have been obtained [62]. A particular benefit of low-temperature development for ZEP-520 is the much improved LER, as demonstrated in Fig. 3.31. The ZEP-520 resist lines in Fig. 3.31a were developed at room temperature, while the lines in Fig. 3.31b were developed at –4°C. It clearly shows the ZEP-520 resist molecules have much lower kinetics in low-temperature developer solution.

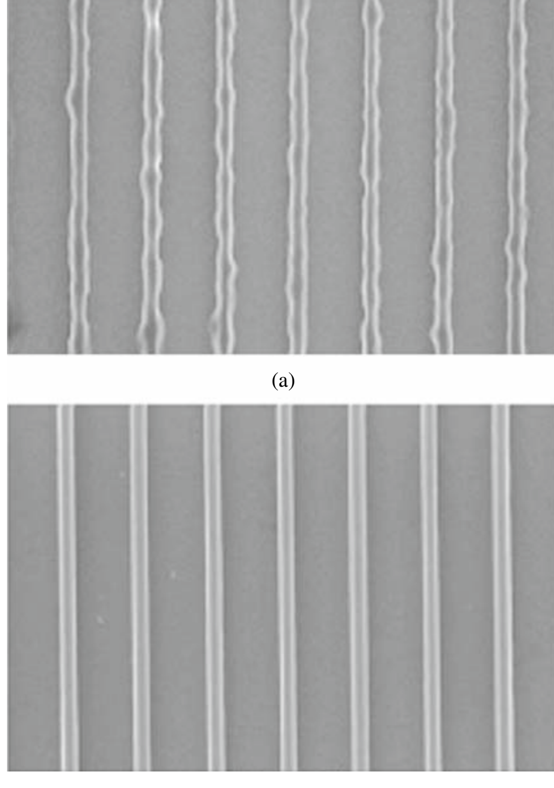


Fig. 3.31 Comparison of edge roughness of ZEP-520 resist lines (40-nm wide) developed at (**a**) room temperature and (**b**) at –4°C temperature (Reprint from [62] with permission)

The reason that cold development works for both PMMA and ZEP-520 resists is believed to be due to the variation of molecule weight distribution before and after electron exposure. Upon electron exposure, chain scissions occur, resulting in short chains and low molecular weight molecules in the exposed region. Low-temperature developer reduces the kinetic energy of molecules. Only those molecules with very short chain can be dissolved by the developer, which effectively suppressed the influence of scattering electrons because the polymer molecules exposed by low-energy scattering electrons cannot participate in the dissolving process. The same mechanism can explain why low-strength developer can help to achieve high-resolution resist patterns. Table 3.3 shows qualitatively the relation between developer concentration and resolution. On the other hand, positive chemically amplified resists such as UV3 do not show any improvement in low-temperature developers because there is no molecular weight degradation process during e-beam exposure of these resists [62]. However, cold development does show better results for some negative chemically amplified resists such as NEB-22 [63] and SAL-601 [64].

Table 3.3 Influence of developer concentration on resist resolution

Developer concentration MIBK:IPA)	Sensitivity	Resolution
1:3	Low	Extremely high
1:2	Medium	Very high
1:1	High	High
Pure MIBK	Very high	Low

While low-temperature development helps to enhance resolution of PMMA and ZEP-520, it works the opposite for HSQ resist. The contrast of HSQ resist is actually enhanced at high development temperature, as shown in Fig. 3.32, where development temperature varies from 18°C to 50°C, including ultrasonic-assisted development [65]. The sensitivity of HSQ is decreasing but the contrast is increasing with the increase of development temperature. It is not clear exactly why the high-temperature development has caused the decrease of e-beam exposure sensitivity of HSQ, as most other resists become more sensitive at high development temperature. However, the benefit of high-temperature development of HSQ is apparent. High-resolution dense patterns can be resolved when developed at high temperature.

3.5 Ion Sputtering and Redeposition

It has been discussed previously that FIB is not suitable for exposure of resists. Focused ion beam is almost exclusively used for direct nanofabrication, including sputtering to remove material or deposition to add material at sub-100 nm resolution. Sputtering is a process in which incident ions transfer their energy to atoms in solid material and these atoms, upon acquiring enough kinetic energy,

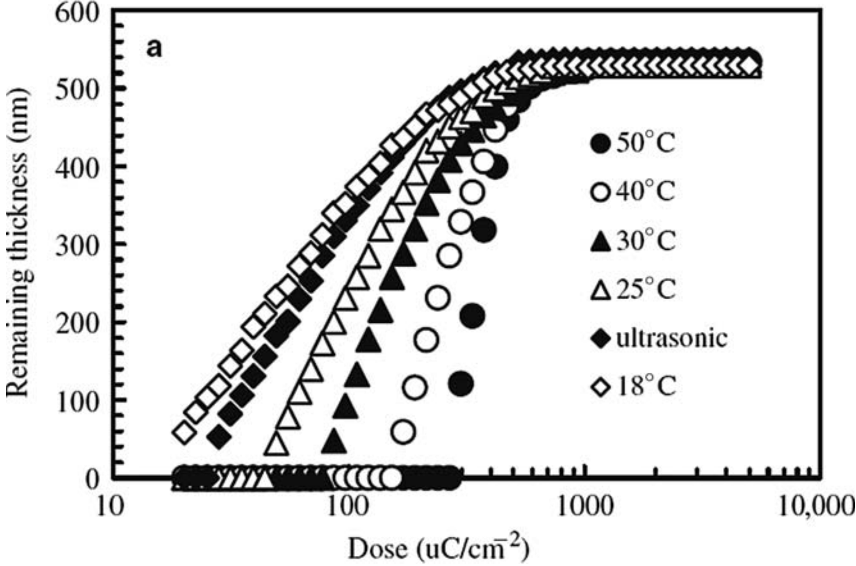


Fig. 3.32 Contrast curves of HSQ resist at different development temperatures

escape from solid material surface. Ion sputtering is not a one-to-one process. An ion beam bombarding a target material can generate a large number of recoiling atoms. These recoiling atoms can further pass their kinetic energy to surrounding atoms, generating more recoiling atoms. Some of the recoiling atoms at the vicinity of target surface may have enough kinetic energy to break off from the binding of surface energy into free space, becoming sputtering atoms. The number of atoms sputtered off by an incident ion is characterized by sputtering yield. Sputtering yield is a function of ion energy, incident angle, and density of target material [66].

Although a FIB system can produce a fine ion beam with beam diameter less than 10 nm, it is not always possible to achieve sub-10 nm structures by ion sputtering. The key limiting factor is the redeposition of sputtered material. Ion sputtering is a physical process. The sputtered material atoms have to land somewhere. In an open space, they can spread in a wide area and do not create significant redeposited structures. In a confined space, redeposition of sputtering material can be a major problem. One method to reduce the redeposition is to reduce the dwelling time of an ion beam at each sputtering point, for example, to sputter with fast and multiple scan of ion beam. At a fixed ion current, fast and multiple scan causes much less redeposition of sputtering atoms than slow and single scan, because multiple scan can sputter off those deposited atoms from previous sputtering. Figure 3.33 gives measured ion sputtering rate of niobium film at different ion beam dwelling time [67]. The two curves represent ion beam spots scanning with 50% overlap and without overlap. The curves indicate that the sputtering rate is the highest at very short

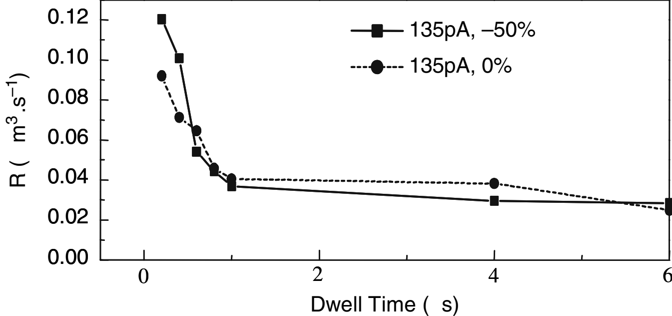


Fig. 3.33 Ion sputtering rate of niobium film at different ion beam dwelling time

beam dwelling time. It drops sharply and becomes constant with further increase of beam dwelling time, as the ion sputtering removal and redeposition become balanced.

The most effective technique to reduce the redeposition is to incorporate chemical gases into the sputtering process. Figure 3.34 compares 50-nm interdigital finger structures patterned by ion sputtering etching without gas assistance and with gas assistance. Because of the tightly confined spaces between the fingers, they are barely resolvable by ion sputtering alone (Fig. 3.34a). With the assistance of XeF_2 gas, the 50-nm fingers are clearly separated (Fig. 3.34b).

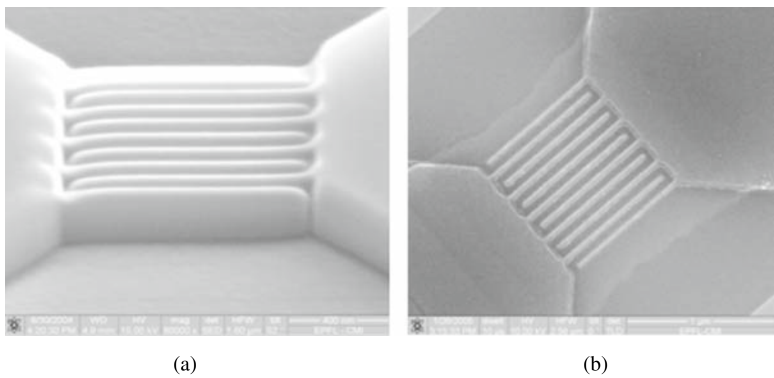


Fig. 3.34 Ion sputtering patterning of 50-nm interdigital fingers without XeF_2 gas assistance (**a**) and with XeF_2 gas assistance (**b**) (Reprint from [68] with permission)

Chemical gas-assisted sputtering process is essentially a combination of ion sputtering and chemical etching processes, because ionized chemically active gases such as chlorine and fluorine can violently react with silicon, gallium

arsenide, or aluminum to produce volatile compounds. The process works as follows: a small amount of chemically active gas is introduced into the sputtering chamber; gas molecules then absorb to target surface; ion bombardments ionize the absorbed gas atoms; these gaseous ions react with target surface material and convert the target material into volatile compounds which are then pumped away by the FIB vacuum system. Gas enhancement to ion sputtering etching can be characterized by enhancement coefficient E, which is expressed as Eq. (3.13):

$$E = 1 + (E_0 - 1)\theta \text{ with } \theta = \frac{1}{1 + \frac{\phi J}{F}}, \tag{3.13}$$

where θ is the steady-state surface coverage of reactive gas ($\theta < 1$), J is ion current density, ϕ is a constant for a particular gas–substrate combination, and F is the reactive gas flux density over the surface. E_0 is the maximum enhancement factor (when $\theta = 1$). It is apparent that the surface coverage of gas molecules is directly proportional to the gas flux but inversely proportional to the ion current density. There exists a balance of supply and consumption of reactive gas molecules at which the gas-enhanced ion sputtering etching reaches maximum. The chemical gases used in the assisted sputtering have to match the target material. Certain gases only react with certain target materials. Table 3.4 has listed some of the combination of chemical gases and target materials and their enhancement factors [69]. This property can be utilized to selectively etch a target if it is composed of different materials.

Table 3.4 Enhancement factors of gas-assisted ion sputtering etching

Gases	Al	S_i	S_iO_2	W
Br_2	8–16	5–6	0	0
Cl_2	7–10	0	0	0
XeF_2	0	7–12	7–10	7–10

3.6 Charged Particles Projection Lithography

Although charged particles, whether electrons or ions, can be focused into a fine beam, and therefore, can do very high resolution lithography, the fundamental limit is the low throughput because of its sequential nature in exposure. While optical lithography has standard throughput of at least 60 wafers per hour, a single 50-nm e-beam has to have a scanning speed of 20 km s^{-1} and beam dw elling time of 2.5 ps in order to compete with this throughput [70]. None of single-e-beam lithography systems can ever dream of such a high scanning speed and none of e-beam resists have such high sensitivity. The only solution is to increase the pixel number per single shot of exposure, that is, large beam size or multiple beams in a single shot.

The earliest attempt to increase the throughput of e-beam lithography was the development of shaped beam system. IBM pioneered the development during 1970s [71]. The so-called shaped beam is to use various shaped apertures to define the e-beam so that the actual pixel exposed onto a wafer is no longer a single Gaussian beam of very fine size but a large beam to fill a particular shape of exposure pattern. Shaped e-beam lithography played a key role in quick turnaround manufacturing of ICs in the early days when bipolar transistor-based IC was dominant, because every design needed a customization and e-beam direct write provided a quick turnaround without the need of making photomasks [72]. The shift from bipolar to complementary metal-oxide semiconductor (CMOS)-based ICs in the early 1990s has gradually phased out the shaped e-beam direct write in IC manufacturing. With the continuous decrease of minimum circuit feature dimension and increase of wafer size, direct write by shaped e-beam no longer met both the resolution and the high-throughput requirements. For example, to expose a 300-mm wafer at 65-nm critical feature dimension, it takes 50–60 h for patterning an entire wafer with a Leica-shaped beam system [73]. Massively parallelism in charged particle exposure was needed. Along this direction came the projection lithography either through a mask or in multibeams.

Several types of e-beam projection lithography (EPL) systems were developed over the last 15 years. There was the cell projection system from Hitachi (HL-800D) [74], the projection reduction exposure with variable axis immersion lens (PREVAIL) system from IBM [75], and the scattering with angular limitation projection e-beam lithography (SCALPEL) system from Lucent Technologies [76]. All of these e-beam lithography technologies rely on exposure through a mask. Typical mask types for e-beam projection lithography are either scattering stencil mask as shown in Fig. 3.35a, or scattering membrane mask as shown in Fig. 3.35b. Scattering stencil mask is mainly used in cell projection and PREVAIL systems and scattering membrane mask is used in SCALPEL system. The stencil masks are typically made of silicon of a few micrometers thick. The membrane masks are typically 100–200-nm thick so that high-energy electrons can easily penetrate through the mask, with 50–100-nm thick absorber made of high-density materials. They are called “scattering” mask because the “opaque” and “clear” regions of these masks rely on large and small scattering angles of electrons passing through the mask. It has been described previously that electrons scatter in a solid material and the scattering range depends on the density of substrate material. Electrons will have larger scattering angles when passing through denser region of the mask, such as the silicon layer in the case of stencil mask (Fig. 3.35a) or the absorber layer in the case of membrane mask (Fig. 3.35b). On the other hand, electrons passing through the open region (stencil mask) or thin membrane region (membrane mask) will have smaller scattering angles. If there is an aperture underneath the mask, which only allows electrons with small scattering angles to pass through, then all electrons with large scattering angles are blocked out, which creates an equivalent “opaque” image on resist for the mask region where electrons undergo large scattering range.

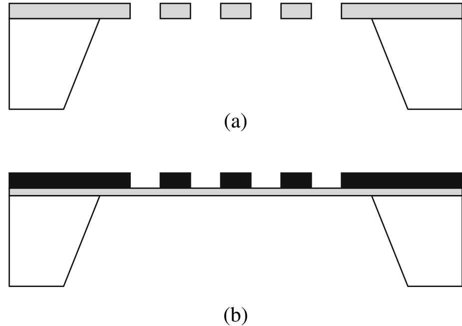


Fig. 3.35 (**a**) Scattering stencil mask. (**b**) Scattering membrane mask

Typical comparison of shaped beam-based direct-write lithography and mask-based projection lithography is given in Table 3.5 [77]. The comparison shows that projection through a mask can increase the throughput several folds. The latest Nikon NSR-EB1A e-beam stepper has achieved resolution capability of 70-nm dense lines and 50-nm isolated lines [78]. However, there are also fundamental and engineering issues that pose great challenge to further developing EPL into a mass-manufacturing tool.

The key engineering issue is the mask. The mask, whether stencil type or membrane type, is difficult to make, to inspect, and to repair. This has been an age-old problem when X-ray lithography was in its peak development era. A full 200-mm diameter silicon stencil mask with only 2-μm thickness is mechanically fragile and any stress in the silicon membrane can cause deformation of mask pattern geometry. To improve the mechanical strength of the mask, the 200-mm circular area has to be divided into 1×1 mm square of subfields with struts (reinforce ribs). The membrane mask has to be very thin in order to let most electrons pass through. At sub-100 nm lithography resolution, patterning and etching through several-micrometer-thick silicon membrane for stencil mask becomes difficult to maintain the quality of pattern geometry. Membrane mask is easier to achieve high resolution and pattern quality because the absorber layer is much thinner.

Table 3.5 Comparison of conventional and projection e-beam lithography

	Shaped e-beam direct write	E-beam projection
Typical resolution (nm)	130	100
Acceleration voltage (kV)	50	100
Projection pattern	aperture	mask
Image reduction	Several tens	4:1
Subfield size (μm)	5	250
Deflection (mm)	2	5
Beam current on wafer (μA)	2.5	25
Throughput (wafer per hour)	3–10 (200 mm)	40 (200 mm), 20 (300 mm)

The key fundamental issue is the space charge effect. As seen in Table 3.5, the current for EPL is 10 times greater than for conventional shaped beam lithography. The beam current has to be high in order to achieve high wafer throughput. High current results in strong Coulomb interactions between electrons, which leads to beam blur and deterioration of lithography resolution. It is estimated that at beam current of 25 μA the total blur due to Coulomb interaction can be as high as 80 nm, taking into account the numerical aperture of beam focusing [79]. Further optimizing the focusing column length or increasing beam energy to more than 100 keV may slightly improve the beam blur. The fundamental limit in resolution due to space charge effect is always there as long as the beam current is high.

In order to avoid using a mask and high beam current in EPL, a number of multibeam approaches have been developed [72]. With the multiple beamlets well separated, the Coulomb interactions become negligible. By scanning individual beams to form an exposure pattern collectively, a mask is no longer needed. There are different ways to produce multiple beamlets and to control them. The electrons can be emitted from a single source [80] or from a photocathode illuminated by multiple laser beams (MAPPER) [81, 82]. In the case of single electron source, multiple e-beams are produced by flooding electrons through an array of blanking apertures [80]. Scanning of multiple e-beams can be controlled collectively through a single lens [82], or individually through multiple apertures [80] or by moving the wafer stage relative to the photocathode where the multiple e-beams are generated [81]. Multibeam lithography technology, or commonly called maskless lithography (ML2), is still in very early stage of development. Nevertheless, it represents a future direction of lithography for mass IC manufacturing, particularly at 32-nm technology node downwards and when the cost of masks becomes prohibitively high. Steady progress has been made in the multibeam technology in the last few years. One of the systems, MAPPER, has demonstrated writing 45-nm dense patterns in 2007 [83]. Initially, ML2 will be used as low-volume production tool at throughput of about 5 wafers per hour, and may gradually be developed into a high-volume lithography technology.

In parallel with the development of EPL, ion projection lithography (IPL) also once received great attention in the late 1990s. Ion projection lithography was proposed as an optional technology competing with EPL for next-generation lithography (NGL) [84]. The main reason to develop IPL was the fact that ions scatter little in resist and substrate, hence no proximity effect in ion lithography, and resists are several orders of magnitude more sensitive to ions than to electrons. For IPL, hydrogen or helium ions are used because they have greater penetration depth in resists than other heavy ions such as gallium. However, IPL suffers the same drawback as the EPL, which is the need of a mask. For IPL, a stencil mask has to be used because ions cannot penetrate through a thin membrane. Although 50-nm pattern resolution was demonstrated by IPL [85], the development of IPL was short lived, mainly because the leading IC manufacturers favored EUVL as the choice of NGL. Although not being developed

into a lithography tool for IC manufacturing, IPL still plays a role in nanoscale patterning either by exposure of resists or by direct sputtering machining. A newly developed projection maskless patterning (PMLP) tool has successfully avoided using the delicate stencil mask and can produce up to 40,000 programmable beams working parallel, using H+, He+, Ar+, C60- or cluster ions [86]. With PMLP technology 2D and 3D design data can be transferred directly into almost any material surface of a substrate without using resist. With a novel ion optics system of 200× image reduction, the PMLP tool has demonstrated 16-nm half-pitch lines. Even 11-nm half-pitch lines are visible. It provides a viable technology for making 32-nm node complex photomasks, which the current e-beam mask-making tools are struggling to achieve. It is predicted to be possible to have more than 900,000 programmable beams with <10 nm resolution capability, so that making a 22-nm node leading-edge complex mask can be done in less than 10 h and making a nanoimprint template can be done in less than 1 h [86].

References

1. Haller, I., M. Hazakis, and R. Srinivasan, *High resolution positive resists for electron beam exposure*. IBM J. Res. Develop., 1968. **12**(251).
2. Mayadas, A.F. and R.B. Laibowitz, *One-dimensional superconductors*. Phys. Rev. Lett., 1972. **28**: p. 156.
3. Yamazaki, K. and H. Namatsu, *5-nm-order electron-beam lithography for nanodevice fabrication*. Jpn. J. Appl. Phys., 2004. **43**(6B): p. 3767.
4. Swanson, L.W., *Liquid metal ion sources – mechanism and application*. Nucl. Instrum. Methods Phy. Res., 1983. **218**(1–3): p. 347.
5. Cui, Z., *Electron beam lithography (Chapter 3)*, in *Micro-Nanofabrication Technologies and Applications*. 2006, Springer.
6. Harris, L.A., *Introduction to Electron Beam Technology*, R. Bakish, Editor. 1962, John Wiley & Sons.
7. Cui, Z. and L. Tong, *Optimum geometry and space charge effect in vacuum microelectronic devices*. IEEE Electron Dev., 1993. **ED-40**(2): p. 448.
8. *GEMINI field emission electron optics: Core technology for the LEO SUPRA range*. [cited; Available from: www.smt.zeiss.com].
9. Cui, Z. and L. Tong, *A new approach to simulating liquid metal ion sources*. J. Vac. Sci. Technol., 1988. **B6**(6): p. 2104.
10. Prewett, P.D. and D.K. Jefferies, *Characteristics of a gallium liquid metal field emission ion source*. J. Phys. D: Appl. Phys., 1980. **13**: p. 1747.
11. Hawkes, P.W. and E. Kasper, *Principles of Electron Optics*. 1989, Academic Press.
12. Rouse, J., X. Zhu, and E. Munro, *Solution of electron optics problems with space charge in 2D and 3D*. Proc. SPIE, 1995. **2522**: pp. 375–386.
13. Broers, A.N., *High-resolution thermionic cathode scanning transmission electron microscope*. Appl. Phys. Lett., 1973. **22**: pp. 610–612.
14. Bracher, B.H. and P.D. Prewett, *Electron beam spot size calculation and software*, in *Rutherford Appleton Laboratory Internal Report* 1997.
15. Munro, E., *Electron and ion optical design software for integrated circuit manufacturing equipment*. J. Vac. Sci. Technol., 1997. **B15**(6): p. 2692.
16. Murata, K., T. Matsukaw, and R. Shimizu, *Monte Carlo calculations on electron scattering in a solid target*. Jpn. J. Appl. Phys., 1971. **10**(6): p. 678.

17. Parikh, M. and D.F. Kyser, *Energy deposition function in electron resist films on substrates*. J. Appl. Phys., 1979. **50**(2): p. 1104.
18. Cui, Z., *MOCASEL: A total solution to electron beam lithography simulation*. Proc. SPIE, 1999. **3676**: p. 494.
19. Parikh, M., *Self-consistent proximity effect correction technique for resist exposure (SPECTRE)*. J. Vac. Sci. Technol., 1978. **15**(3): p. 931.
20. Parikh, M., *Correction to proximity effects in electron beam lithography. II. Implementation*. J. Appl. Phys., 1979. **50**(6): p. 4378.
21. Hintermaier, M., et al., *Proximity correction using computer aided proximity correction (CAPROX) evaluation and application*. J. Vac. Sci. Technol., 1991. **B9**(6): p. 3043.
22. Owen, G. and P. Rissman, *Proximity effect correction for electron beam lithography by equalization of background dose*. J. Appl. Phys., 1983. **54**(6): p. 3573.
23. Jackel, L.D., et al., *Beam energy effects in electron beam lithography: The range and intensity of backscattered exposure*. Appl. Phys. Lett., 1984. **45**: p. 698.
24. Wu, B. and A.R. Neureuther, *Energy deposition and transfer in electron beam lithography*. J. Vac. Sci. Technol., 2001. **B19**(6): p. 2508.
25. Kyser, D.F., *Spatial resolution limits in electron beam lithography*. J. Vac. Sci. Technol., 1983. **B1**(4): p. 1391.
26. Xia, X., et al., *Fabrication of near-infrared and optical meta-materials on insulating substrates by lift-off using PMMA/Al stack*. Microelectron. Eng., 2007. **84**(5–8): pp. 1144–1147.
27. Cui, Z., *TEMP: A software package for simulating resist heating in e-beam lithography*. Proc. SPIE, 1998. **3331**: pp. 420–430.
28. Babin, S., P. Hudek, and I. Kosic, *Quantitative measurement of the resist heating in a variable shaped electron lithography*. J. Vac. Sci. Technol., 1997. **B15**(2): p. 311.
29. Yang, H., et al., *Low-energy electron-beam lithography of ZEP-520 positive resist*, *Proceedings of the 1st IEEE International Conference on Nano/Micro Engineered and Molecular Systems*. 2006. p. 391.
30. Lee, Y.H., et al., *Low voltage alternative for electron beam lithography*. J. Vac. Sci. Technol., 1992. **B10**(6): p. 3094.
31. Yoshizawa, M., et al., *Challenges to ultra-thin resist process for LEEPL*. J. Photopolym. Sci. Technol. 2004. **17**(4): pp. 581–586.
32. McCaord, M.A. and T.H. Newman, *Low voltage, high resolution studies of electron beam resist exposure and proximity effect*. J. Vac. Sci. Technol., 1992. **B10**(6): p. 3083.
33. T.P.H. Chang, D.P. Kern, and L.P. Murray, *Arrayed miniature electron beam columns for high throughput sub-100 nm lithography*. J. Vac. Sci. Technol., 1992. **B10**: p. 2743.
34. Spallas, J.P., et al., *A manufacturable miniature electron beam column*. Microelectron. Eng., 2006. **83**: p. 984.
35. Utsumi, T., *Low-energy e-beam proximity lithography (LEEPL): Is the simplest the best?* Jpn. J. Appl. Phys., 1999. **38**(12B): p. 7046.
36. Utsumi, T., *Present status and future prospects of LEEPL*. Microelectron. Eng., 2006. **83**: pp. 738–748.
37. Lindhardt, J. and M. Scharff, *Energy dissipation by ions in the keV region*. Phys. Rev., 1961. **124**: p. 128.
38. Cui, Z., *A Monte Carlo programme for simulating ion penetration in amorphous solid targets,* in *Rutherford Appleton Laboratory Internal Report* 1993.
39. Matsui, S., et al., *Lithography approach for 100 nm fabrication by focused ion beam*. J. Vac. Sci. Technol., 1986. **B4**: p. 845.
40. Randall, J.N., et al., *Masked ion beam resist exposure using grid support stencil masks*. J. Vac. Sci. Technol., 1985. **B3**(1): p. 58.
41. Yang, H., et al., *Low-energy electron-beam lithography of hydrogen silsesquioxane*. Microelectron. Eng., 2006. **83**: pp. 788–791.

42. Yasin, S., D.G. Hasko, and H. Ahmed, *Fabrication of <5 nm width lines in poly(methylmethacrylate) resist using a water:Isopropyl alcohol developer and ultrasonically-assisted development*. Appl. Phys. Lett., 2001. **78**(18): pp. 2760–2762.
43. Hu, W., et al., *Sub-10 nm electron beam lithography using cold development of poly (methylmethacrylate)*. J. Vac. Sci. Technol., 2004. **B22**(4): p. 1711.
44. *Product data sheets*. [cited; Available from: http://www.microchem.com/].
45. Broers, A.N., et al., *Electron-beam fabrication of 8 nm metal structures*. Appl. Phys. Lett., 1976. **29**(9): pp. 596–598.
46. Broers, A.N., *Resolution limits for electron-beam lithography*. IBM J. Res. Develop., 1988. **32**(4): p. 502.
47. Ishii, T., H. Nozawa, and T. Tamamura, *Nanocomposite resist system*. Appl. Phys. Lett., 1997. **70**(9): p. 1110.
48. Dentinger, P.M. and J.W. Taylor, *Increasing plasma etch resistance of resists using fullerene additives*. J. Vac. Sci. Technol., 1997. **B15**(6): p. 2575.
49. You, H.-C., F.-H. Ko, and T.-F. Lei, *Resist nano-modification technology for enhancing the lithography and etching performance*. Microelectron. Eng. 2005. **78–79**: pp. 521–527.
50. Merhari, L., et al., *Nanocomposite resist systems for next generation lithography*. Microelectron. Eng., 2002. **63**(4): pp. 391–403.
51. Namatsu, H., et al., *Nano-patterning of a hydrogen silsesquioxane resist with reduced linewidth fluctuation*. Microelectron. Eng., 1998. **41–42**: pp. 331–334.
52. *Results are from NanoBeam Ltd.* (http://www.nanobeam.co.uk). 2008.
53. Grigorescu, A.E., et al., *10 nm lines and spaces written in HSQ, using electron beam lithography*. Microelectron. Eng., 2007. **84**(5–8): pp. 822–824.
54. Grigorescu, A.E., et al., *Influence of the development process on ultimate resolution electron beam lithography, using ultrathin hydrogen silsesquioxane resist layers*. J. Vacuum Sci. Technol. B, 2007. **25**(6): pp. 1998–2003.
55. Cui, Z., et al., *Profile control of Su-8 photoresist using different radiation sources*. Proc. SPIE, 2001. **4407**: p. 119.
56. Schneider, A., et al., *SU-8 resist and X-ray LIGA technique to produce 1 mm high Ni gear wheel*, in *Proc. 2nd International Workshop on Microfactories*. 2000.
57. Cui, Z., et al., *Comparative study of AZPN114 and SAL601 chemically amplified resists for e-beam nanolithography*. J. Vac. Sci. Technol., 1998. **B16**(6): p. 3284.
58. Cui, Z. and P.D. Prewett, *Comparative study of AZPF514 and UVIII chemically amplified resists for e-beam nanolithography*. Microelectron. Eng., 1999. **46**(1–4): p. 255.
59. Nordquist, K.J., D.J. Resnick, and E.S. Ainley, *Comparison of negative resists for 100 nm electron-beam direct write and mask making applications*. J. Vac. Sci. Technol., 1998. **B16**(6): p. 3289.
60. Chen, W. and H. Ahmed, *Fabrication of 5–7 nm wide etched lines in silicon using 100 keV electron-beam lithography and polymethylmethacrylate resist*. Appl, Phys. Lett., 1993. **62**(13): p. 1499.
61. Hasko, D.G., S. Yasin, and A. Mumtaz, *Influence of developer and development conditions on the behavior of high molecular weight electron beam resists*. J. Vac. Sci. Technol., 2000. **B18**(6): p. 3441.
62. Ocola, L.E. and A. Stein, *Effect of cold development on improvement in electron-beam nanopatterning resolution and line roughness*. J. Vac. Sci. Technol., 2006. **B24**(6): p. 3061.
63. Ocola, L.E., et al., *Lithography for sub-60 nm resist nanostructures*. J. Vac. Sci. Technol., 1999. **B17**(6): p. 3164.
64. Yang, H., et al., *Comparative study of e-beam resist processes at different development temperature*. Microelectron. Eng., 2007. **84**(1109–1112).
65. Chen, Y., H. Yang, and Z. Cui, *Effects of developing conditions on the contrast and sensitivity of hydrogen silsesquioxane*. Microelectron. Eng., 2006. **83**(1119–1123).

66. Sigmund, P., *Theory of sputtering I: Sputtering yield of amorphous and polycrystalline targets*. Phys. Rev., 1969. **184**: p. 383.
67. Fu, X., et al., *Gas assisted etching of niobium with focused ion beam*. Microelectron. Eng., 2005. **78–79**(29–33).
68. Santschi, C., et al., *Interdigitated 50 nm Ti electrode arrays fabricated using XeF2 enhanced focused ion beam etching*. Nanotechnology, 2006. **17**: pp. 2722–2729.
69. Abramo, M., L. Hahn, and L. Moskowicz, *Gas assisted etching: An advanced technique for focused ion beam device modification*. Proc. ISTFA, 1994. **439**.
70. Pease, R.F., et al., *Prospects for charged particle lithography as a manufacturing technology*. Microelectron. Eng., 2000. **53**: pp. 55–60.
71. Pfeiffer, H.C., *Recent advances in electron beam lithography for high-volume production of VLSI devices*. IEEE Trans. Electron Dev., 1979. **ED-26**(4): p. 663.
72. Pfeiffer, H.C., *The history and potential of maskless e-beam lithography*. Microlithography World, 2005. **February**.
73. Pain, L., et al., *Transitioning of direct e-beam write technology from research and development into production flow*. Microelectron. Eng., 2006. **83**: pp. 749–753.
74. Nakayama, Y., et al., *Electron-beam cell projection lithography: A new high-throughput electron-beam direct-writing technology using a specially tailored Si aperture*. J. Vac. Sci. Technol., 1990. **B8**(6): p. 1836.
75. Pfeiffer, H.C. and W. Stickel, *PREVAIL- an e-beam stepper with variable axis immersion lenses*. Microelectron. Eng., 1995. **27**: p. 143.
76. Berger, S.D. and J.M. Gibson, *New approach to projection-electron lithography with demonstrated 0.1 μm linewidth*. Appl. Phys. Lett., **57**(2): p. 153.
77. Okamoto, K., et al., *High-throughput e-beam stepper lithography*. Solid State Technology, 2000. **May**.
78. Yanabe, M., *Status and issues of electron projection lithography*. J. Microlith. Microfab. Microsyst., 2005. **4**(1): p. 011005.
79. Mkrtchyan, M., et al., *Space charge effects in e-beam projection lithography*. Solid State Technology, 2000. **July**.
80. Eder-Kapl, S. et al., *Projection mask-less lithography (PML2): First results from the multi beam blanking demonstrator*. Microelectron. Eng., 2006. **83**: pp. 968–971.
81. Kampherbeek, B.J., et al., *An experimental setup to test the MAPPER electron lithography concept*. Microelectron. Eng., 2000. **53**: pp. 279–282.
82. Coyle, S.T., et al., *Prototype raster multibeam lithography tool*. J. Vac. Sci. Technol., 2002. **B20**(6): p. 2657.
83. News article: *Mapper demos massively parallel e-beam lithography*. Semicondutor International, 2007. **September**.
84. Gross, G., *Ion projection lithography: Next generation technology?* J. Vac. Sci. Technol., 1997. **B15**(6): p. 2136.
85. Hirschera, S. et al., *Ion projection lithography below 70 nm: Tool performance and resist process*. Microelectron. Eng., 2002. **61–62**: pp. 301–307.
86. Platzgummer, E., H. Löschner, and G. Gross, *Projection maskless patterning (PMLP) for the fabrication of leading-edge complex masks and nano-imprint templates*. Proc. SPIE, 2007. **6730**: p. 108.

Chapter 4
Nanofabrication by Scanning Probes

4.1 Introduction

Both photons and charged beams are capable of delineating sub-100 nm patterns, as described in Chapters 2 and 3. However, one question has been missing: that is, at what cost? To break into sub-100 nm scale, many "tricks", apart from short wavelength and high numerical aperture (NA), have to be used in photon-based lithography, and very sophisticated charged optical system is required in charged particle-based lithography. For low-cost nanoscale patterning technologies, scanning probe lithography (SPL) is definitely an alternative to expensive photon or charged beam techniques.

Scanning probe lithography appeared shortly after the advent of scanning tunneling microscope (STM). Researchers actually stumbled onto it when they observed that some lines were left on the sample surface after an extended period of scanning of STM probe [1]. It is now known as local oxidation process initiated by a probe. At the time when the lines were observed, the mechanism was unclear. It was said "under the influence of air some oxidation process seems to occur and maybe SiO_2 is formed". Although the SiO_2 lines drawn by the STM tip were nowhere comparable in quality to those patterned by optical or electron beam (e-beam) lithography, it was the start of drawing patterns at sub-100 nm scale without photons or e-beam, and a STM was a very simple and cheap setup. Since then there was a burgeoning research and development in SPL, because scanning probe-based lithography could be done in any laboratories with a scanning probe microscope (SPM). The invention of atomic force microscope (AFM) further fueled the development.

Conventional lithography using either photons or charged beams always has to rely on a polymer material (resist) as an imaging layer. Scanning probe lithography, however, can be implemented with diverse mechanisms. Thin polymer resist can be exposed by electrons emitted from a probe. The surface of a sample can be modified under the influence of very high local electric field, such as local oxidation, field-induced decomposition or thermal melting. Materials can be transported from a probe to the sample surface to form deposition patterns, or patterns are formed simply by scratching the surface with a probe mechanically. New ways of patterning with a scanning probe are continuously

Z. Cui, *Nanofabrication*, DOI: 10.1007/978-0-387-75577-9_4,

© Springer Science+Business Media, LLC 2008

being developed. Some of the patterning techniques have been commercially exploited to become a dedicated nanolithography tool. Most of the techniques are still within the laboratory as a by-product of a SPM.

The most attractive feature of SPL is its simplicity. It makes the sub-100 nm patterning accessible to anyone with a SPM. However, one has to be aware of its limitation. Although a STM or an AFM has atomic resolution capability, a STM- or AFM-based lithography does not have the same resolution capability as its imaging, because the mechanism of sensing the existence of an atom is different from making the structural change of surface. The very proximal interaction between a probe and a sample surface has determined that the interaction is extremely "superficial". Patterns made by SPL are never more than a few tens of nanometers. Therefore, special measures have to be taken when transferring the "superficial" patterns into a substrate. Imaging with a SPM is slow, so does with the SPL. Scanning probe lithography is in fact even slower because there has to be enough energy transfer into a substrate in order to induce a structural change of the surface. Scanning probe lithography generally has the scanning speed of sub-micrometer-to-micrometer per second, which is many orders of magnitude slower than a conventional e-beam lithography. The working area of a SPL is also small, making it suitable only for patterning a small area of surface. There has been progress in developing multiple probe systems, so that parallel patterning process can be carried out, very much like the multiple e-beam patterning process described in Chapter 3. Nevertheless, SPL is quite a handy technique for making sub-100 nm structures. It is a useful and versatile tool in current nanoscience and nanotechnology development.

4.2 Principles of SPMs

Scanning tunneling microscope was invented in 1981 by Binnig and Rohrer [2], who received Noble Prize in Physics for the invention in 1986. Figure 4.1 schematically illustrates the basic configuration of a STM. It consists of a sharp conductive tip (stylus) which is usually a tungsten wire sharpened electrochemically. The tip assembly is controlled by a piezoelectric transducer which moves the tip up and down to control the gap between the tip apex and the sample surface with sub-nanometer precision. Although the radius of tip apex may be a few nanometers or a few tens of nanometers, at the very end of the tip there are always a few atoms of tip material protruding closest to the sample surface at a distance about 0.5–1 nm, as illustrated by the inset picture in Fig. 4.1. When a bias voltage (mV to a few volts) is applied between the tip and the sample surface, a tunneling of electrons between the tip atoms and sample surface atoms can occur if their gap distance is within $\sim$1 nm, as explained by quantum mechanics principle. For suitably sharpened tip, gap distance, and bias voltage, the electron tunneling can be confined only between single atoms. An additional piezoelectric transducer in the tip assembly can

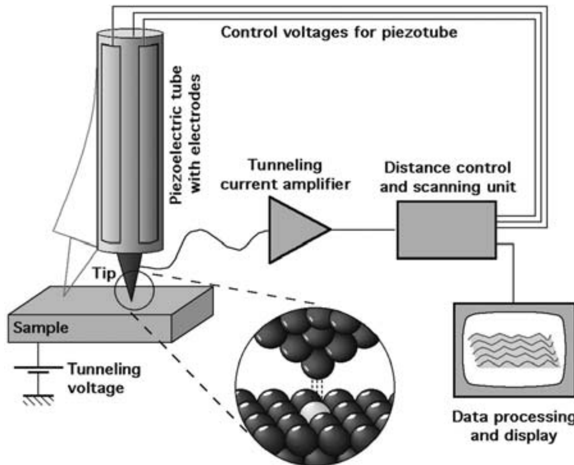


Fig. 4.1 Schematic view of STM (Reprint from Wikipedia)

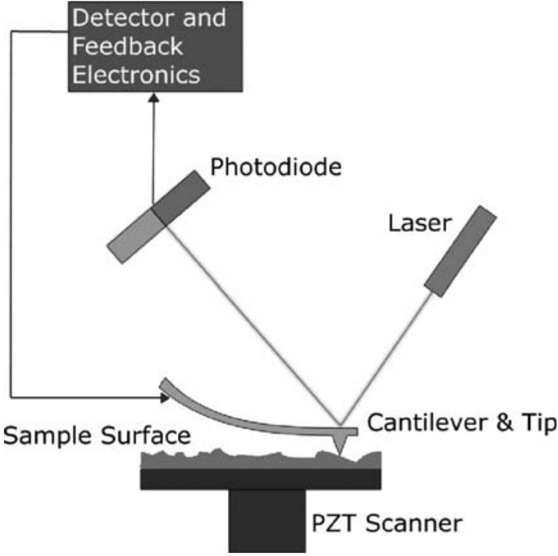
produce raster scan of the tip over small region of the sample surface, creating a two-dimensional (2D) image of the sample surface with atomic resolution.

The key technique in a STM operation is the feedback control of tunneling current. As the tunneling current is exponentially dependent on the tip–sample distance, the feedback control loop can either keep a constant tunneling current or keep a constant tip–sample distance, with sub-angstrom precision vertically and sub-nanometer precision laterally. In the constant current mode, tip–sample distance across the scanning region is varied, creating a 2D topographical map of the surface. In the constant gap mode, variation of electron density of the surface atoms is detected. In either case, a 2D map of surface atoms can be created.

Atomic force microscope was invented shortly after the STM [3]. As schematically illustrated in Fig. 4.2, the key difference between a STM and an AFM is the tip assembly. In an AFM, the tip is mounted on a flexible cantilever beam while the STM tip is rigidly mounted. Atomic force microscope operates by measuring attractive or repulsive forces between the tip and the sample surface. The forces cause the cantilever beam to deflect, which is detected optically and translated to surface topographical variation. A raster scan of the tip over a surface region generates a 2D map of the surface structure with atomic resolution.

Atomic force microscope can be operated either in contact mode or non-contact mode, also called vibrating (tapping) mode. In contact mode, a feedback control loop can either keep a constant height of the tip over the sample surface, or keep a constant contact force between the tip and sample surface. In the tapping mode, the AFM tip is literally tapping the sample surface at high frequency. In all these modes, a 2D map of surface topographical image can be created.

Atomic force microscope is actually a much more versatile technique than the STM. A STM relies on tunneling electrons for imaging, which requires a vacuum environment. Atomic force microscope can be operated in air.



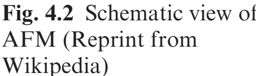


Fig. 4.2 Schematic view of AFM (Reprint from Wikipedia)

Scanning tunneling microscope requires both the tip and sample are conductive, while AFM can be used to image any materials, even in liquid. With a conductive tip, AFM can be operated like a STM to provide typical capabilities of a STM. Over the years, many other types of scanning probe microscopic techniques have evolved from the basic concept of AFM, such as measuring lateral force using a cantilever tip, measuring magnetic force, electrostatic force, Kelvin potential, surface conductivity, capacitance, temperature, etc. [4].

The third major SPM, in addition to STM and AFM, is the near-field scanning optical microscope (NSOM), or sometimes called scanning near-field optical microscope (SNOM). Near-field scanning optical microscope was invented also shortly after the advent of STM [5]. Figure 4.3 schematically illustrates the basic configuration of a NSOM, where a light source is used as a scanning probe which is only a few nanometers above the imaging surface. The light source is usually a sharpened optical fiber connected to a laser source. The fiber may be made of glass or silica, etched in hydrofluoride (HF) acid or heated in a laser and pulled to yield a sharp tip. It must be coated with a metal (usually aluminum but sometimes other metals such as gold) to prevent leakage of the electric field laterally through the probe. At the end of the glass fiber, there is a small opening aperture to let the laser light out, which forms the optical probe. A constant height of fiber probe above a sample surface is maintained by a shear force feedback system. The system employs a turning fork (shown in Fig. 4.3) which is commonly found in digital watches. The optical fiber is attached to one leg of the tuning fork which is oscillated by a digital circuit. As the probe is in contact with the sample surface, the oscillation is damped (as a result of shear forces) and this effect is detected and fed into the control electronics to adjust the height of the probe (using a piezoelectric scanner). A far-field photon

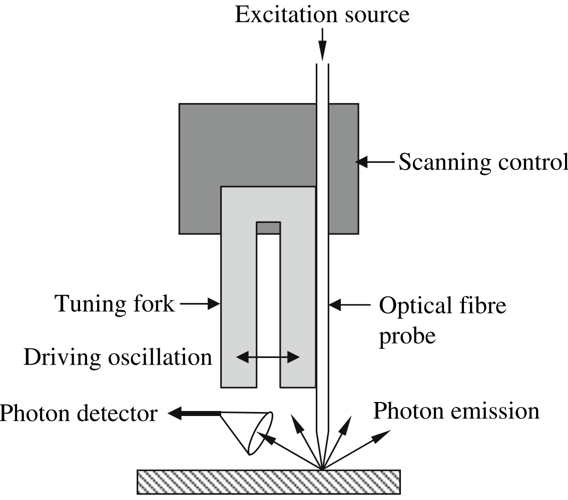


Fig. 4.3 Schematic view of NSOM

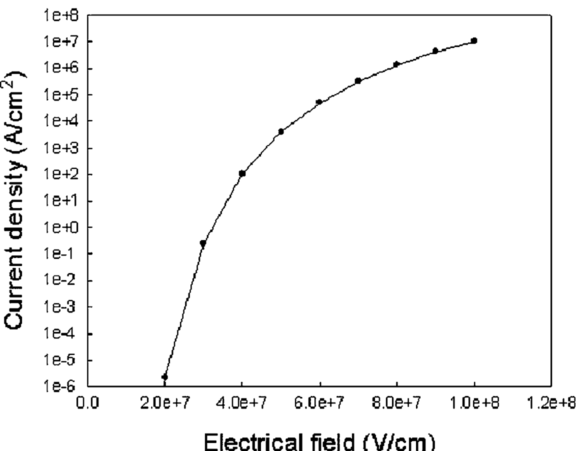
detector is placed underneath the sample to detect the transmitted image of sample surface, or above to detect reflected optical signal. Other configurations of NSOM use either a bent glass fiber acting like a cantilever in a AFM or directly use a AFM-style cantilever in which the pyramidal tip is hollow and has a small see-through aperture. The principle of a NSOM is based on the propagation of near-field optical evanescent wave, which is no longer limited by diffraction. The evanescent wave has been introduced in the description of near-field optical lithography in Section 2.6 of Chapter 2. Imaging by evanescent light wave has resolution far less than the illuminating wavelength. For example, the first demonstration of NSOM achieved 25-nm resolution of imaging with a light source at wavelength of 488 nm [5]. The resolution was only one-twentieth of the illumination wavelength.

All the three SPMs have been explored for nanopatterning, either by direct exposure of resists or modification of surface. They have the common features of high resolution capability, normally below 100 nm, and accurate alignment and positioning because they are after all microscopes. They also suffer the same drawbacks, such as low throughput and limited pattern area. In the following sections, each of the patterning mechanisms is described. Their nanopatterning capability and limits are analyzed.

4.3 Exposure of Resists

4.3.1 Field Electron Emission

It is well known that electron emission occurs at very high electrical field. The high-performance SEMs and e-beam lithography systems have been utilizing



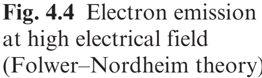


Fig. 4.4 Electron emission at high electrical field (Folwer–Nordheim theory)

field emission cathodes to produce high current density and low energy spread electrons, which has been described in previous chapter. Field electron emission process is governed by the Folwer–Nordheim theory [6], as shown in Eq. (4.1):

$$J = k_1 E^2 \exp\left(-\frac{k_2}{E}\right), \tag{4.1}$$

where J is the field emission current density, E is the field strength, and k_1 and k_2 are constants related to the work function of field emitter. Figure 4.4 shows a plot of field emission current density versus the field strength, assuming the field emitter is made of tungsten which has a work function of 4.5 eV. There is a threshold field ($\sim 2\times 10^7$v cm^{-1}) above which electron emission starts. With the increase of field strength, the emission current density increases rapidly.

There are two ways to achieve high field strength on an electron emitter: using a very sharp emitter or bringing the counter electrode as close as possible to the emitter. Figure 4.5 shows an example of emitter and counter electrode configuration. The emitter has a radius of 50 nm at its apex and the counter electrode is above the emitter with a few tens of nanometers gap. The electrical field at the surface of emitter is numerically calculated using a home-developed computer simulation program [7]. Figure 4.6 shows the surface field strength as the function of gap distance between the emitter and the counter electrode. It is apparent that the field strength can be above the field emission threshold ($\sim 2\times 10^7$ v cm^{-1}) even at 30 V of applied voltage, as long as the emitter tip is close to the counter electrode within 50-nm range. This establishes the basis for electron exposure of resists using a STM or an AFM.

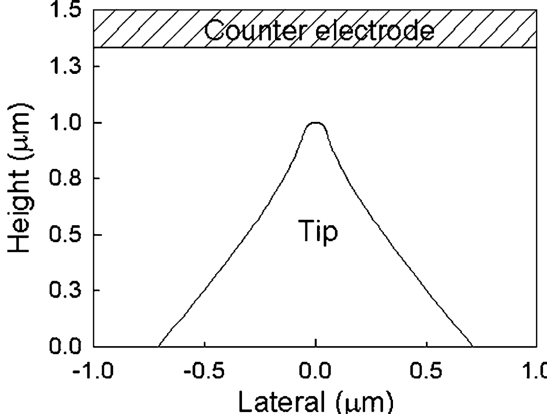


Fig. 4.5 Schematic of field emitter and counter electrode and their relative dimensions

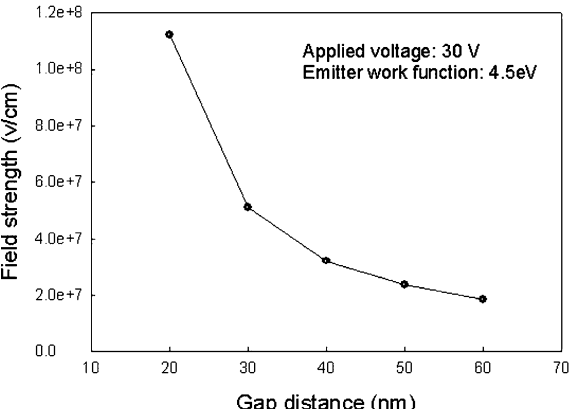
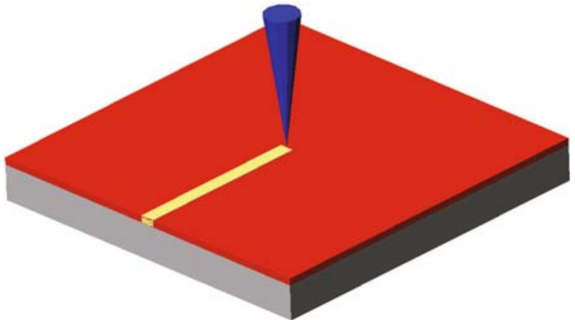


Fig. 4.6 Field strength vs. gap distance between a probe tip and counter electrode

4.3.2 Exposure of Resist by STM

As indicated in Fig. 4.6, emission of electrons can occur in a STM because of the close proximity of the STM tip to a sample surface. Exposure of resist with a STM is simply moving the STM tip on top of a conductive substrate coated with a thin layer of electron-sensitive polymer, while a bias voltage is applied between the substrate and the STM tip, as illustrated in Fig. 4.7. As described in previous section, an AFM can be operated as a STM if a bias voltage is applied between the cantilever tip and a conductive substrate, together with incorporation of a current feedback loop to control the field emission current from the AFM tip [8]. Therefore, exposure of resists by an AFM works the same way as STM.

Exposure of polymethylmethacrylate (PMMA) resist with a STM was first reported in late 1980s. Metal lines 22-nm wide were made with exposure of 20-nm thick PMMA resist and lift-off at 20-V applied voltage and scanning



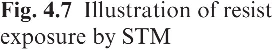
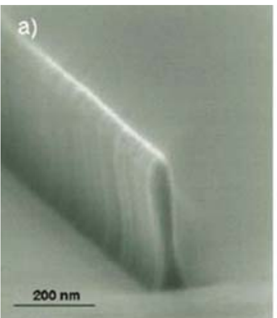


Fig. 4.7 Illustration of resist exposure by STM

speed of 1 μm s^{-1} [9]. Exposure of chemically amplified resist SAL601 using a STM was also reported [8]. When exposing a resist, the STM is normally operated at constant current mode. The applied bias voltage to the tip is much higher, usually a few tens of volts, than when it is operated as a microscope, which is normally a few volts. Exposure of resists with a STM is no different from a conventional e-beam lithography described in Chapter 3, except that electrons emitted from a STM tip are of very low energy and in close proximity to the resist layer. There are some unique features one has to be aware of in STM-based e-beam lithography.

1. Very thin resist layer (<50 nm) has to be used. As seen in Fig. 4.6, the tip has to be very close to the substrate which acts as the counter electrode, in order to generate field electron emission at low bias voltage. The STM tip and the sample substrate are only separated by the resist layer which has to be very thin. On the other hand, the emitted electrons have kinetic energy of only a few tens of volts. They cannot travel very far in resist, as explained in Chapter 3, which requires the resist layer to be thin so that it can be fully exposed. Thin resist layer may have the difficulty of pattern transfer by reactive-ion etching (RIE), as has been mentioned many times before. Multilayer resist system obviously does not work here, because it increases the gap distance between the STM tip and the conductive substrate. It has been shown that chemical-amplified resists generally have high etch resistance. For example, SAL601 resist has higher etch selectivity when used as mask to etch silicon (>5:1) than PMMA (~1:1). So an option in STM-based resist exposure is to use a thin layer of chemically amplified resist. Figure 4.8a shows a 50-nm line etched into 300-nm depth of silicon using STM-exposed SAL601 resist as the etch mask. When PMMA resist is exposed with STM, the only possible route of pattern transfer is by metal lift-off. Another possibility is to use the lift-off metal pattern as the etch mask to further transfer into silicon. Figure 4.8b is such an example where PMMA was first exposed by a STM and metal line pattern was produced by lift-off. The metal line was then used as a mask to etch into silicon.

2. Unlike conventional e-beam systems where a series of electron lenses are employed to focus the electrons into a very fine beam, a STM does not have



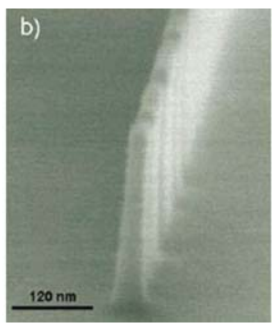


Fig. 4.8 Silicon line patterns produced by STM exposure of SAL601 resist and reactive-ion etching (**a**) and PMMA resist with subsequent lift-off of metal line and RIE using the metal line as mask (**b**) (Reprint from [10] with permission)

lenses between its tip and substrate. Electrons emitted from the STM tip are generally divergent. The degree of divergence depends on both the applied voltage and the separation between tip and substrate [11]. For example, a tip of 10-nm apex radius forms about a 10-nm diameter beam if the separation is 10 nm. The beam diameter becomes 30 nm if the separation is 30 nm, which almost linearly increases with the separation. High bias voltage also causes large divergence of emitted electrons, due to the increase of tip–substrate separation in order to maintain constant current working mode. A 10-nm beam at 20 V of bias voltage can become 40 nm at 40-V bias voltage [11]. Because there has to be a separation between the tip and substrate in order to accommodate the resist layer and the applied bias voltage has to be high enough to generate field electron emission, the resultant beam of electrons is always bigger than the source size. This is the reason why the resolution of STM lithography can never match the atomic resolution of STM imaging.

3. Low energy exposure is the key feature of STM-based lithography. The biggest advantage of low-energy e-beam lithography is it is low or free of proximity effect, as already described in Chapter 3. In conventional e-beam lithography, low-energy electrons (<5 keV) can increase the efficiency of resist exposure, or resists appear to be more sensitive to low-energy electrons. However, this is not the case in STM exposure of resists. Although the field emission electrons from a STM is normally below 50 eV, the efficiency of resist exposure by such low-energy electrons is much lower, or STM exposure of resists requires much higher exposure dose than conventional e-beam lithography. Figure 4.9 compares the linewidth versus exposure dose by conventional e-beam lithography (Fig. 4.9a) and by STM exposure (Fig. 4.9b). Different linewidths were created by multiple scans of beam in 40-nm grid separation. From the comparison it becomes apparent that STM requires 30 times higher exposure dose than conventional e-beam lithography to expose the same patterns. This implies a fundamentally different exposure mechanism in play with STM exposure of resists. Conventional e-beam lithography systems deliver an e-beam beam with high energy into a resist layer. The electrons undergo scattering and slow down until they lose all energy. The STM tip, however, emits electrons which have very low initial

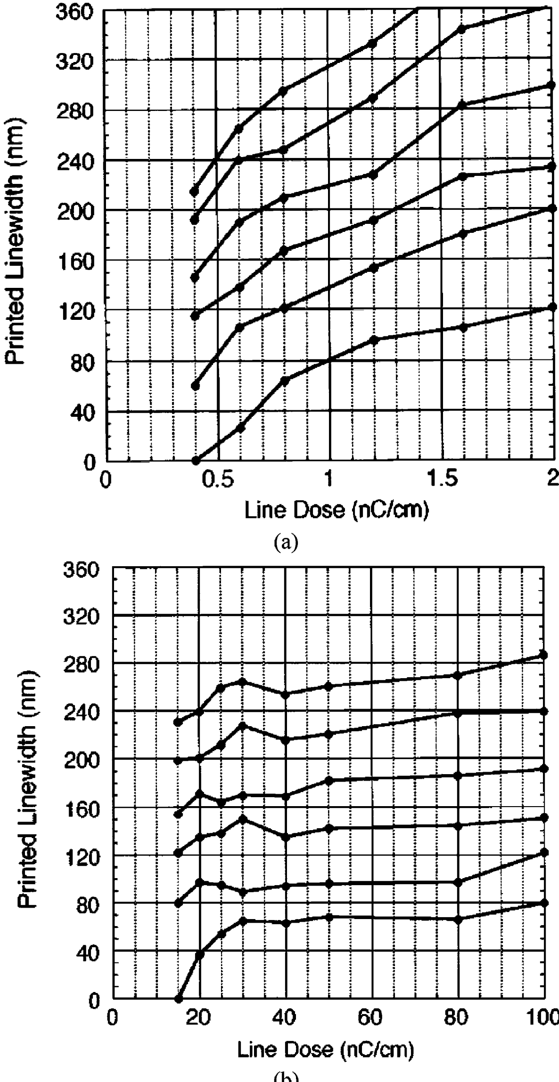


Fig. 4.9 Comparison of line patterns vs. exposure dose (**a**) for conventional e-beam lithography at 30 kV and (**b**) for STM lithography at 40–60 V (Reprint from [10] with permission)

energy (a fraction of eV) when they enter the resist layer. They are accelerated toward conductive substrate. During the process they lose energy to inelastic scattering with resist molecules as well as gain energy from the high field. Such process is perceived less efficient in breaking the molecular chain of a polymer resist than in the case of electrons with initial high energy.

Extremely high exposure dose can only mean slow exposure. Coupled with slow scan speed ($\sim$1 μm s^{-1}), resist exposure by STM is of very low throughput. This would not be a serious problem for laboratory research in which only a

small pattern area and a handful of devices are needed. To propose the technique as an alternative to conventional industrial lithography technologies, the throughput issue has to be seriously dealt with. Massively parallel exposure by array of STM tips has been developed, which will be introduced in later part of this chapter.

The comparison in Fig. 4.9a and b also reveals that STM exposure of resist has much less proximity effect. This is seen from the variation of linewidth with the increase of exposure doses. In conventional e-beam lithography, increase of exposure dose will result in increase of exposed feature dimension because at higher exposure dose large-range scattering electrons acquire enough energy to contribute to the exposure of resists. This is not the case for STM exposure of resists. Increase of exposure dose shows nearly no change of linewidth, as seen in Fig. 4.9b, which means no electrons are found outside the exposure region.

4.3.3 Exposure of Resist by NSOM

It is quite natural to think of using an optical probe to deliver photons to realize optical lithography. While NSOM was invented in mid-1980s [5, 12], its real application in exposure of photoresist was quite a few years later [13]. The exposure is based on near-field evanescent photons, which have been extensively described in Section 2.6 of Chapter 2. The optical probe of NSOM normally has an aperture about 100 nm. One would expect the exposed photoresist features should at least be of the same order. However, experimental results have rarely demonstrated that the NSOM exposure of photoresist can achieve pattern dimension matching the size of optical probe [14]. One of the reasons was that the early results were obtained using a much thick photoresist layer (200–500 nm) [15]. As it is known, the evanescent wave decays exponentially in resist layer within about 40-nm thickness. The resist layer has to be very thin. With thin resist layer, the dimension of exposure feature is limited by the aperture size of optical probe and the distance between the probe and resist surface [16].

To achieve small-size aperture (<100 nm) in an optical probe is not easy, because the end of a tapered glass fiber has to be cut flat to create a window for letting out the light. An alternative is to use apertureless probe. An aperture in an optical probe is created by coating a metal thin film on the surface of a sharpened glass fiber, so that light is only allowed to transmit through the end of fiber, as illustrated in Fig. 4.10a. The glass fiber probe can also be without the metal coating, as shown in Fig. 4.10b, which is called apertureless. However, this type of optical probe is less favorable because of the less optimum light confinement in the probe. The probe can also be a metal tip illuminated by a light wave, as shown in Fig. 4.10c. At the extremity of metal tip, strong localized electrical field can be generated by the excitation of illuminating light wave [17], dubbed as “lightning rod” effect [18], which can act as an optical probe. Metal

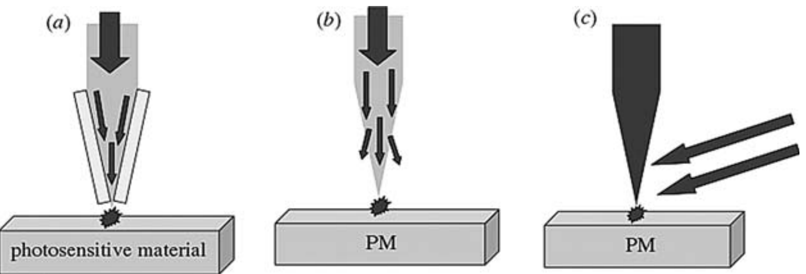


Fig. 4.10 Optical probe types in NSOM: (**a**) glass tip with aperture, (**b**) glass tip without aperture, (**c**) metal tip without aperture (Reprint from [18] with permission)

tip is much easier to make than a glass fiber tip and can be made extremely sharp. Using metal tip apertureless NSOM 40-nm resolution lithography has been achieved.

The optical lithography by NSOM has been extended to pattern alkanethiol molecules on a thin film of gold with extremely high resolution. Alkanethiol molecule is known to absorb spontaneously onto gold surface and form a dense and well-ordered monolayer. The alkanethiol self-assembled monolayer (SAM) is very stable in gold etchant – therefore, can act as an etching mask. Alkanethiol molecules can be oxidized, upon exposure under UV light in the presence of oxygen, to become alkylsulfonate which only binds to gold weakly and can be removed by rinsing or immersion in thiol solution. It is found that exposure of the alkanethiol SAM on gold with an apertured NSOM tip can produce patterns much smaller than the aperture size of NSOM tip. Alkanethiol patterns as small as 20 nm have been created even though the aperture size of NSOM tip is estimated to be about 50 nm [19]. The pattern resolution is found to be dependent on the morphology of gold film. The finer the grain size of polycrystalline gold film, the smaller the pattern dimension, which leads to believe that the asperities of gold grains have acted as metal tips to generate local electric field by the excitation of NSOM light wave. Such highly localized field initiated the oxidation of alkanethiol SAM. There is also a host of other SAMs that have been explored for SPL applications [20].

4.4 Local Oxidation Lithography

In a STM or an AFM setup, the tip is normally negatively biased, acting as cathode, and sample is grounded acting as anode. For a silicon sample, the surface is normally passivated by dipping into buffered HF (H_2O:HF at 25:1) to remove the native oxide (an oxide of a few atomic layers is naturally formed when silicon surface is exposed to air), leaving a single atomic layer of hydrogen

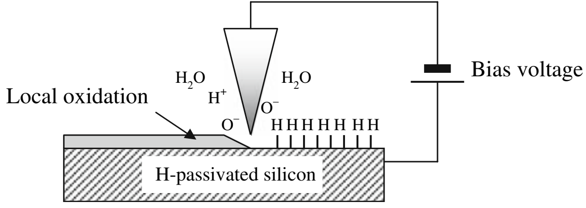


Fig. 4.11 Schematic illustration of field-induced local oxidation

on surface, also called H-passivated or H-terminated [21]. When a STM or an AFM tip is brought to a few nanometers above the surface of a sample, even a few volts of applied voltage between the tip and sample surface can generate electric field up to the magnitude of $\sim 10^8$ V cm^{-1}. At such a high electrical field, molecules can be broken up into ions. Local oxidation of sample surface takes place when there are water molecules, due to humidity in air, around the extremity of tip. For silicon, the monolayer of hydrogen after passivation can be dissociated from silicon surface. The silicon surface is then oxidized at the tip–sample high field region, as schematically illustrate by Fig. 4.11.

The local oxidation reaction takes the following steps [22], as expressed in Eq. (4.2):

$$rmSi + 2h^+ + 2(OH^-) \rightarrow Si(OH)_2 \rightarrow SiO_2 + 2H^+ + 2e^-. \qquad (4.2.)$$

Water molecules (H_2O) are ionized at the high electrical field. Negative ions (OH^-) are driven toward the sample surface (anode). Silicon is oxidized. Positive hydrogen ions are attracted to the tip (cathode) to combine with the released free electrons to become hydrogen gas molecules (H_2). Apart from silicon, a variety of other materials can also be oxidized by STM or an AFM, including compound III–V semiconductors, silicon carbide, and a number of metals such as titanium, tantalum, aluminum, molybdenum, nickel, and niobium. For metallic surface, the oxidation follows a slightly different reaction [21], as expressed in Eq. (4.3):

$$rmM + nH_2O \rightarrow MO_n + 2nH^+ + 2ne^-. \qquad (4.3.)$$

There are a number of factors that can influence the oxidation process, including applied bias voltage, dwelling time of tip at one location of oxidation, and ambient humidity. The growth rate of oxide can be measured by the height of oxide grown out of sample surface. For oxidation of silicon, it was found that the height of SiO_2 linearly increases with the increases of applied bias voltage (–2–20 V) to the tip and dwelling time (10 ms–1000 s) [23], and this has been confirmed by an analytical model [24]. However, the oxide growth rate is high only at the beginning of oxidation (first 50 ms), then slows down and drops. This is because the increase of oxide thickness has reduced the field strength

across the tip and substrate silicon. The oxide growth practically stops at field strength $<10^7$V cm^{-1}, unless the bias voltage is further increased.

Humidity plays an important role in the formation of oxide structure during tip-induced oxidation process. Figure 4.12 compares oxide line features made by AFM-induced oxidation of silicon at different humidity levels. At 61% of humidity, the oxide lines were measured around 90 nm (Fig. 4.12a). It reduced to around 22 nm at 14% of humidity level although all other conditions were the same (Fig. 4.12b). Humidity level not only can influence the amount of water molecules being supplied to the oxidation reaction, but also can influence the electrical field distribution around the tip–sample region. At high humidity level, a water film is formed, which can increase the lateral distribution of electrical field, creating an equivalent "defocus" effect. Therefore, the oxide growth area becomes wider.

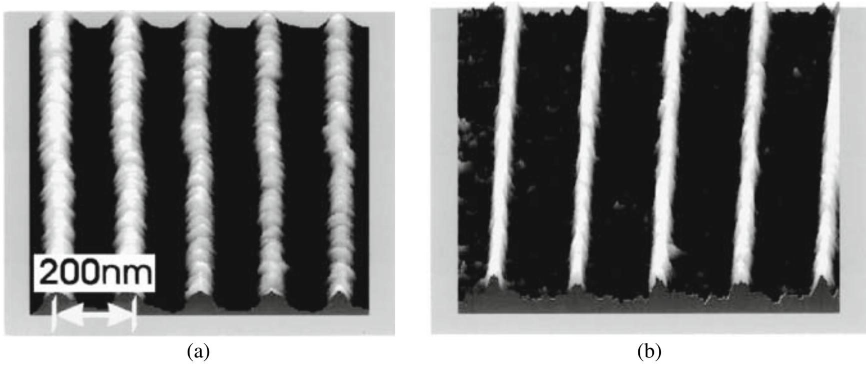


Fig. 4.12 Atomic force microscope images of oxide lines produced at (**a**) 61% of humidity and (**b**) 14% humidity (reprint from [23] with permission)

Local oxidation was first discovered with STM scanning over silicon surface [25]. However, STM is not the best instrument for local oxidation lithography. This is due to the way STM operates. As described earlier, a STM operates either in constant current mode or constant distance mode. The bias voltage depends on either the tunneling current or the tip–sample distance. When oxide grows underneath the tip, both the tunneling current and the tip–sample distance (formation of oxide bump) are changing, which will make the bias voltage change because of the feedback control loop. Damage of STM tip often occurs because the STM often fails to recognize the increase of surface height as the oxide layer grows. When measuring the oxide height during oxidation process, STM can give false measurement, overestimating the height because of the decrease of tunneling current through the oxide layer.

Atomic force microscope is, on the other hand, better suited for local oxidation lithography. Atomic force microscope holds two distinct advantages over

the STM: (1) the bias voltage which is the root cause of oxidation can be applied independently of the feedback control of the tip–sample spacing; and (2) real oxide height is directly measured without involving tunneling current [26]. Atomic force microscope working in tapping mode is even better because it eliminates lateral shear forces and overcomes tip–sample adhesion forces (e.g., capillarity). These properties avoid damages on the surface and improve imaging and lithography resolution. Table 4.1 compares the performances of STM, AFM in contact mode (CM-AFM), and AFM in tapping mode (TM-AFM) for local oxidation lithography.

Table 4.1 Comparison of STM and AFM for local oxidation lithography (Reproduced from [26] with permission)

	STM	CM-AFM	TM-AFM
Speed	$< 8\ \mu m\ s^{-1}$	$< 128\ \mu m\ s^{-1}$	$< 20\ \mu m\ s^{-1}$
Resolution	$\geq$10 nm	$\geq$30 nm	$\geq$10 nm
Thickness	$\leq$1.5 nm	$\leq$8 nm	$\leq$4 nm
Reproducibility	$<$50%	$>$50%	$>$75%

Unlike exposure of resists, patterned oxide can be directly used either as a mask to etch into silicon, because of the high selectivity between silicon dioxide and silicon in RIE, or as dielectric barriers in variety of electronic devices. One of the first nanoelectronic devices operating at room temperature was fabricated by using local oxidation nanolithography [27]. Other devices, in which SPM-based local oxidation lithography has played key roles, include data storage memories, conducting nanowires, side-gated field-effect transistors, single-electron transistors, superconducting quantum interference devices, quantum points and rings, microlens, and templates for the growth of biomolecules and conjugated materials [22].

4.5 Additive Nanofabrication

4.5.1 Field-Induced Deposition

Though the resist exposure and local oxidation lithography have been the main nanofabrication techniques using a scanning probe, there are a wide range of other techniques under the name of SPL. One of the techniques is to form nanoscale patterns by transferring materials from the tip to the surface of a sample under high electrical field.

It is well known that ions or atoms can be directly pulled out of material surface under extremely high electrical field. The phenomenon is called field evaporation [28]. The extremely high field can be obtained from a sharp tip close to a counter electrode. As mentioned earlier, STM is such a system that

high field exists at the STM tip at only a few volts of bias voltage. Field evaporation of material from a STM tip was first observed in early 1990s. In the experiment, a gold STM tip was used and negative pulse voltage of 3–4 V was applied to the tip. Gold dot was formed on substrate surface typically 15–20 nm across its base and 2–3 nm in height [29]. Material deposition was easily observed from a gold tip because gold has relatively low threshold field for field evaporation (3.5 V A^{-1}) [30], and gold surface is inert to chemical contamination. If a tungsten tip was used in combination with a gold substrate, a pit was formed instead of a mound [29]. This is because tungsten has much higher threshold in field evaporation (5.7 V A^{-1}). Instead of evaporation from tip to substrate, the substrate gold atoms evaporated onto the tungsten tip, forming a pit in the substrate.

Field evaporation alone cannot completely explain the material deposition process from a STM tip. Heating by field emission current may also be responsible for the deposition. Field emission current occurs at much lower threshold field than that of field evaporation. For gold tip, the field emission current becomes considerable at the field as low as $\sim$0.6 V A^{-1} [31]. High field emission current can heat up the extremity of STM tip, causing melting and flowing of tip material. This can also explain why the material deposition from the tip is sustainable despite continuous loss of material from the tip. Melting and flowing of tip material can sustain the supply of atoms toward the apex of tip so that the tip would not become blunt due to the evaporation, which is in much the same way as the working of liquid metal ion sources, as described in Section 3.2.2 in Chapter 3. For this reason, a negative bias to the tip is preferable because negative bias is the correct configuration for field emission to occur. Experimental observation also confirmed that negative bias of the tip produced much stable deposition [31].

Another speculation about the mechanism of tip material deposition process is the formation of nanobridge between the tip and sample surface [32]. The extremely high field at the apex of a tip can cause the tip to elongate. For metallic tip with 50-nm radius and 0.75 nm above a metallic plane, the pulling force at the tip apex is of 6×10^{-10} N. Coupling with the heating of field emission current, the tip can be pulled down to touch the sample surface, so that a liquid metal bridge is formed between the tip apex and sample surface. Deposition then takes place through this liquid bridge. This is very plausible because liquid metal ion sources work in the same way whereby liquid apex elongates at high field strength to form an extremely sharp emitter [33]. The existence of a liquid metal bridge was also indirectly confirmed from the measurement of resistance between STM tip and substrate. The elongation keeps the tip always in sharp geometry and maintains stable deposition over an extended period. Because of the sharpening process at high field, the very end of the tip becomes much smaller than its original geometry. Deposited lines as small as 4–5 nm in width have been made, with minimum separation of 1 nm between the deposited features [32]. One must be aware that field-induced deposition is found to be much easier with metallic tip such as gold or platinum

because they are soft and have low melting point and relatively low threshold of field evaporation.

Field-induced deposition can also be done with the assistance of organometallic gases [34]. Like the local oxidation by STM in ambient humidity, organometallic gas molecules are decomposed at the high field around the extremity of STM tip and metal atoms are deposited on substrate surface directly under the STM tip. The process is often called STM-induced chemical vapor deposition (CVD). The tip is negatively biased. There is a threshold bias voltage for different precursor gases. The threshold bias is 27-V voltage for iron carbonyl gas, but 15 V for tungsten carbonyl. The deposition is normally done at emission current from 100 pA to 500 pA [34]. Unlike the local oxidation which has limit in the height of oxide pattern, high-aspect-ratio patterns can be produced by the gas-assisted deposition. The diameter of deposited dots is mostly dependent on the bias voltage. Higher voltage increases the gap between the tip and the top of the dot, which increases the effective diameter of field emission electrons, as explained earlier, thus increasing the diameter of deposited dots. The pressure of precursor gas within the range from 50 mTorr to 300 mTorr has little effect on the deposition rate or the characteristics of deposition. The deposited film contains about 50% of metal, with rest being carbon contamination and small amount of oxygen. A range of other materials can also be deposited, based on the same principle.

Atomic force microscope works the same way as STM and can transfer materials from tip to substrate at a biased voltage. As AFM tips are normally made of silicon or silicon nitride. Metallic materials have to be coated onto the AFM tip in order to make the deposition happen. Once it is coated with metals, it becomes no different from a STM tip, and deposition takes place under high electrical field. As AFM can work in liquid solution, it is possible to initiate electrochemical deposition using an AFM tip. Copper has been electrochemically deposited onto GaAs surface by immersing the AFM tip into a mixture of $CuSO_4 + H_2SO_4$ solution [35]. The electrodeposition of Cu is strongly dependent on the structure of the substrate, applied potential, and concentration of Cu^{2+} ions in the precursor solution. Based on the same principle, many other metallic materials can be electrochemically deposited with AFM.

4.5.2 Dip-Pen Nanolithography

The most well known additive nanofabrication technique using a scanning probe is the Dip-pen nanolithography (DPN). Writing with a fountain pen on paper is a familiar example of how ink is transported from the pen to a substrate. Dip-pen nanolithography works on the same principle. Although DPN technology was first reported in 1999 [36], similar phenomena have been found since the invention of AFM. It was found long ago that when used in air,

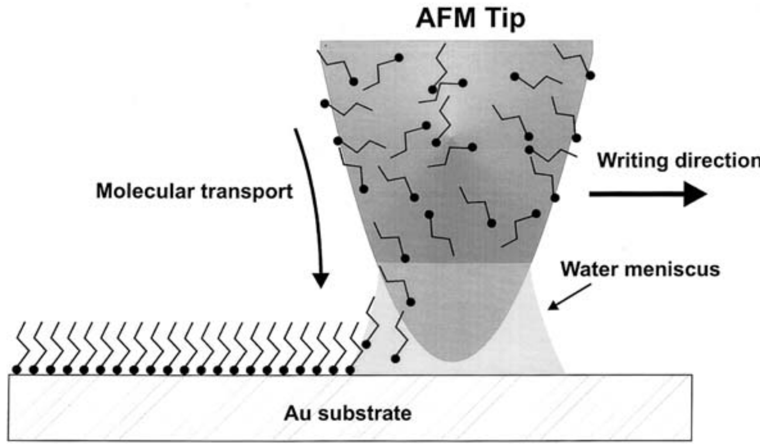
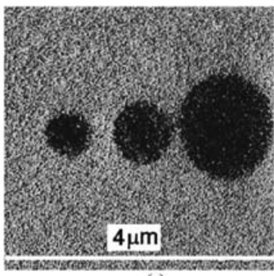


Fig. 4.13 Schematic of Dip-pen nanolithography (Reprint from [36] with permission)

condensed water is formed between the AFM tip and sample surface because of humidity. Depending on the humidity level and wetting property of the substrate, water molecules could move between the AFM tip and substrate surface, which was difficult to control and has been plaguing the AFM imaging operation. The DPN technique made use of this phenomenon. By using special liquids whose molecules can self-assemble onto substrate surface once in contact with the surface, the transport of molecules only goes in one direction, that is, from tip to substrate. In this way molecule patterns can be formed by scanning the AFM probe [36]. Figure 4.13 shows schematically how the ink is deposited from a probe to substrate. There is a liquid meniscus formed between the tip and substrate, which provides the route for liquid molecules to move toward the substrate surface.

In order to ensure this one-way transport, the preferred choice of ink–substrate combination is the thiol-based molecular ink and gold substrate. Thiol-group molecules are known to form stable SAM on gold [37]. One of such thiol-group molecules used in the first DPN experiment was 1-octadecanethiol (ODT). Figure 4.14a is the lateral force microscope (LFM) image of ODT dots deposited onto polycrystalline gold surface. The three dots represent AFM tip coated with ODT in contact with gold substrate for 2, 4, and 16 min. By moving the tip along the surface, line patterns can be drawn with ODT ink, as shown in Fig. 4.14b. The factors that can influence the resolution of DPN patterning are (1) the grain size of substrate, very much like the texture of a paper when writing with a fountain pen; (2) surface chemisorption property which determines the diffusion of molecules after deposition; (3) tip dwelling time on the surface, and (4) humidity level which affects the size of meniscus formed between the tip and substrate. Rough surface, high absorption surface, long dwelling time, and high



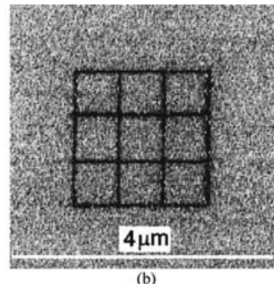


Fig. 4.14 Patterns made with DPN: (**a**) *dots* of different sizes representing tip-substrate contact time of 2, 4, and 16 min, (**b**) *line* patterns drawn by scanning tip (Reprint from [36] with permission)

humidity level can all contribute to large size of deposited patterns. Dip-pen nanolithography technique can routinely produce sub-100 nm patterns, as by all other SPL techniques. However, to make features below 50 nm requires optimization of various conditions. The smallest features around 10–15 nm by DPN have been demonstrated in Chad Mirkin's group [38].

Although DPN started with the ink–substrate combination of thiol-gold, many other combinations of ink and substrate materials have been explored in recent years. A comprehensive review of the research and development work in DPN is given in Ref. [39]. The DPN technology has been commercially exploited and has become a successful nanopatterning tool, with potential applications in many nanoscience and nanotechnology areas [40].

4.6 Subtractive Nanofabrication

4.6.1 Electrochemical Etching

Scanning tunneling microscope and AFM can be used for etching substrate materials in acidic solution. One of the early experiments made etching of silicon and GaAs surface by scanning a STM tip over the substrate surface in a buffered HF solution [41]. In order for the STM to work in liquid, the STM tip was coated with a wax, which ensured excellent insulation and was electrochemically inert in the electrolyte. The tip was positively biased at 1.4 V and tunneling current was 1 nA. By continuously scanning 25 s over 200×200 nm area, a 2-nm depth of etched pit was formed. The etching mechanism was speculated to be STM-induced oxidation of silicon followed by the oxide etched in buffered HF, because no etching was found if the STM scanned on silicon surface in H_2SO_4 solution instead of buffered HF. If the etching is due to the oxidation of silicon surface, then the oxidation is quite different from previously described local oxidation process in which the STM tip is negatively biased. With negatively biased STM in a buffered HF solution, etching of silicon surface was also observed. In this configuration, the etching mechanism was

thought to be due to direct attack of HF to the H-terminated silicon surface under the electrical field [42]. The STM tip does not have to operate in an acidic solution. Etching has been observed in a damp CO_2 atmosphere. Gas-assisted etchings of superconductive thin film and n-GaAs have also been reported [43, 44]. So far electrochemical etching of material using scanning probes is only of limited success. The quality of etched nanostructures is poor and the mechanism of etching is still not fully understood.

4.6.2 Field-Induced Decomposition

It has long been discovered that STM or AFM tips can directly make pits or hillocks on a sample surface at pulsed bias voltage. The material transportation from a STM or AFM tip has been described in Section 4.5.1. Excluding mechanical indentation or thermal melting, the only cause of pit formation is probably related to the high field generated by the scanning probe. In the presence of a high field, the desorption activation energy for a neutral adsorbate remains relatively unchanged at room temperature, but the ionic potential energy function is greatly perturbed. If the field is high enough ($\sim$10 V nm^{-1}), the potential barrier for ionization may become comparable to the thermal energy of the adsorbate, in which case ionization can occur. Once ionized, the adsorbate follows the electric field lines away from the surface. The energy barrier that normally dominates desorption kinetics is effectively circumvented by the reduction of the barrier to ionization. So the surface atoms escape like evaporation at high electrical field.

The most-studied materials for tip–sample interaction are gold and graphite. A number of studies were reported on tungsten tips interacting with gold surface either in air or in vacuum. Holes were formed ranging from a few nanometers to a few tens of nanometers [45]. The holes were only a few nanometers deep. In one experiment, a sudden jump of current spike was detected, which led to the suggestion that the hole in gold surface was formed from field evaporation of gold atoms toward the tungsten tip [46], as described in Section 4.5.1. Highly oriented pyrolytic graphite (HOPG) is the most widely used material for STM and AFM surface imaging. Holes were made on HOPG surface whether intentional or unintentional if the bias voltage was raised higher than normal imaging bias (mV to 3 V) [47]. The diameter of the hole is typically 2–10 nm. The higher the bias voltage is the bigger the holes become. However, holes can only be formed in air. Reducing the air pressure to below 50 Torr can cause hole formation to be degraded, and below 10 Torr it ceases completely. Venting the evacuated chamber with 1 atmosphere of O_2 or N_2 did not cause the hole formation to resume, but venting with air, or with 20 mTorr of H_2O vapor, the hole formation starts again [45]. Based on the observation, field-induced chemical reaction has been proposed to explain this phenomenon [48]. To sublimate carbon atoms directly from graphite, the binding energy of 7.43 eV has to be broken down by the external

field. If there are H_2O molecules participating in the process, only 1.82 eV of external energy is required to decompose graphite into CO and H_2 gases.

4.6.3 Thermomechanical Indentation

When an AFM tip is heated up, it can make dents in low-melting-point polymer materials. In this way, patterns of nanometer scale can be made into polymer surface. The first thermal mechanical patterning by AFM was reported in early 1990s [49], in which the AFM tip was heated by a pulsed laser beam and in contact with a PMMA substrate. The AFM tip was heated to above the softening point of PMMA (120°C). The size and depth of indentation depends on the temperature of the tip and the force of the tip pressing into the polymer. At about 10^{-7}N pressing force, the dents were typically about 100 nm across and 10–20-nm deep. The AFM tip can also be heated by building a microscale heater into the cantilever beam, which makes the thermomechanical indentation system much compact. An example of such integrated cantilever is schematically shown in Fig. 4.15.

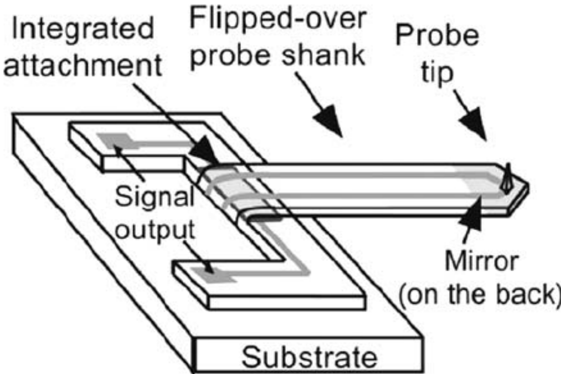


Fig. 4.15 Schematic of integrated heater-cantilever-tip assembly (Reprint from [50] with permission)

The initial concept of thermomechanical indentation has later been taken up for high-density storage application [51]. In this way, the binary data "1" and "0" are represented by dent and no dent on the polymer surface. For high-density data storage, there are two key parameters, the data density and data writing rate. The density is directly related to size of dents. Thermomechanical writing is a combination of applying a local force by the cantilever/tip to the polymer layer and softening it by local heating. Initially, the heat transfer from the tip to the polymer through the small contact area is very poor. It only improves as the contact area increases. This means that the tip must be heated to a relatively high temperature (about 400°C) to initiate the melting process. Once melting has commenced, the tip is pressed into the polymer, which increases the

heat transfer to the polymer. The volume of melted polymer increases, which leads to the increase of indentation size. Such a highly nonlinear heat-transfer process makes it very difficult to achieve small tip penetration and thus small bit sizes. Early work used thick polymer layer, which was found difficult to control the size of indentation. This problem could be overcome by using very thin polymer layers deposited on Si substrates. The Si substrate prevents the tip from penetrating farther than the film thickness allows, and it enables more rapid transport of heat away from the heated region because Si is a much better conductor of heat than the polymer. With a 40-nm film of PMMA, bit sizes ranging between 10 nm and 50 nm have been achieved. One drawback emerged with this approach, which was the increased tip wear, because of the frequent contact between the Si cantilever tip and the hard Si substrate during writing. A 70-nm layer of cross-linked photoresist (SU-8) was added between the Si substrate and the PMMA film to provide a soft landing for the cantilever tip. With this configuration, dents of 40 nm in 120-nm pitch were made in the PMMA layer, which is equivalent to 400 Gb in^{-2} data storage density [51].

4.6.4 Mechanical Scratching

Perhaps the most straightforward nanoscale patterning by a scanning probe is direct scratching, or plowing, of a sample surface. By applying a force on an AFM tip or by controlling the cantilever deflection during scanning, mechanical scratches on hard or soft surfaces have been made with widths of tens of nanometer and depths of a few nanometer. Figure 4.16 shows three 500-nm long grooves scratched by a Si_3N_4 cantilever tip on a two-layer GaSb/InAs film. Figure 4.16a is the AFM image of scratched grooves and Fig. 4.16b is the depth profile of these grooves. The stiffness of the cantilever is 0.37 N m^{-1} and the applied forces vary from 100 nN to 150 nN with a scanning rate of 0.02 μm s^{-1}. The grooves produced on the 5-nm GaSb/20-nm InAs substrate are about 65 nm in width and 4 nm in depth.

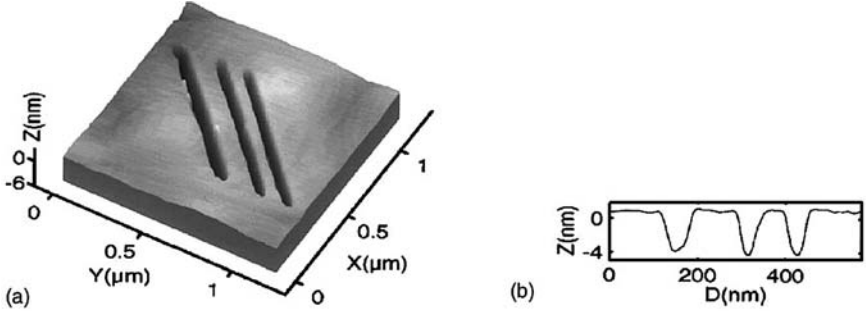


Fig. 4.16 (**a**) AFM image of scratched grooves. (**b**) Depth profiles of these grooves (Reprint from [52] with permission)

Scratching hard surface can wear the tip and reduce its lifetime. Coating the AFM tips with hard materials or making the tip directly in hard materials such as diamond can improve its lifetime. Both carbon-coated Si_3N_4 tips and single-crystal diamond tips have been used for mechanical scratch patterning, achieving scratch depth up to 80 nm at the applied force up to 30 μN [53].

The resolution of mechanical scratching depends on the size of the tip and the depth of scratch. It is of the highest resolution among all the SPL techniques described so far, because the geometry of a tip is fixed while electrical field or field emission electrons spread a much larger range than the boundary of a solid tip apex. Even a large tip can produce a small scratch by controlling the force and depth. For example, a 3-nm wide groove was made using a v-shaped 30-nm radius Si_3N_4 tip with applied force of 100 nN [54]. The downside of such high-resolution scratch patterns is the very shallow depth. The mechanical scratch patterns made with standard AFM tips are no more than a few nanometers deep. For example, the above 3-nm groove was only about 1-nm deep. Such shallow patterns are of very limited use. One way to extend the use of these mechanically scratched shallow patterns is to employ the top surface imaging (TSI) concept which has been widely used in conventional optical or e-beam lithography. One of the reported examples used a bilayer resist system consisting of a 3-nm thick Ti film on top of a 65-nm thick PMMA layer. Atomic force microscope with Si tip scratched the 3-nm Ti film with a tip force of a few millinewtons and the scratched pattern is then transferred to the underneath PMMA layer by an isotropic O_2 RIE [55].

Soft materials such as polymer are obviously very easy to scratch. Atomic force microscope has been used for direct patterning of PMMA resist, and metal nanowires can be created by subsequent lift-off process. In this case, normal silicon AFM tips can be used with applied force of about micronewtons. The scratching depth of 15 nm was achieved and more than 50 scratches could be performed with the same AFM tip because of the less damage to the tip with polymer material [56]. However, because the mechanical scratching is simply shoveling away the material from the plowing path, generation of debris along the plowing path is inevitable. For shallow scratching, debris may be less a problem. For soft materials, plowing by a scanning tip may simply squeeze the material to both sides of plowing path, leaving additional topography on the patterned surface. Figure 4.17 shows such an example. Figure 4.17a is the AFM image of a nanogroove formed by AFM plowing on PMMA surface. The bright edges on both sides of the groove represent higher than surface plane topography. Figure 4.17b is the profile of the groove. The bumping edges are over 20-nm higher than the PMMA surface plane, which is caused by squeezed material from the scratched trench.

If the film of polymer can move away while being plowed by an AFM tip, then the bumping edge on the sides of groove can be avoided. Such was the idea of using a strained polymer film for AFM plow patterning [57]. The idea was to use a conducting polymer poly(3-hexylthiophene-2,5-diyl) (P3HT) whose molecules are packed in a head-to-tail fashion. This leads to a physical effect such

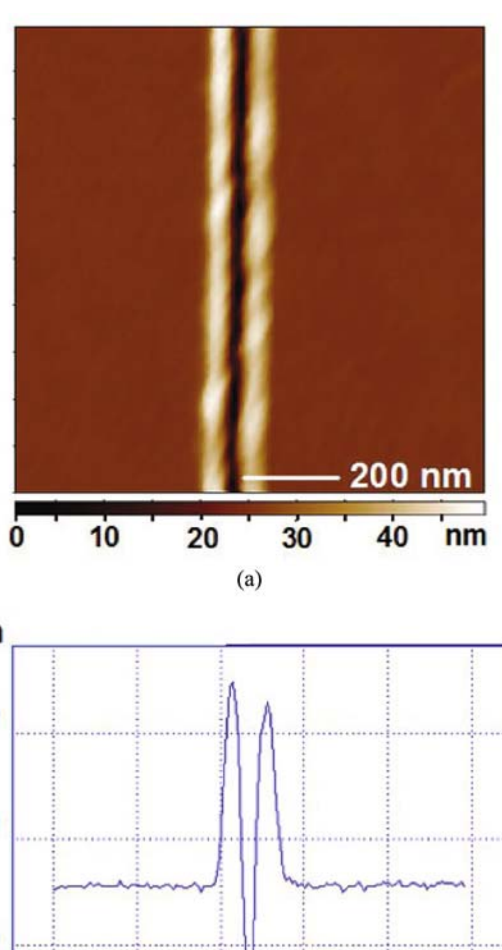


Fig. 4.17 (**a**) AFM image of plowed PMMA groove by silicon tip. (**b**) Profile of the groove in (**a**) (Reprint from [56] with permission)

that an internal tensile strain is sustained within the spin-coated P3HT films. When an AFM tip plows through the film, the internal tensile strain pulls away the film from the cutting, leaving an open trench at the plowing path. By varying the polymer film process conditions and AFM plowing parameters, trenches with widths spanning from 40 nm to 2.3 μm have been fabricated [57]. The final trench width is also determined by the film adhesion to the substrate, which acts against the released strain. Therefore, different trench widths can

be achieved either by varying the total internal strain, for instance, by varying the film thickness, or by modifying the film adhesion to substrate surface. In general, the more hydrophobic the substrate surface, the less the adhesion of the P3HT film. Surfaces with an enhanced hydrophobicity can be prepared by applying a SAM, such as octadecyltrichlorosilane (OTS).

Just like the STM and AFM can individually pick and place atoms on a surface, they can also modify a monolayer of molecules by shoveling away the molecules. Working with SAMs is much easier because it requires less force in the mechanical plowing process and causes less wear to the tip. A number of studies have investigated thiol-functionalized DNA on gold surface with AFM. By applying force slightly higher than what is needed for AFM imaging (30–50 nN), DNA molecules were mechanically shaved away and the gold surface was exposed [58]. A patterned SAM is created. In this case the patterned SAM can act as a mask either for subsequent patterning of gold or for adding some other molecules to the patterned area. One of the experiments immersed the SAM-covered substrate in a solution which contained molecules different from the SAM. When the SAM molecules were shaved away by AFM, they were replaced by the reactive molecules in the solution [59]. There are a whole range of techniques to pattern SAMs with a scanning probe, which has been thoroughly reviewed in Ref. [20].

4.7 High-Throughput SPL

The superior resolution capability of SPL has fueled widespread R&D effort in 1990s to develop it into a competing or an alternative technology for wafer-scale nanolithography. However, as C.F. Quate pointed out, four criteria dictated the viability of SPL to compete with existing lithographic technologies [60]:

- resolution,
- alignment accuracy,
- reliability,
- throughput (wafers per hour).

The fundamental strengths of SPL are their resolution ability, as has been described in the above sections, and alignment accuracy because of their atomic-scale imaging capability. Reliability is somewhat an unproven issue as different SPL methodologies may have different reliability problems, such as the resist exposure, local oxidation, and DPN technique. The key nagging issue for the SPL is throughput. The throughput issue can be tackled in two aspects, increasing the speed of single probe and using multiple probes.

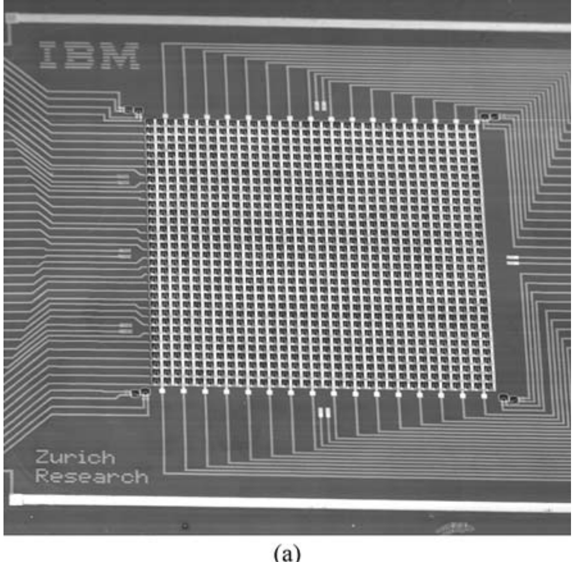
Scanning probe microscope is notoriously slow as a microscope. When SPM is compared to other real-time forms of microscopy such as the scanning electron microscope (SEM), searching for the desired feature on a sample is an extremely time-consuming process. Normal STM or AFM scans at speed of

1–10 μm s^{-1}. To increase the speed of probe scanning, there are issues of feedback control loop, which has to work at high frequency, and of probe assembly itself which has to move at high speed. The conventional piezoelectric tube which controls the height of probe has low resonance frequency and cannot respond to height variation at high-speed scanning of probe. One way to increase the speed is to operate an AFM at constant height mode so that the piezoelectric tube does not have to respond to height variation during high-speed scan. Combined with the image capturing scheme used in SEM, a scan speed of 1 mm s^{-1} was achieved [61]. Another development was to replace the piezoelectric tube with a ZnO piezoelectric actuator integrated to the cantilever. ZnO film was coated onto the cantilever beam, which bent the cantilever up to 4 μm in vertical direction over the topographic features on a sample surface. The integrated cantilever–actuator assembly could achieve much higher resonance frequency than a piezoelectric tube. Working at constant force mode, the integrated cantilever probe achieved 3 mm s^{-1} scanning speed over a sample surface containing 2-μm vertical steps [62].

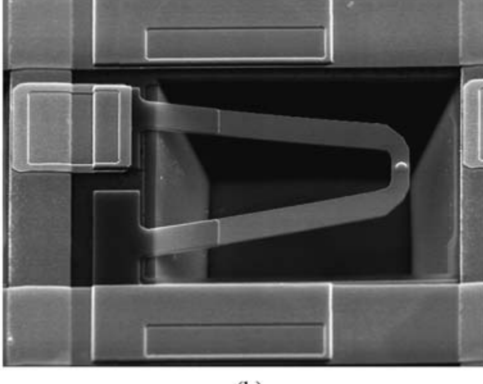
Having a high-speed scanning probe is only half of the story for high-throughput lithography. Surface modification has to take place at high speed, too. The surface modification for the purpose of lithography includes either exposure of resists or local oxidation. Early work of SPL was using PMMA as the resist. Although it is of high resolution, PMMA is too low in sensitivity. Coupling with low efficiency of electron exposure by SPL (described in Section 4.3.2), much lower throughput can be expected. Use of high-sensitivity resists such as chemically amplified resists is one possible route to improve the throughput. The chemically amplified resist SAL601 has been used for STM lithography, which is capable of achieving scan speed of 1 mm s^{-1} [63]. An interesting development was the use of spin-on-glass (SOG) as a resist to achieve high-speed lithography [64]. Spin-on-glass can be spin-coated like a resist and has been commonly used in IC manufacturing as a planarization film. The siloxane type of SOG consists of a methyl (CH3) group and silanol groups (Si–OH) bonded to the Si atoms in the Si–O backbone. It was found that electron exposure of SOG could decompose the methyl group, thereby decreasing the organic content of SOG, resulting in the exposed region soluble in 100:1 buffered oxide etching (BOE) solution. The selectivity of exposed and unexposed regions in BOE solution is 20:1, making it a perfect resist layer for SPL. The most striking feature of the SOG resist is its high sensitivity to electron exposure. Line features with 100-nm width were achieved at scanning speed as high as 3 mm s^{-1}. The minimum 40-nm line features were obtained at scanning speed of 180 μm s^{-1}, 1 nA current, and 74-V bias voltage [64].

Even at 3 mm s^{-1} scanning speed, patterning with a single probe is nowhere near the throughput level that the current mainstream lithographic technologies have. The only way to further increase the throughput of SPL is to use multiple probes and to pattern in parallel fashion. There is no fundamental hurdle to build multiple probes. It is very much an engineering implementation issue to extend from single probe to multiple probes. Each probe has to be

individually controllable, which requires multiple channels current and feedback control circuits and built-in sensor and actuator for each individual probe. The first attempt was the construction of two-probe system. The exposure of SAL601 at 1 mm s^{-1} scanning speed for each probe was demonstrated [65]. Making massively parallel scanning tips became possible with the development of MEMS (micro-electro-mechanical system) fabrication technology. IBM demonstrated the first 32×32 cantilever array chip for high-density data storage application, that is, the Millipede [51]. Figure 4.18a shows the complete cantilever array chip with interconnects to surrounding control circuits and Fig. 4.18b is one of the cantilevers with a silicon tip at the top. The cantilever array was fabricated on a silicon-on-insulator (SOI) wafer, requiring 11 levels of



(a)



(b)

Fig. 4.18 (**a**) 32×32 cantilever array chip. (**b**) Individual cantilever (Reprint from [66] with permission)

photolithography patterning and multisteps RIE and metal depositions [66]. Because of the batch processing nature similar to the manufacturing of ICs, the resultant 1024 cantilevers were identical and highly uniform.

Another successful example of parallel SPL is the DPN technology. Dip-pen lithography has been introduced in Section 4.5.2. Because the DPN is mostly for delivering liquid molecules to a sample surface, control of emission current or heating current for individual tips is not necessary. The fabrication of large array of DPN tips is relatively simple. The tips can be constructed in 2D array, instead of 2D matrix, as those shown in Fig. 4.18a. 55,000 "dip-pens" in 2D arrays were fabricated by a self-aligning process [67]. The dip-pen array has been used to generate 88 million dots. Each dot is around 100-nm diameter and 400-nm spacing to the adjacent dots.

There is still a long way to go before the SPL can be used as a high-throughput tool to compete with the current mainstream lithographic technologies for IC manufacturing. However, scanning probe-based lithography can have many applications other than competing for IC manufacturing. Examples have already been demonstrated, such as the Millipede for high-density data storage and DPN array for patterning protein arrays. In this aspect, the SPL is rather a complement of existing semiconductor fabrication process. It must be said that a large user base and wide application spectrum are still in the realm of single-probe instruments. New patterning techniques and capabilities using SPM have been continuously developed and they are rapidly becoming important in many areas of science and technology, particularly in the current development of nanoscience and nanotechnology [68].

References

1. Ringger, M., et al., *Nanometer lithography with the scanning tunnelling microscope*. Appl. Phys. Lett., 1985. **46**(9): p. 832.
2. Binnig, G. and H. Rohrer, *Scanning tunnelling microscopy*. Helv. Phys. Acta, 1982. **55**(6): pp. 26–735
3. Binnig, G., C.F. Quate, and C. Gerber, *Atomic force microscope*. Phys. Rev. Lett., 1986. **56**(9): p. 930.
4. West, P. and A. Ross, *An Introduction to Atomic Force Microscopy Modes*. 2006, Pacific Nanotechnology, Inc.
5. Pohl, D.W., W. Denk, and M. Lanz, *Optical stethoscopy: Image recording with resolution $\lambda/20$*. Appl. Phys. Lett., 1984. **44**(7): pp. 651–653
6. Folwer, R.H. and L.W. Nordheim, Proc. R. Soc. London, 1928. **A119**: p. 173.
7. Cui, Z. and L. Tong, *Optimum geometry and space-charge effects in vacuum microelectronic devices*. IEEE Trans. Electron Devices, 1993. **40**(2): p. 448.
8. Soh, H.T., K.W. Guarini, and C.F. Quate, *Resist exposure using field-emitted electrons*, in *Scanning Probe Lithography*. 2001, Kluwer Academic
9. McCord, M.A. and R.F.W. Pease, *Lift-off metallization using poly(methyl methacrylate) exposed with a scanning tunnelling microscope*. J. Vac. Sci. Technol., 1988. **B6**(1): p. 293.
10. Wilder, K., et al., *Electron beam and scanning probe lithography: A comparison*. J. Vac. Sci. Technol., 1998. **B16**(5): p. 3864.

11. Mayer, T.M., D.P. Adams, and B.M. Marder, *Field emission characteristics of the scanning tunnelling microscope for nanolithography*. J. Vac. Sci. Technol., 1996. **B14**(4): p. 2438.
12. Betzig, E., et al., *Near-field scanning optical microscopy (NSOM) – development and biophysical applications*. Biophys. J. 1996. **49**(1): pp. 269–279.
13. Froehlich, F.F., T.D. Milster, and R. Uber, *High-resolution optical lithography with a near-field scanning subwavelength aperture*. Proc. SPIE, 1993. **1751**: pp. 312–320.
14. Leggett, G.J., *Scanning near-field photolithography – surface photochemistry with nanoscale spatial resolution*. Chem. Soc. Rev., 2006. **35**: pp. 1150–1161.
15. Smolyaninov, I., D.L. Mazzoni, and C.C. Davis, *Near-field direct-write ultraviolet lithography and shear force microscopic studies of the lithographic process*. Appl. Phys. Lett., 1995. **67**(26): p. 3859.
16. Riehn, R., et al., *Near-field optical lithography of a conjugated polymer*. Appl. Phys. Lett., 2003. **82**: p. 526.
17. Novotny, L., R.X. Bian, and X.S. Xie, *Theory of nanometric optical tweezers*. Phys. Rev. Lett., 1997. **79**: pp. 645–648.
18. Royer, P., et al., *Near-field optical patterning and structuring based on local-field enhancement at the extremity of a metal tip*. Phil. Trans. R. Soc. Lond., 2004. **A362**: pp. 821–842.
19. Sun, S. and G.J. Leggett, *Matching the resolution of electron beam lithography by scanning near-field photolithography*. Nano Lett., 2004. **4**(8): pp. 1381–1384.
20. Kramer, S., R.R. Fuierer, and C.B. Gorman, *Scanning probe lithography using self-assembled monolayers*. Chem. Rev., 2003. **103**(11): pp. 4367–4418.
21. Day, H.C. and D.R. Allee, *Selective area oxidation of silicon with a scanning force microscope*. Appl. Phys. Lett., 1993. **62**(21): p. 2691.
22. Garcia, R., R.V. Martinez, and J. Martinez, *Nano-chemistry and scanning probe nanolithographies*. Chem. Soc. Rev., 2006. **35**: pp. 29–38.
23. Avouris, P., T. Hertel, and R. Martel, *Atomic force microscope tip-induced local oxidation of silicon: Kinetics, mechanism, and nanofabrication*. Appl. Phys. Lett., 1997. **71**(2): p. 285.
24. Stievenard, D., P.A. Fontaine, and E. Dubois, *Nanooxidation using a scanning probe microscope: An analytical model based on field induced oxidation*. Appl. Phys. Lett., 1997. **70**(24): p. 3272.
25. Dagata, J.A., et al., *Modification of hydrogen-passivated silicon by a scanning tunneling microscope operating in air*. Appl. Phys. Lett., 1990. **56**: p. 2001.
26. Fontaine, P.A., E. Dubois, and D. Stievenard *Characterization of scanning tunneling microscopy and atomic force microscopy-based techniques for nanolithography on hydrogen-passivated silicon*. J. Appl. Phys., 1998. **84**(4): p. 1776.
27. Snow, E.S., et al., *A metal/oxide tunneling transistor*. Appl. Phys. Lett., 1998. **72**: p. 3071.
28. Muller, E.W. and T.T. Tsong, *Field Ion Microscopy. Principle and Applications*. 1969, Elsevier.
29. Mamin, H.J., et al., *Gold deposition from a scanning tunneling microscope tip*. J. Vac. Sci. Technol., 1991. **B9**(2): p. 1398.
30. T.T. Tsong, *Field ion image formation*. Surf. Sci., 1978. **70**: pp. 211–233.
31. Chang, C.S., W.B. Su, and T.T. Tsong, *Field evaporation between a gold tip and a gold surface in the scanning tunneling microscope configuration*. Phys. Rev. Lett., 1994. **72**(4): p. 574.
32. Houel, A., et al., *Direct patterning of nanostructures by field-induced deposition from a scanning tunneling microscope tip*. J. Vac. Sci. Technol., 2002. **B20**(6): p. 2337.
33. Cui, Z. and L. Tong, *A new approach to simulating liquid metal ion sources*. J. Vac. Sci. Technol., 1988. **B6**(6): p. 2104.
34. McCord, M.A. and D.D. Awschalom, *Direct deposition of magnetic dots using a scanning tunneling microscope*. Appl. Phys. Lett., 1990. **57**(20): p. 2153.
35. Koinuma, M. and K. Uosaki, *AFM tip induced selective electrochemical etching and metal deposition on p-GaAs(100) surface*. Surf. Sci., 1996. **357–358**: pp. 565–570.
36. Piner, R.D., et al., *Dip-pen nanolithography*. Science, 1999. **283**: pp. 661–663.

37. Xia, Y. and G.M. Whitesides, *Soft lithography*. Angew. Chem. Int. Ed., 1998. **37**: p. 550.
38. *Mirkin's group*. [cited; Available from: http://chemgroups.northwestern.edu/mirkingroup/].
39. Ginger, D.S., H. Zhang, and C.A. Mirkin, *The evolution of Dip-pen nanolithography*. Angew. Chem. Int. Ed., 2004. **43**: pp. 30–45.
40. *Nano Ink Corp*. [cited; Available from: http://www.nanoink.net/].
41. Nagahara, L.A., T. Thundat, and S.M. Lindsay, *Nanolithography on semiconductor surfaces under an etching solution*. Appl. Phys. Lett., 1990. **57**(3): p. 270.
42. Ye, J.H., et al., *Local modification of n-Si(100) surface in aqueous solutions under anodic and cathodic potential polarization with an in situ scanning tunneling microscope*. J. Vac. Sci. Technol., 1995. **B13**: p. 1423.
43. Thomson, R.E., J. Moreland, and A. Roshko, *Surface modification of YBa2Cu3O7-delta thin films using the scanning tunneling microscope: Five methods*. Nanotechnology, 1994. **5**: p. 57.
44. Kaneshiro, C. and T. Okumura, *Nanoscale etching of GaAs surfaces in electrolytic solutions by hole injection from a scanning tunneling microscope tip*. J. Vac. Sci. Technol., 1997. **B15**: p. 1595.
45. Shedd, G.M. and P.E. Russell, *The scanning tunneling microscope as a tool for nanofabrication*. Nanotechnology, 1990. **1**: pp. 67–80.
46. Li, Y.Z., et al., *Writing nanometer-scale symbols in gold using the scanning tunneling microscope*. Appl. Phys. Lett., 1989. **54**: p. 1424.
47. Schneir J, et al., *Creating and observing surface features with a scanning tunneling microscope*. Proc. SPIE, 1987. **897**: p. 16.
48. Kondo, S., et al., *Surface modification mechanism of materials with scanning tunneling microscope*. J. Appl. Phys., 1995. **78**: p. 155.
49. Mamin, H.J. and D. Rugar, *Thermomechanical writing with an atomic force microscope tip*. Appl. Phys. Lett., 1992. **61**: p. 1003.
50. Basu, A.S., S. McNamara, and Y.B. Gianchandani, *Scanning thermal lithography maskless, submicron thermochemical patterning of photoresist by ultracompliant probes*. J. Vac. Sci. Technol., 2004. **B22**(6): p. 3217.
51. Vettiger, P., et al., *The millipede – more than one thousand tips for future AFM data storage*. IBM J. Res. Dev., 2000. **44**(323).
52. Magno, R. and B.R. Bennett, *Nanostructure patterns written in III–V semiconductors by an atomic force microscope*. Appl. Phys. Lett., 1997. **70**: p. 1855.
53. Filho, H.D.F., et al., *Metal layer mask patterning by force microscopy lithography*. Mater. Sci. Eng., 2004. **B112**: p. 194.
54. Muller, M., et al., *Controlled structuring of mica surfaces with the tip of an atomic force microscope by mechanically induced local etching*. Surf. Interface Anal., 2004. **36**: p. 189.
55. Hu, S., et al., *Fabrication of silicon and metal nanowires and dots using mechanical atomic force lithography*. J. Vac. Sci. Technol., 1998. **B16**: p. 2822.
56. Chen, Y., J. Hsu, and H. Lin, *Fabrication of metal nanowires by atomic force microscopy nanoscratching and lift-off process*. Nanotechnology, 2005. **16**: pp. 1112–1115.
57. Jones, A.G., et al., *Highly tunable, high-throughput nanolithography based on strained regioregular conducting polymer films*. Appl. Phys. Lett., 2006. **89**: p. 013119.
58. Zhou, D., et al., *Use of atomic force microscopy for making addresses in DNA coatings*. Langmuir, 2002. **18**: p. 8278.
59. Xu, S. and G. Liu, *Nanometer-scale fabrication by simultaneous nanoshaving and molecular self-assembly*. Langmuir, 1997. **13**: pp. 127–129.
60. Quate, C.F., *Scanning probes as a lithography tool for nanostructures*. Surf. Sci. 1997. **386**: pp. 259–264.
61. Barrett, R.C. and C.F. Quate, *High speed, large-scale imaging with the atomic force microscope*. J. Vac. Sci. Technol., 1991. **B9**: p. 302.
62. Manalis, S.R., S.C. Minne, and C.F. Quate, *Atomic force microscopy for high speed imaging using cantilevers with an integrated actuator and sensor*. Appl. Phys. Lett., 1996. **68**(6): p. 871.

63. Marrian, C.R.K., E.A. Dobisz, and J.A. Dagata, *Electron-beam lithography with the scanning tunnelling microscope*. J. Vac. Sci. Technol., 1992. **B10**(6): p. 2877.
64. Park, S.W., et al., *Nanometer scale lithography at high scanning speeds with the atomic force microscope using spin on glass*. Appl. Phys. Lett., 1995. **67**(16): p. 2415.
65. Wilder, K., et al., *Nanometer-scale patterning and individual current-controlled lithography using multiple scanning probes*. Rev. Sci. Instrum., 1999. **70**(6): p. 2822.
66. Despont, M., et al., *VLSI-NEMS chip for parallel AFM data storage*. Sens. Actuators, 2000. **80**: pp. 100–107.
67. Salaita, K., et al., *Massively parallel Dip–pen nanolithography with 55000-pen two-dimensional arrays*. Angew. Chem. Int. Ed., 2006. **45**: pp. 7220–7223.
68. Tseng, A.A., A. Notargiacomo, and T.P. Chen, *Nanofabrication by scanning probe microscope lithography: A review*. J. Vac. Sci. Technol., 2005. **B23**(3): p. 877.

Chapter 5
Nanofabrication by Replication

5.1 Introduction

Making sub-100 nm structures can also be done in a simple way, that is, by replication. As long as there is a mold, or a stamp, or a master, which has sub-100 nm surface relief structures, these nanostructures can be replicated in the similar fashion as stamping out millions of compact disks (CDs). This was the idea proposed in 1995 when Stephen Y. Chou first reported sub-25 nm holes made in polymethylmethacrylate (PMMA) polymer with an imprinting mold and he coined the word "nanoimprint" [1]. Nanoimprinting lithography (NIL) technology has since undergone phenomenal growth. There is an annual international conference on NIL since 2002 (International Conference on Nanoimprint and Nanoprint Technology, "NNT"). Many other conferences have dedicated sessions on NIL. Several commercial companies have been established, ranging from producing nanoimprinting tools, nanoimprint stamps, to exploring commercial applications of NIL technology.

Originally, NIL was proposed to replace the expensive optical lithography or electron beam (e-beam) lithography, with the same level of resolution or better but at a fraction of cost. This was not surprising from the analogue of CD manufacturing. Nanoimprinting lithography is a parallel process, just as the optical lithography through a mask, but without the diffraction-induced limit on imaging resolution as the optical lithography has. Resolution wise, it is on par with e-beam lithography, but without the proximity effect that e-beam lithography suffers. Nanoimprinting lithography had achieved sub-10 nm resolution patterning in its early development a decade ago [2]. After more than a decade of development, the goal of succeeding optical lithography to be the next-generation lithography (NGL) tool is nearer. In 2003, the International Technology Roadmap for Semiconductors (ITRS) formerly included NIL as one of the NGL candidates [3]. Trial runs on integrated circuit (IC) manufacturing with NIL are being carried out. Though optical lithography has succeeded in pushing IC manufacturing into 32-nm technology node, the NIL technology still has a fair chance to succeed optical lithography at 22-nm node or below.

Z. Cui, *Nanofabrication*, DOI: 10.1007/978-0-387-75577-9_5,

© Springer Science+Business Media, LLC 2008

Though it is simple and of high resolution capability, NIL in its original form was not without problems and limitations. The original form of nanoimprinting required relatively high temperature to soften the polymer and relatively high pressure to press the stamp into the polymer. Heating and cooling cycle prolongs the nanoimprinting process, which is not ideal for high-throughput manufacturing. Requirement for high pressure can increase the wear of stamp and is restricted in certain applications. Although nanoimprinting based on thermal press is still being widely used today, there have been several other forms of NIL which moved away from high-temperature and high-pressure processes. Notably among the derivatives of NIL is the UV-cured nanoimprinting. Liquid instead of solid polymers are used and once a liquid polymer is molded into shape it is solidified by photopolymerization. Some other techniques have further moved away from the original concept of nanoimprinting. Techniques such as microcontact and soft lithography can no longer be called imprinting at all because the stamps are not pressed into a polymer medium to produce a surface relief pattern. In these techniques, a pattern of new material is transferred to a substrate surface by simply bringing the stamp in contact with the surface. It is a printing rather than imprinting process. However, one thing in common for all the derivatives, including the original form of nanoimprinting itself, is the replication process and the dependence on a stamp.

Nanoimprinting lithography has been the subject of active research and development in the last 10 years. Vast number of papers has been published. This chapter is not intended to cover all the areas of nanoscale replication techniques, but to focus on the issues related to patterning sub-100 nm features, the various ways to achieve this dimension scale, and limitations of each technique. The details of nanoimprinting technology and its many applications have been adequately covered in a book published a few years ago [4].

5.2 Thermal Press Nanoimprint

Nanoimprint relying on heating and pressing was the first proposed replication process at nanometer scale [1]. It can often be confused with hot embossing, as hot embossing refers to the same replication process that relies on heating and pressing a stamp into plastics or other ductile surfaces. The general consensus is that hot embossing refers to thermal press imprint of large structures which is above a micrometer or tens of micrometers, while nanoimprint refers to making small structures often less than a micrometer. Therefore, the process to make plastic microfluidic devices, which are normally of tens or hundreds of micrometers in structure dimension, is called hot embossing, not nanoimprinting. Nanoimprint has much broader meaning as described in the previous section, as some of the nanoimprint techniques may not involve heating and pressing at all.

For thermal press nanoimprint, the basic process is schematically illustrated in Fig. 5.1. A thin layer of polymer (100–200-nm thickness) coated on a flat substrate is first heated to about 50°C–100°C above its glass transition temperature (T_g) so that it is softened. A stamp with surface relief structures is pressed into the thin polymer layer at pressure around 50–100 bars (725–1450 psi) pressure, depending on the viscosity of molten polymer. The imprint depth is slightly less than the polymer layer thickness so that the stamp surface does not have hard contact with the substrate to prevent any damage to the stamp. The next step is to decouple the stamp from the polymer layer (demolding) while it is still soft. The temperature and pressure at different stages of imprinting are plotted in Fig. 5.2. Normally at about 10°C below the T_g the pressure is released. When cooling down to about 50°C, the stamp and the imprinted sample are separated. After the separation of stamp, an impression of the stamp pattern is left in the polymer layer. The final step is to remove the residual polymer inside the imprinted pattern areas by reactive-ion etching (RIE), to lay bare the substrate surface. The polymer patterns obtained after removal of residual polymer are in the same form as those obtained by optical lithography or e-beam lithography. There are a number of key factors which are related to faithful replication of sub-100 nm features in the thermal press nanoimprinting process, such as the stamp, the polymer, the separation of stamp and polymer (demolding), and alignment if multilevel nanoimprinting is needed. These factors are discussed in the following sections.

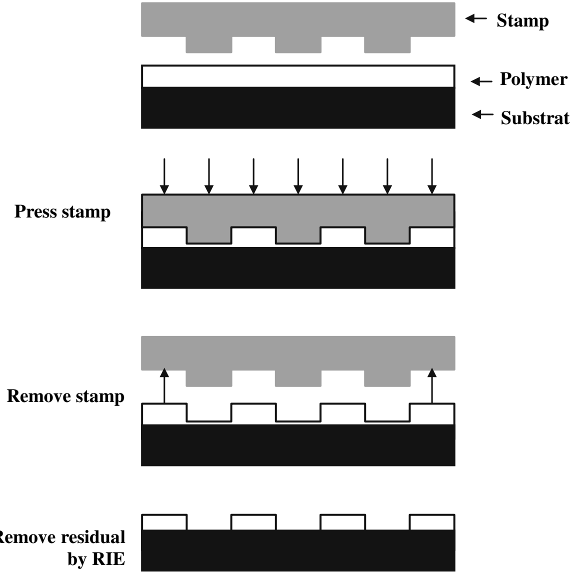
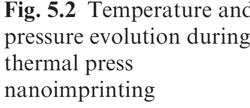


Fig. 5.1 Schematic of nanoimprinting process



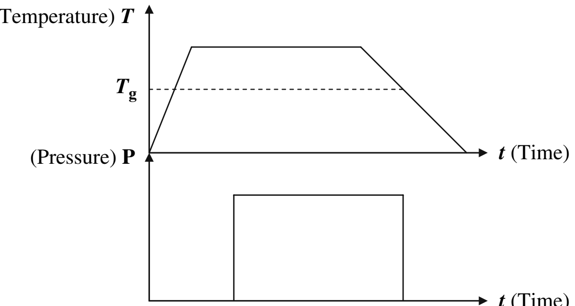


Fig. 5.2 Temperature and pressure evolution during thermal press nanoimprinting

5.2.1 Nanoimprint Stamps

As a master in the replication process, the stamp plays the key role in replicating nanoscale structures in polymer, because the replicated structures are only as good as the master structures. For thermal press nanoimprint, the stamp is made of hard material and should be able to withstand heating and pressure without significant wear. Nanoimprint stamps can be made of many materials. However, they have to be compatible with conventional lithography process. The commonly used materials for stamps are silicon, silicon dioxide, silicon nitride, or various metals. For nanoimprinting of sub-100 nm features, the stamps are exclusively made by e-beam lithography patterning of resist and then transferred into substrate materials. Making sub-100 nm patterns by e-beam lithography has its own problems, such as proximity effect, which has been extensively discussed in Chapter 3. Once the resist patterns are made, they can be transferred by RIE to silicon, or silicon dioxide, or silicon nitride layer on silicon. They can also be transferred by lift-off to make metal structures on silicon. A range of metallic materials can be used, such as gold, aluminum, or chrome. For high-resolution e-beam lithography, PMMA is the best choice of resist. However, PMMA is known to have poor selectivity in RIE. In this case, PMMA can be used as an intermediate imaging layer. After development of PMMA, a lift-off of thin chrome layer (~20 nm) is carried out to transfer PMMA pattern into chrome pattern (be aware of the change of pattern polarity). Then the chrome pattern is used as etch mask to etch into silicon, silicon dioxide, or silicon nitride up to 400-nm depth [5].

Stamps made of lift-off metal do not last very long because the metal patterns do not always stick well to silicon substrate. Stamps made of oxide or nitride are much more durable. However, silicon substrate is a brittle material. It can be easily broken under a nonuniform pressure load. The best and most durable stamp is made of nickel. The nickel stamp fabrication process is schematically illustrated in Fig. 5.3. The fabrication starts with e-beam lithography of PMMA resist on standard silicon wafer. After development of exposed PMMA, a thin metal layer (10–20 nm) of either gold or nickel is coated to cover the whole surface of PMMA layer. Then electroplating of nickel is carried out using the

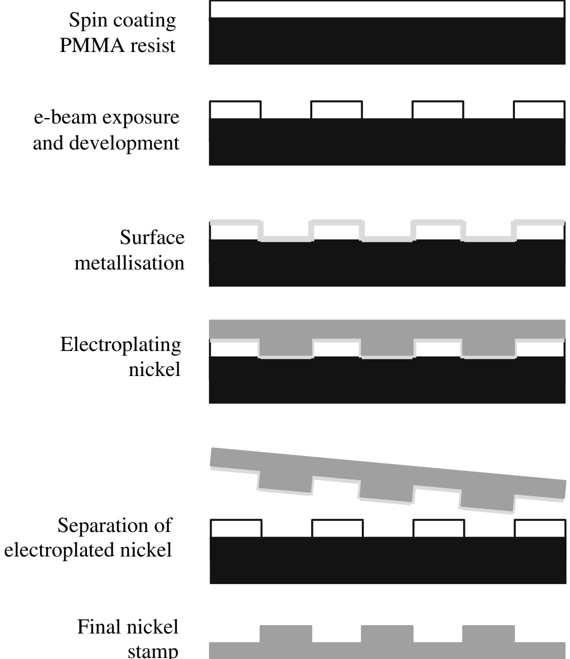


Fig. 5.3 Process for fabricating nickel stamp

metal-covered silicon wafer as the cathode. The electroplated nickel fills up the PMMA pattern cavities and continuous electroplating makes the total thickness to 100–200 μm. Finally, the nickel sheet is separated from the silicon wafer by peeling off the silicon wafer if the nickel sheet is thick and strong enough, or by dissolving the PMMA layer in an acetone bath. The nickel sheet can be fixed onto a metal holder plate to serve as the stamp. A large size of stamp can be made by repeated stamping of a small stamp over a large size of substrate (step-and-stamp nanoimprinting), which is similar to the working of optical stepper. Full 4″-size nickel stamp has been made by the step-and-stamp nanoimprinting with a 3×3 mm^2 silicon master stamp [6].

Stamps can also be made without the pattern transfer step, if the patterned medium is hard enough. The hydrogen silsesquioxane (HSQ) resist, which has been introduced in Chapter 3, has such a property. As previously described, HSQ resist can turn into silicon dioxide upon electron beam exposure or heating at high temperature. Line patterns of 20 nm with 60-nm period have been fabricated in 30-nm thick HSQ by e-beam exposure. With a post-development baking at 500°C, the HSQ becomes essentially silicon dioxide and is hard enough to be used as a stamp for nanoimprint [7].

It must be noted that the above stamps are only for thermal press nanoimprinting. Other nanoimprinting techniques may use different stamps. For example,

UV-cured nanoimprint has to use transparent stamps and microcontact printing may use soft stamps. These other stamps will be described in later sections.

5.2.2 Nanoimprint Polymers

Although a high-resolution stamp is the prerequisite for high-resolution NIL, it is all down to the polymers to realize the resolution capability of the stamp. The basic requirements on a nanoimprinting polymer are low glass transition temperature (T_g) to enable low imprinting temperature, low viscosity to facilitate easy flow of polymer during filling stamp cavity process, low shrinkage after imprinting to maintain pattern fidelity, and high resistance to RIE to serve as a good mask for pattern transfer by RIE. A number of factors can influence the final outcome of nanoimprint, such as applied pressure, heating temperature, thickness of imprinting layer, pattern density, and feature size on the stamp. An imprinting polymer, once heated above its T_g, has to fill the cavities of stamp by compressed flow under applied pressure. The force F required for the squeezing flow can be expressed by Eq. (5.1) [8]:

$$F = -\frac{3\pi R^4}{4h_0^3}\frac{dh}{dt}\eta_0, \tag{5.1}$$

where R is the radial flow distance, h_0 is the initial thickness of polymer layer, dh/dt is the speed of polymer layer being pressed in depth direction, and η_0 is the viscosity of the polymer. The viscosity of polymer depends on both its molecular weight and heating temperature. The higher the molecular weight the more viscous the polymer becomes. The viscosity of 350 kg mol^{-1} PMMA can be more than three orders of magnitude higher than 25 kg mol^{-1} PMMA at the same temperature [9]. With the increase of heating temperature, the viscosity of PMMA decreases exponentially. Figure 5.4 shows the viscosity of 110 kg mol^{-1} PMMA versus temperature [10].

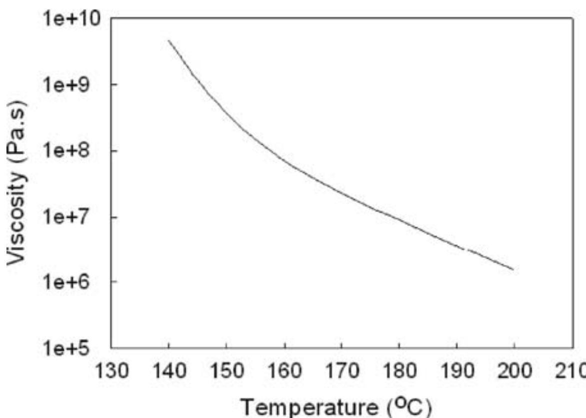


Fig. 5.4 Viscosity of 110 kg mol^{-1} PMMA vs. temperature

The above formula tells that high pressure is required to achieve high speed of imprinting. High pressing force is also required when imprinting either thin polymer layer or large stamp cavity or high viscosity of polymer. High imprinting temperature is also preferred. The direct consequence of low imprinting temperature is the prolonged imprinting time at certain pressure. Imprinting, which may be completed in a few seconds at 200°C, can take hours at 140°C [11]. Thick polymer layer can make the imprinting easier. However, imprinting with thick polymer layer will leave too much residual. The final RIE step to remove the residual will take longer time. Because the RIE process is indiscriminative, it etches both the residual and the imprinted polymer pattern structure, resulting in pattern damage and distortion with excessive RIE. To have the appropriate thickness of residual layer, the amount of polymer should ideally be calculated according to the cavity volume of stamp to make sure the cavities are not underfilled and in the meantime leaving the minimum thickness of residual layer.

Smaller features are easier to imprint than larger features. For example, 500-nm periodical lines can be imprinted at 100-bars pressure and 140°C–150°C temperature in 5 min. At the same pressure and temperature, 100-μm periodical lines cannot be imprinted at all, unless the temperature rises to 200°C [11]. In this aspect, sub-100 nm periodical features are actually easier to imprint because the squeezing flow distance of polymer is much less. The most difficult imprinting task is the imprinting of both large and small patterns. While small feature can be imprinted quickly, large feature cavities may not get filled up completely, as shown in Fig. 5.5 [12]. The dependence of polymer filling on feature dimension and density is very much like the proximity effect in e-beam lithography. In fact, it can be called the imprinting proximity effect [13].

The stamp design has serious impact on the imprinting process if the imprinting proximity effect is to be avoided. An example is shown in Fig. 5.6. For the same pattern structures, the positive stamp structure in Fig. 5.6a (protrusions) causes much less displacement of polymer during imprint, compared to the negative stamp structures in Fig. 5.6b (cavities). In general, pattern design for NIL stamps should make sure that the cavities and protrusions are uniformly distributed over the entire stamp area, to avoid large difference in

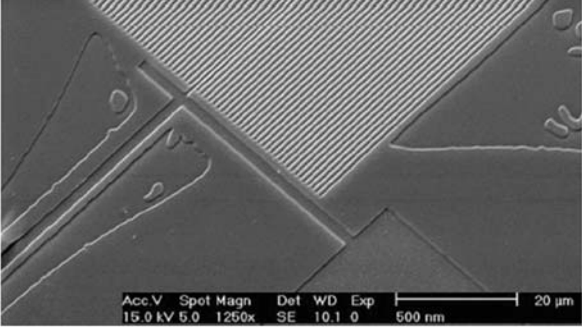


Fig. 5.5 Voids in imprinted structure due to nonuniform distribution of patterns (Reprint from [12] with permission)

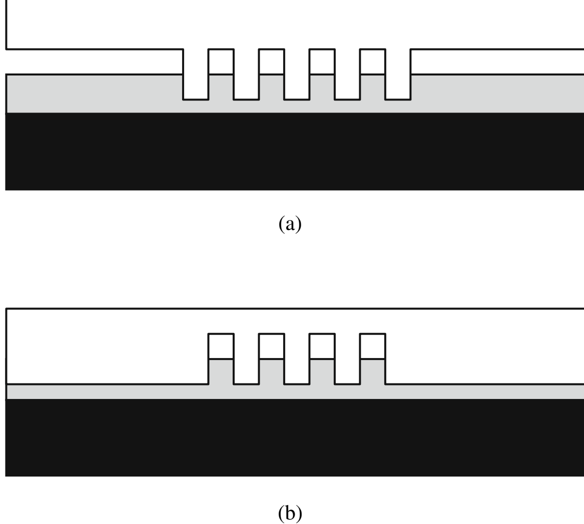


Fig. 5.6 Imprint with different stamp structures: (**a**) Positive stamp with less displacement of polymer required, (**b**) negative stamp with more displacement of polymer required

polymer squeezing flow distance. This can reduce the occurrence of incomplete filling and variation of residual layer thickness. Optimizing the pressure, temperature, and imprinting time in combination with appropriate polymer material property can sometime minimize the proximity effect and achieve uniform filling of stamp cavities with dimension from sub-100 nm to millimeter [14].

Polymethylmethacrylate (PMMA) is the commonly used polymer in nanoimprint, because it is readily available in e-beam lithography. The other polymer, also widely used in NIL, is polystyrene (PS). Polystyrene is a common thermoplastic polymer and has much lower viscosity than PMMA at the same temperature [15]. The T_g of PMMA is 95°C –105°C and the T_g of PS is 110°C. These polymers are adopted for NIL because of their commercial availability, but they may not be the best-suited materials. In recent year, a range of polymers have been developed specifically for nanoimprint, such as the NXR series [16] and mr-I series nanoimprinting polymers [17]. These polymers are tailored to have high resolution capability, wafer-scale uniformity, fast response and excellent flow, low processing temperature and pressure, strong adhesion to substrates, and are also stable at room temperature. They are generally better in pattern transfer with RIE or lift-off.

Thermoplastic polymers can either deform plastically or elastically under heating and pressure. Plastic deformation results in permanent change of shape, which is required for imprinting. If it is elastic deformation, the polymer can recover to its original shape. There is no clear boundary between plastic and elastic deformations. It can be controlled either by the heating temperature or the pressing time. In nanoimprinting of thermoplastic polymers, the recovery can cause the residual layer reshapes after the separation of stamp and sample.

In some cases it may recover almost to the level of the initial layer thickness. The recovery phenomenon is particularly noticeable at low imprint temperature and for large imprinted features [18].

Small imprinted features may be free from the elastic recovery problem. They however cannot avoid the internal stress induced by pressing of stamp. Free-standing nanoscale polymer structures are known to be unstable. There are simulations indicating a strong reduction in the fundamental elastic moduli of the material inside a nanostructure when the width of the nanostructure decreases below 50 nm [19]. The polymer structures at sub-100 nm become very weak. On the other hand, large shear force can exert on the polymer structures during stamp pressing into the polymer. Internal stress can build up in these small polymer structures, weakening their rigidity and leading to structure distortion. The imprinting-induced internal stress can be manifested when the polymer structures are annealed at temperatures near the T_g of the polymer, at which temperature the internal stress can be released, if there are any, but low enough not to cause any visible flow of polymer. Pattern structure decay in the form of reduced height and enlarged base width has been observed at the annealing temperature. One example showed that line patterns of 210-nm high and 140-nm wide in 15 kg mol^{-1} PMMA decayed into 55-nm high and 180-nm wide after annealed 46 min at temperature of $T_g + 15°C$ [20]. The loss in height does not correspond to a proportional increase in width because the loss in height shows up as an increase of the residual layer thickness. The imprinting-induced internal stress is a general phenomenon in thermoplastic polymers. High-molecular-weight polymers are more viscous, and therefore can build up more internal stress. With the internal stress and the elastic recovery mentioned earlier, the pattern structure shape by thermal press imprint can become time-dependent.

5.2.3 Demolding

A high-quality stamp and good compatible polymer imprinting medium cannot guarantee good-quality nanoimprinting pattern structure unless the demolding process is also successful. Demolding is to separate the stamp from imprinted polymer structure after cooling down below the glass transition temperature of polymer. Distortion or damage of imprinted patters can happen if the polymer sticks to the stamp surface, or there is a friction drag due to roughness of stamp sidewalls, or there is a trapping of polymer due to negative slope of stamp sidewalls.

Polymer sticking to stamp is the common problem encountered in NIL because of the large surface contact area and high shear force between polymer and stamp. There are three ways to reduce the surface adhesion: (1) to choose a stamp material which has low surface energy, such as diamond-like carbon (DLC). Diamond-like carbon is known to have low surface energy. Diamond-like carbon film can be patterned to serve as the nanoimprint stamp which is not only of low polymer adhesion but also of high wear resistance [21]; (2) to mix mold release agent with polymer [22]; (3) to coat the surface of stamp with an

antistiction layer. Teflon™ or Teflon-like fluorocarbon or fluorosilane release agent (1H,1H,2H,2H-perfluorodecyl-trichlorosilane) [12] are all able to modify the stamp surface to make the surface hydrophobic and of low energy. Among the three methods, the antistiction coating is most widely used. The coating can be done in liquid phase or vapor phase. However, the antistiction layer becomes less effective as the stamp cavity becomes very small, for example, a few tens of nanometers, where only a few polymer macromolecular chains are in action during imprinting.

The sidewall angle of stamp has critical influence on demolding, hence on the final imprinted polymer structures. Figure 5.7 shows three different profiles of imprint stamp. It is obvious that the stamp profile shown in Fig. 5.7a is preferable because it is much easier to pull out the stamp from imprinted polymer layer. The demolding of stamp shown in Fig. 5.7b will depend on the surface roughness of sidewalls because sidewalls exert high shear force on polymer structure during demolding. The stamp profile shown in Fig. 5.7c is the most undesirable profile. Imprinted polymer structures are in effect trapped inside the stamp structures. Distortion and damage to the imprinted polymer structures are inevitable. Stamps made from RIE or lift-off are mostly of tapered or vertical sidewall profiles and generally do not have problem in demolding as long as the sidewall surface is treated with antistiction agents.

For stamps with vertical sidewall profile, the higher the pattern aspect ratio (height to width) the more difficult it becomes to demold. High-aspect-ratio structures have large sidewall contact surface. There is a high shear stress at the vertical interface between stamp and polymer. During demolding, the stress is concentrated at the base area of polymer pattern, as illustrated in Fig. 5.8a. As demolding takes place at a temperature lower than the polymer T_g, polymer becomes brittle at the low temperature. Pattern fractures can happen around the base area, as the experimental result shown in Fig. 5.8b.

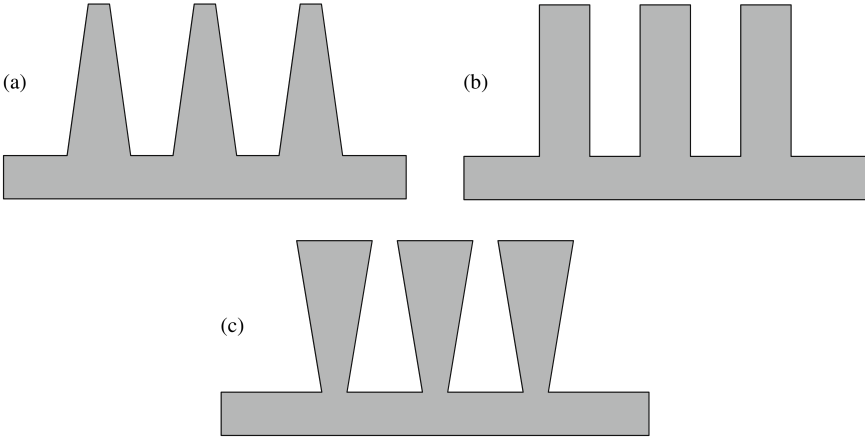


Fig. 5.7 Profiles of nanoimprint stamp: (**a**) tapered sidewall, (**b**) vertical sidewall, and (**c**) undercut sidewall

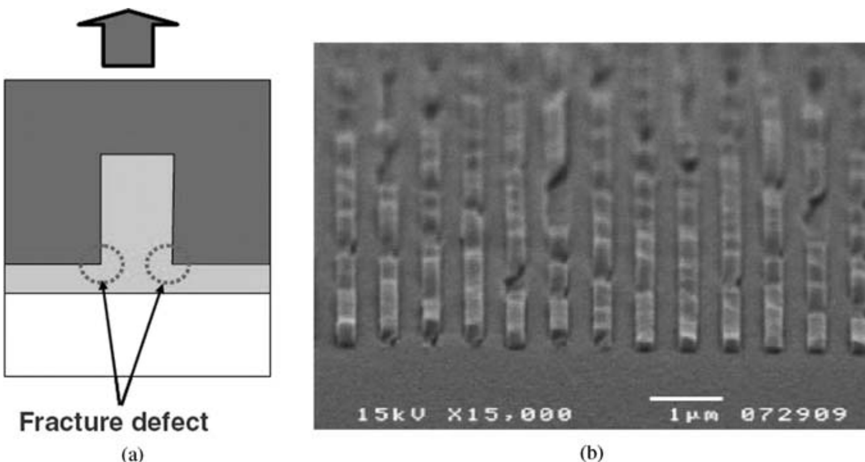


Fig. 5.8 (**a**) Illustration of stress concentration during demolding. (**b**) PMMA patterns after stamp release (Reprint from [23] with permission)

Using low-molecular-weight (M_w) polymer to facilitate imprinting, because of its low viscosity, is generally in conflict with the requirement of safe demolding, because low-M_w polymer is also low in shear modulus. This can become a serious issue when low-M_w polymer is used for high-aspect-ratio nanoimprinting. For the sake of safe demolding, the pattern aspect ratio by nanoimprint is rarely above 3:1. For higher-aspect-ratio nanoimprint, one solution is to use double-layer polymers, as schematically shown in Fig. 5.9. The double-layer polymer system consists of high M_w as the base layer and low M_w as the top layer. As it is shown, when the double-layer polymers squeezed into the stamp cavity, the high-M_w polymer is at the base region. Because of the high shear modulus of high-M_w polymer, it can withstand the high shear force during demolding, keeping the high-aspect-ratio polymer pattern structure intact. Figure 5.10 shows the high-aspect-ratio patterns by imprinting of the double-layer polymer system.

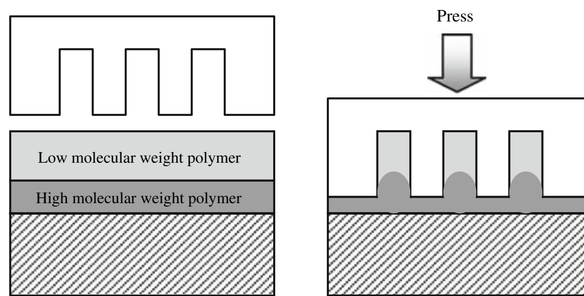


Fig. 5.9 Schematic of double-layer nanoimprinting

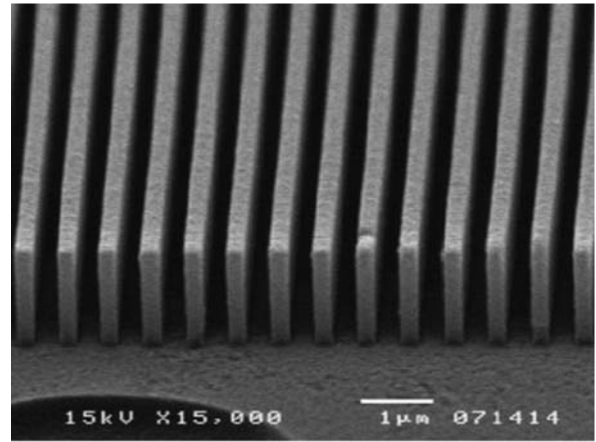


Fig. 5.10 High-aspect-ratio pattern imprinted using double-layered polymer process. *Top layer*: PMMA (M_w = 15 k), 0.7 μm; *base layer*: PMMA (M_w = 996 k), 1.5 μm, 170°C, 150 bars, 250-nm lines; mold *pattern height*: 1.3 μm (Reprint from [23] with permission)

5.2.4 Alignment

Nanoimprinting lithography is superior to other microlithography techniques in single-level patterning. For multilevel patterning, the alignment becomes a big issue. The stamps used in thermal press NIL are generally made of silicon or nickel. They are nontransparent. In order to align the master to an underlying substrate, an open window can be fabricated in the master and the alignment can be carried out through the open window. However, to etch through a silicon wafer to open the alignment window, the etch-through hole has to be a few tens to a hundred micrometers. Alignment through such a large hole cannot achieve high alignment accuracy. There are some other factors which also affect the alignment accuracy.

Another alignment technique is to fabricate alignment marks at the back of a stamp. Before moving the stamp over a substrate, a video camera takes the image of alignment marks on the substrate and stores in a computer. Once the master is placed on top of the substrate, an optical microscope connected to a video camera will take the picture of alignment marks on the back of the stamp and compare them with the substrate alignment marks already stored in the computer. Both alignment mark images are displayed on a computer screen. By moving around the master relative to the substrate, while aligning the marks on computer screen, an accurate alignment can be achieved. In the early version of NIL systems when the NIL apparatus was a simple hot press without dedicated alignment facility, the aforementioned alignment was done in a maskaligner with a special holder. Once the alignment between the stamp and substrate was done, they were clamped mechanically and manually placed into the imprinter stage for imprinting. With extreme care, sub-micrometer alignment accuracy could be achieved over a 4″ substrate [24]. Later versions of nanoimprinters have the built-in online alignment facility [25]. Though the alignment accuracy is still at sub-micrometer level, the alignment process has become automatic.

The third alignment technique for NIL with nontransparent stamp is to insert a miniaturized camera between the stamp and substrate prior to pressing down the stamp [26], as illustrated in Fig. 5.11. In this way, the alignment marks do not have to be on the back of stamp and the alignment is in situ and automatic. The alignment technique has been applied to step-and-stamp imprinting lithography (SSIL). Step-and-stamp imprinting lithography utilizes a small stamp to imprint a large-size substrate [27]. As shown in Fig. 5.11, by moving the stage which holds the substrate the stamp imprints onto the substrate in step-and-repeat fashion similar to an optical stepper. Dynamic alignment is ensured in each step.

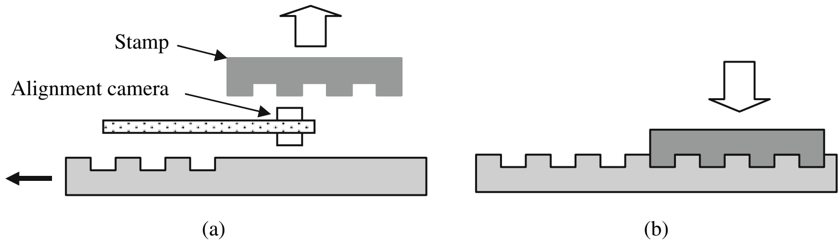


Fig. 5.11 Alignment scheme in step-and-stamp imprint lithography: (**a**) step (**b**) stamp

Apart from the limit in optical imaging resolution during alignment, there are other fundamental limits in the alignment accuracy for thermal press nanoimprint:

1. *Misalignment caused by mechanical pressure*. This may be caused by nonparallelism between top and bottom platens or nonflatness of the substrate such as bowing induced by internal stress from previous processing steps. If such nonparallelism exits, a prealigned position before applying imprinting pressure can be offset once a pressure is applied between top and bottom platens.
2. *Misalignment caused by different thermal expansions between a master and a substrate*. If the master and substrate are made of different materials, their different thermal expansion coefficients can result in misalignment once they are heated to high temperature. For example, an oxide stamp to imprint on a 4″-size silicon substrate can have 100 nm/°C thermal expansion difference at the rim of 4″ silicon wafer. The effect becomes more significant when a silicon stamp is used to imprint on a GaAs substrate [28]. Therefore, for high alignment accuracy it is imperative to avoid high temperature and high pressure.

5.3 Room Temperature Nanoimprint

Thermal press nanoimprint is not always a desirable lithography technique. In some applications, the substrate materials cannot withstand high pressure or may change property at high temperature. For example, cobalt platinum

magnetic thin film is stable only below a certain temperature, InP is very fragile at high uniaxial stress, and not to mention that most of the biological samples cannot survive high temperature and high pressure. From the throughput point of view, high-temperature nanoimprintng platen needs to cool down, which consumes most of the time in an imprinting cycle. As long as thermoplastics are used as the imprinting medium, the requirements for high temperature and high pressure are inevitable. Alterative imprinting medium is the key for the realization of room temperature NIL (RT-NIL).

One of such new polymers is Hydrane HS2550 [29]. This polymer has glass transition temperature at 0°C–10°C and melting transition point at 30°C, which makes the polymer very suitable for low-temperature NIL. The best nanoimprint results were obtained at 50°C and 5-bars pressure. If the imprinting is performed at room temperature, then 30 bars of pressure is needed. Though not exactly a RT-NIL, it is much improved compared to PMMA. At 40-bars pressure, PMMA (950 kg mol^{-1}) requires heating up to 180°C and PMMA (50 kg mol^{-1}) requires 150°C. One of the problems with thermoplastics is the recovery or relaxation, as described in Section 5.2.2. The relaxation is generally more serious when polymer deformation is forced at low temperature. The relaxation effect in Hydrane HS2550 is more significant than in PMMA, which makes it difficult to imprint sub-100 nm dense pattern structures in the polymer. The remedy is to imprint in succession of pressures. By imprinting at increased pressure of 10 bars, 20 bars, and 30 bars with 1-min interval, 75-nm line and space patterns have been achieved [29].

The breakthrough in RT-NIL was the use of HSQ as the imprinting medium. Hydrogen silsesquioxane has been introduced in e-beam lithography. It has the property of silicon dioxide when exposed by electrons or heated up at high temperature. For e-beam lithography of HSQ, the resist is normally pre-baked at 120°C–150°C in order to drive out the solvent. Because HSQ is not a macromolecular material, heating up is not going to make it soft again once the resist layer has been dried. Therefore, pre-baking of HSQ can only make the resist layer hard to imprint. Figure 5.12 shows the imprinted depth in HSQ when the resist has undergone pre-baking at different temperatures [30]. The plots in Fig. 5.12 indicate that it becomes harder to imprint when the HSQ layer becomes drier. Once it is dried at 250°C, the HSQ layer cannot be imprinted at all. This means that the HSQ can be imprinted at room temperature even without any pre-baking. In practice, the HSQ layer is normally given 50°C–100°C pre-baking so that the resist is no longer "wet". The plots in Fig. 5.12 show that there is no significant increase of hardness for pre-bake temperature less than 100°C. After imprinting, a post-baking is normally carried out to further dry up the HSQ layer. Conventional thinking is that post-baking can make the HSQ film more resistant to RIE. The experimental investigation found out that with or without post-baking the HSQ etch rate in CHF_3 RIE has no difference [30], which means the post imprinting baking is not necessary.

Hydrogen silsesquioxane is known to be difficult to spin coat large thickness. In order to make high-aspect-ratio imprinted patterns or to obtain desired resist

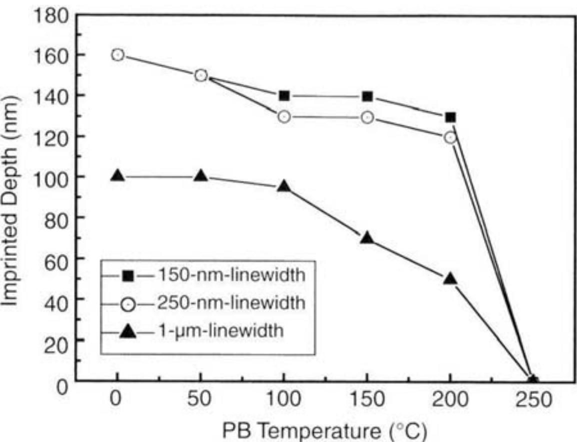


Fig. 5.12 Influence of pre-baking (PB) temperature on imprinted depth in HSQ (imprinting pressure: 220 bars) (Reprint from [30] with permission)

undercut profile for lift-off, multilayer RT-NIL techniques have been developed. In the multilayer resist systems, HSQ is the top layer serving as the imprinting or imaging layer. The bottom layers can be any other resists as long as they can be etched away in RIE. An example of the multilayer nanoimprinting resist system is HSQ on top of PMMA [31]. The nanoimprinting of HSQ/PMMA bilayer resist system and transfer of imprinted pattern into bottom layer by RIE is schematically shown in Fig. 5.13. The bilayer resist system is formed by first spinning 150-nm thick PMMA on a silicon substrate. After baking the PMMA layer in oven at 180°C for 1 h, the HSQ resist (Fox-12 from Dow Corning) was spun on top of the PMMA layer at 9000-rpm spin speed, resulting in a 40-nm thickness of HSQ top layer. The HSQ/PMMA resist stack was then baked on a hotplate at 150°C for 2 min and followed by oven bake of

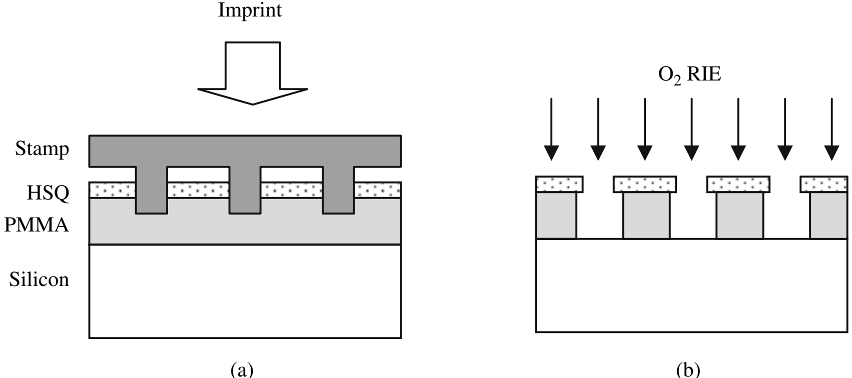


Fig. 5.13 Schematic of HSQ/PMMA bilayer resist for RT-NIL

20 min at 180°C. Nanoimprinting of the bilayer resist was done by a silicon stamp at 440 bars followed by oxygen RIE to transfer the HSQ pattern into the PMMA layer. Minimum 70-nm line features have been achieved. Normally a thin residual layer remains after imprint, which has to be removed by RIE. In the aforementioned process, such a step was not necessary. This may be due to the fact that the HSQ was so thin and had been hard baked, and there was the PMMA layer underneath to cushion the press of stamp. The stamp just broke through the thin HSQ layer into the PMMA resist.

The above bilayer system offers a number of advantages over conventional thermal press nanoimprinting and room temperature nanoimprinting based on single-layer HSQ: (1) feature profile is independent of imprinting. High aspect ratio is achieved in RIE of PMMA layer; (2) the high etch resistance of HSQ to oxygen plasma allows very thin HSQ layer to be used as the imprinting layer; (3) no need to remove residual HSQ after imprinting, resulting in better feature fidelity; (4) the underneath PMMA can act as a buffer layer to provide a soft landing for silicon stamp, which reduces the chance of damaging the stamp; (5) no intermediate layer is needed, because there is no intermixing between the HSQ and PMMA; (6) there is no need to coat release agent on silicon stamp because only 40-nm thick HSQ needs to penetrate through. The stamp can be repeatedly used without additional cleaning process.

Other polymers have also been used as the bottom layer, such as photoresists or poly(vinel alcohol) [30] or spin on glass (SOG) [32] or polyelectrolyte multilayer films [33]. The imprinting and pattern transfer processes are similar to what is illustrated in Fig. 5.13.

5.4 UV-Cured Nanoimprint

Although nanoimprinting can be performed at room temperature as described in the previous section, it is still not a low-pressure process. A truly room temperature and low-pressure process is the UV-cured NIL (UV-NIL). Figure 5.14 schematically illustrates the basic steps of UV-NIL. It is almost the same as the thermal press NIL shown in Fig. 5.1. The key differences are the use of transparent stamp and UV-curable polymer. The UV-curable polymers have viscosity ranging from 50 mPa s to 200 mPa s in liquid form at room temperature. Little pressure ($<$1 bar) is needed to press the stamp into the liquid polymer. The solidification of liquid polymer is done by UV illumination through the transparent stamp. UV-NIL was first proposed in mid-1990s [34]. The technology has since gained more and more popularity as it has solved many problems associated with the thermal press NIL.

5.4.1 Transparent Stamps

The obvious choice of substrate for transparent stamps is the photomask blank plate. The use of photomask plate for making UV-NIL stamps has three advantages:

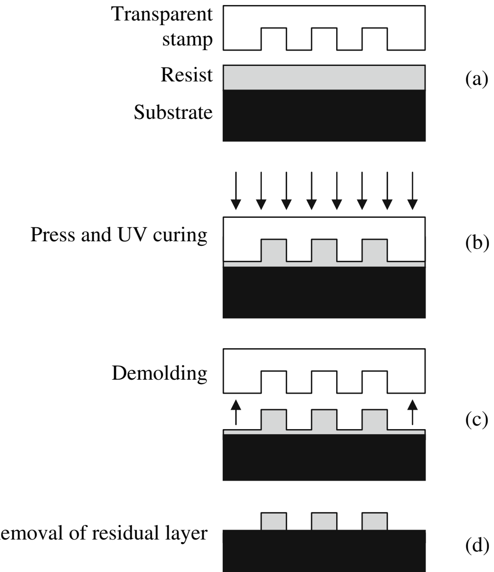


Fig. 5.14 Schematic of UV-NIL: (**a**) Basic components, (**b**) UV exposure to cure the polymer, (**c**) demolding, and (**d**) RIE to remove residual polymer

1. The quartz (fused-silica) substrate of a photomask has excellent transparency to UV wavelength.
2. There is already a chrome layer coated on the blank plate. The chrome layer serves both as the masking layer for quartz etch and as a conductive layer during e-beam exposure patterning.
3. There are very mature process technologies for patterning photomasks in mainstream IC manufacturing, particularly the experience in patterning phase-shifting masks (PSMs) which involves high-resolution e-beam lithography and quartz etch and can be directly used for making UV-NIL stamps.

However, UV-NIL stamps have much stringent requirements on both chrome layer patterning and quartz etching. Conventional photomasks, including masks with resolution enhancement features, normally have features above 100 nm because of the 4:1 reduction of optical lithography. Nanoimprinting is 1:1 replication. To imprint 10-nm features, the stamp feature has to be of 10-nm dimension. There are several issues on patterning sub-100 nm UV-NIL stamps in quartz substrate.

First, the chrome layer has to be etched by RIE. At sub-100 nm resolution, only a few e-beam resists are suitable. Those high-resolution e-beam resists may not have the high resistance to RIE of chrome. The typical example is PMMA which is the standard high-resolution e-beam resist, but very poor as a masking layer in RIE. Hydrogen silsesquioxane resist is good in RIE but low in e-beam exposure sensitivity. Chemically amplified resist such as NEB-22 has high

sensitivity but is poor in dry etch resistance. An experimental result revealed that the selectivity of NEB-22 to chrome etching is 5:1; that is, if the etch rate for chrome is 30 nm min^{-1} it would be 160 nm min^{-1} for NEB-22 resist [35]. A standard photomask blank has a chrome layer thickness around 100 nm in order to have adequate opaqueness, which is too thick and requires much thicker resist layer as the etching mask. Bearing in mind that the chrome layer in UV-NIL stamp fabrication only provides a mask for quartz etch, not to block out the light, while the selectivity of chrome to SiO_2 in RIE for pattern transfer into the quartz substrate is 20:1, the chrome layer for UV-NIL stamp fabrication, therefore, does not have to be as thick as that in a standard photomask blank. The thickness of chrome film for UV-NIL stamp fabrication is normally 10–20 nm.

Second, dry etch of quartz is known to be much more difficult than dry etch of silicon. The quartz etch rate in a conventional RIE system is only 30–50 nm min^{-1}. Even with a high power density induction coupling plasma (ICP) etcher the etch rate is about 150–350 nm min^{-1}. Two problems may be encountered in quartz etch, which are microtrenching and microloading. Microtrenching is the vertical overetch that occurs at the foot of sidewalls of a pattern profile. It is caused by local accumulation of ions reflected from sloped sidewall, which enhances the local etch rate at the bottom [36]. Microloading is a common problem in RIE. It is a phenomenon in which larger features are etched faster than smaller features [37]. The two phenomena will be discussed in detail in Section 6.3.6 of Chapter 6. The microtrenching can occur because quartz is an insulating material and charge can easily accumulate at the etched sidewalls. The microloading becomes prominent normally in deep RIE where the large etch depth is required. For stamp fabrication, the etch depth is less than a micrometer. Unless extremely small features coexist with very large features, the microloading effect is usually negligible [38]. Other issues such as sidewall angle and surface roughness also exist. Optimization of etching conditions is necessary for each particular stamp design in order to achieve the best quality of fabrication.

Quartz stamps are rigid, which is not always desirable, particularly when large area of UV-NIL is needed. For large-size substrate, the flatness and waviness is difficult to control within a few micrometers. On the other hand, the structure height of UV-NIL stamp is not more than a few hundreds nanometers. For this reason, the quartz-based stamps are never made in large size. They are usually a few millimeters or a few centimeters square sizes. UV-NIL of large-size substrate is performed in a step-and-repeat manner. For single imprinting of large-size substrate, which is often needed in nanofluidic or nanophotonic applications, flexible stamps are preferred. There are a number of options to make a flexible stamp. Polydimethylsiloxane (PDMS, Dow Corning, Sylgard 184) is a popular choice of stamp material. Polydimethylsiloxane is highly transparent for wavelengths between mid-UV and visible wavelength. Polydimethylsiloxane can be cast into any shape, for example, casting PDMS into a silicon master stamp. Minimum 280-nm imprinted features have been demonstrated with a 4″-size PDMS stamp and UV-NIL at room temperature and pressure of 500 mbar [39]. One must be aware that NIL with flexible stamps

cannot produce the same resolution as rigid stamps because of the distortion effect in soft stamp. While a quartz stamp can produce less than 20-nm imprinted features in low-viscosity resists [40], the soft stamp can only imprint features above 50 nm [41]. To improve the rigidness of stamp while still keeping its flexibility, some variations of PDMS stamps have been proposed. The PDMS stamp can be stuck to a glass back plate to increase its rigidness, or the pattern structures are made into a thin layer of hard polymer (three times harder than PDMS) or a thin glass layer [42] as shown in Fig. 5.15. In this case, the PDMS is sandwiched between two rigid plates. The stamp layer in direct contact with imprinting resist is a hard material. Minimum 23-nm features have been imprinted with the trilayer flexible stamp on a 4″-size substrate [42].

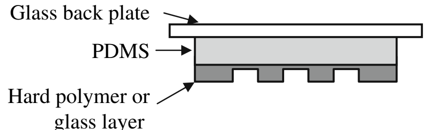


Fig. 5.15 Layer structure of flexible UV-NIL stamp

Quartz stamp is expensive in terms of both the material and the stamp fabrication process. The alternative is to fabricate stamps in polymers, as long as they are transparent to UV wavelength. The aforementioned PDMS is one of the alternatives. Other polymers have been investigated, such as a fluororesin material [43]. The fluororesin has >95% transparency to UV light down to 200-nm wavelength and is a hard material with tensile modulus of 1.2 GPa. The stamp is made by thermal press imprinting with a silicon master stamp. The added benefit of this fluororesin is its natural hydrophobic surface. No anti-stiction coating of stamp is needed. A similar concept has been applied to produce the Intermediate Polymer Stamp (IPS^{TM}) [44]. The polymer stamps can be produced in large quantities just like copying CDs. As a result, the stamp becomes so cheap that it can be for single use and thrown away. In this way, the master stamp can last much longer. A distinctive benefit of the IPS is its "self-cleaning" process. If there are any particulate defects on the master stamp, the IPS copying process can remove these defects. As seen in Fig. 5.16, particles are simply embedded into the polymer stamp and removed during copying the polymer stamp. They would not pose any risk in the next round of copying [44].

UV-NIL does not have to use a quartz stamp. If the UV-curable polymer is coated onto a quartz substrate or any transparent substrate, a normal silicon stamp can be used and the UV curing is done by illuminating through the transparent substrate, as schematically shown in Fig. 5.17. A thermal press NIL of SU-8 photoresist, combined with UV illumination from backside, has been able to produce high-aspect-ratio imprinted dense lines, as the NIL pattern shown in Fig. 5.18. Heating above the T_g of SU-8 promotes a flowable state of SU-8, which helps to fill the dense and deep cavity in silicon stamp. The extra UV curing solidifies the SU-8 and forms the final NIL pattern [45].

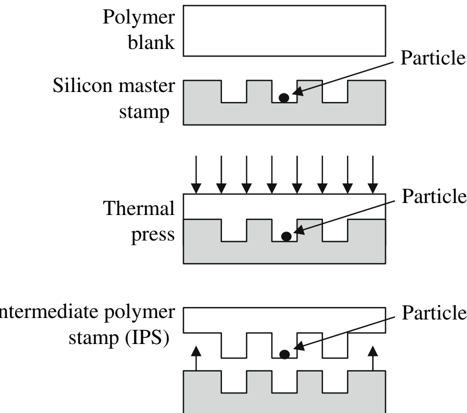


Fig. 5.16 Self-cleaning process in IPS copying process

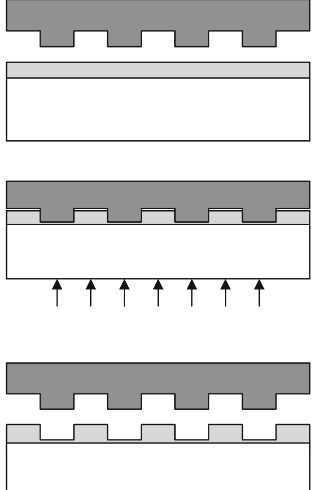


Fig. 5.17 Schematic of UV curing of polymer through a transparent substrate

5.4.2 UV-Curable Polymers

The key enabler of UV-NIL is the UV-curable polymer. UV-curable polymers have existed for long time as standard polymer products. There are UV glues and UV dental sealants. They all rely on photopolymerization process to turn from liquid or paste state into solid state upon UV light irradiation. For the UV-curable polymers to be used in UV-NIL, there are a number of special required properties. The most important properties are as follows:

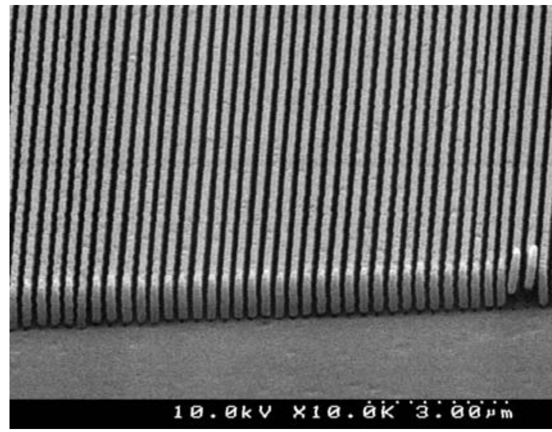


Fig. 5.18 High-aspect-ratio dense lines made in SU-8 photoresist by combining thermal press with UV curing nanoimprint

- low viscosity so that it can fill in the stamp cavity quickly and completely;
- high resolution to replicate sub-100 nm patterns;
- fast UV curing to reduce the cycle time of UV-NIL;
- low shrinkage to ensure pattern fidelity of replication;
- high dry etch selectivity to ensure reliable pattern transfer into substrate materials.

These desirable properties can be obtained by appropriate polymer synthesis. Polymer synthesis is like a cooking process. Basically, different ingredients are added together so that the final mixture possesses flavors of each ingredient in various extents. To meet the above requirements, the UV-curable polymers contain vinyl ethers which are the key ingredient for photoploymerization, organic acrylates to provide the low viscosity of polymer mixture, organic crosslinker for thermal stability and mechanical strength, and photoinitiator which dissociates upon UV irradiation to form radicals which promote the polymerization process [46]. Some polymers may have additives such as silicon-containing acrylate monomers to enhance the dry etch selectivity or fluorinated compounds to facilitate mold release. For example, it is found that a fluorine-based additive (1H, 1H, 2H, 2H, perfluorooctyl triethoxysilane) can migrate to the surface of polymer film during spin coating, which can lower the surface energy and create an antistiction surface to make demolding of stamp easier [47].

The viscosity of UV-curable polymers is preferably below 15 mPa s. In practice it is normally a few tens or a few hundreds of mPa s, which is considerably lower than those used in thermal press nanoimprint which is at least in the order of 10^4 Pa. s [15]. Low viscosity can make the polymer filling of stamp cavity faster. Imprinting pressure can be reduced and residual layer is thinner. The polymer molecule size poses no limit on its high resolution capability as the molecular chains are only 1–2 nm. The resolution of UV-NIL is primarily limited

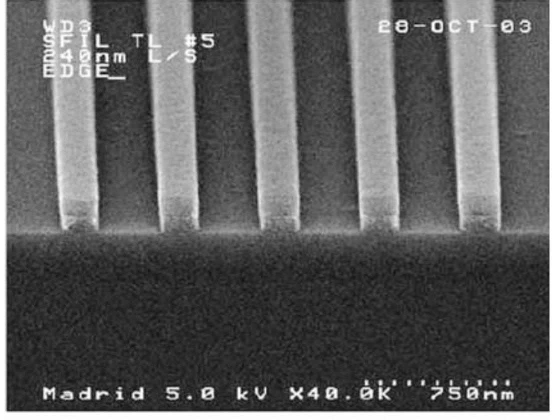
by the structure dimension of stamps. The UV curing time depends on the sensitivity and concentration of photoinitiator. Normally, the required energy for curing the polymers is in the range of a few hundreds of mJ cm^{-2}. The curing time is a few seconds to a minute depending on the power of UV lamp.

One of the critical issues for UV-curable polymers is the shrinkage. UV-curable polymers, once it has been solidified, have typically 5–15% volume shrinkage [34]. The shrinkage is very much dependent on the molecular structures. It is possible to reduce the shrinkage to 3–6% by selecting high molecular weight of polymer components [48]. One thing should be noted that the polymer shrinkage is usually measured on a planar polymer film under uniform pressure and UV irradiation. The measured shrinkage in this way may not be equal to actual shrinkage in replicated pattern structures. For example, the measured shrinkage for planar film is 9% whereas the measured dimensional variation for 100-nm line features shows no change [49]. One possible explanation is that those small features are under higher-than-uniform planar pressure during UV curing. The UV curing under higher pressure may result in a lower shrinkage.

The acrylate-based UV-curable polymers have not only high UV sensitivity but are also sensitive to ambient oxygen. Oxygen molecules dissolved in the polymer can scavenge free radical species, and thus inhibit the polymerization process at the polymer surface at the onset of exposure. This can prolong the required exposure time. For example, a UV-curable polymer has minimum UV exposure dose of 700 mJ cm^{-2} in vacuum, but needs 1.4 J cm^{-2} when exposed in air [48]. A different UV-curable polymer system based on cationic crosslinking of cycloaliphatic epoxies was developed to overcome this problem, because cationic polymerization is not prone to oxygen inhibition, as compared to the free radical polymerization of acrylate monomers. The polymer also exhibits a very good dry etching resistance because of the high silicon content. Furthermore, it has very low shrinkage after curing (only a fraction of that of the acrylate system) [50].

The replicated polymer structures have to be transferred to substrate. There are two stages of RIE involved. First is the etch removal of residual polymer film which is normally less than 100 nm. Second is the RIE of substrate material using the replicate polymer pattern as the mask. For removal of residual film, oxygen plasma is usually used. For pattern transfer to substrate, different etching gases are used depending on the etching chemistry of substrate material. As mentioned earlier, polymers with silicon-containing components can have high resistance in RIE. Reducing the inorganic content may result in much higher etching rate during removal of residual film or pattern transfer that requires oxygen plasma. Figure 5.19 compares the pattern structures of conventional silicon-containing photoresist (SiBER from Shipley containing 19% silicon by weight) and a UV-curable polymer (containing 9% silicon) etched in oxygen plasma. The polymer with low content of silicon fared poorly in oxygen plasma etching. For this particular polymer, a completely different etching chemistry with NH_3 plasma has to be used, which exhibits selectivity of 5:1 to the etching of underneath antireflection coating. High content of silicon may be good in pattern transfer, but makes the removal of the residual film difficult. For example, a UV-curable

Fig. 5.19 Pattern profiles etched in oxygen plasma: (**a**) SiBER photoresist from Shipley. (**b**) A UV-curable polymer (Reprint from [51] with permission)



(a)

(b)

polymer containing inorganic components has dry etch rate of only 10 nm min^{-1} in pure oxygen plasma. An amount of 5% of CF_4 gas has to be added so that the plasma with O_2/CF_4 gas mixture can increase the polymer etch rate to 80 nm min^{-1} [47]. The advantage of pure organic polymer is the ease in removal of residual layer and especially in cleaning of the mold stamp if there are any residues of polymer sticking to the mold stamp. By tuning the polymer components, it is possible to formulate pure organic UV-curable polymer while maintaining etch selectivity to silicon at around 2.5 in C_4F_8/SF_6 plasma [48].

5.4.3 Step-and-Flash Imprint Lithography

One of the NIL technologies, which perhaps has the most publicity and has been hailed as one of the most promising candidates to replace optical

lithography for IC manufacturing, is the step-and-flash imprint lithography (SFIL). The SFIL concept was first proposed in 1999 [52] and it is essentially a double-layer UV-NIL process, as schematically shown in Fig. 5.20. The original UV-NIL, first proposed in 1996 [34], could only work with a thin layer, low viscosity, and UV-curable polymer film. Such a thin polymer film made it difficult for high-aspect-ratio pattern transfer into substrate material. In the SFIL process, a polymer resist, called transfer layer, is coated onto a silicon substrate first. Then a drop of low viscosity and UV-curable silicon-containing polymer liquid, called etch barrier, is dispensed on top of the transfer layer. A transparent stamp coated with release layer is brought close to the transfer layer surface. The liquid polymer is then filled the stamp cavities by capillary action. UV irradiation through the transparent stamp cures the liquid polymer and leaves an imprinted pattern of the stamp. A subsequent RIE transfers the imprinted pattern into the transfer layer which serves as the etch mask for further pattern transfer into the substrate. Same as UV-NIL, it is a room-temperature and low pressure imprinting technology. Similar to the SSIL as shown in Fig. 5.11, the SFIL imprints large size substrates by step-and-repeat of the stamp. Unlike the SSIL which stamps over a substrate already coated with a polymer layer, the liquid drop of etch barrier in SFIL is only dispensed at the location where the stamp is about to make an imprint (drop-on-demand).

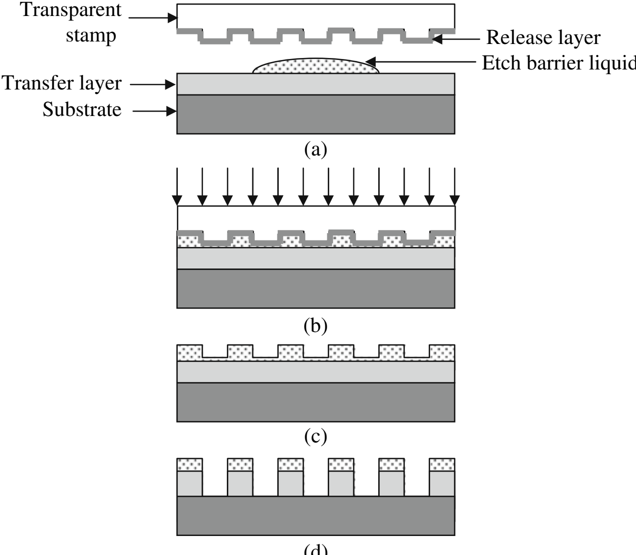


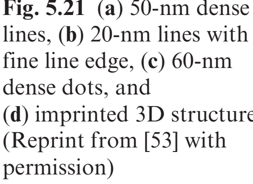
Fig. 5.20 Schematic of SFIL process: (**a**) layer structure, (**b**) imprinting and UV curing, (**c**) demolding, (**d**) RIE into transfer layer

When comparing the SFIL process with the conventional UV-NIL shown in Fig. 5.14, the key difference is the use of thin liquid polymer layer as the etch barrier. The etch barrier material is subject to a number of constraints [53]. It must be dispensable from an automatic fluid dispensing system. As low-viscosity liquid tends to have high vapor pressure, evaporation may change the composition of the liquid material during the course of multiple imprints. The liquid etch barrier must not change significantly in composition between dispensing and imprinting. The liquid must be readily displaced and the cavity in stamp must be filled completely and quickly during the imprinting step. The photopolymerization process should be rapid during exposure. Shrinkage as a result of polymerization must be controlled. The polymer must not stick to the stamp surface and yet it must adhere well to the polymer transfer layer. It must exhibit sufficient rigidity to avoid feature collapse during separation of stamp. It must exhibit some level of temperature stability to withstand the etching temperature, and it must have sufficient etch selectivity during the RIE step to allow high aspect ratios to be generated in the transfer layer.

To meet all the above requirements, many combinations of polymer formulations were investigated. Some of the requirements are conflicting. Therefore, there have to be tradeoffs. For example, the liquid has to be hydrophobic to the stamp but hydrophilic to the transfer layer. The liquid has to be easy to flow, which requires low shear modulus, but mechanically strong to withstand demolding after UV cure. The low viscosity is also in conflict with the UV curing speed, because the viscosity of liquid solution is directly proportional to the oligomers concentration which is responsible for UV curing rate [52]. Proprietary polymer liquids (MonoMat^{TM}) have been developed to meet the above requirements [54].

Step-and-flash imprint lithography, since proposed, has been poised to be a candidate of NGL technology after optical lithography. Step-and-flash imprint lithography is the nearest analogue to optical lithography. It is a parallel process, hence potentially of high throughput. The stamp (template) is transparent, which allows high-precision alignment. There are other advantages that the optical lithography does not have [55]:

- Because no light wave diffraction is involved during imprinting, the only limit in resolution is the resolution of stamp structures, which at current level of e-beam lithography, as described in Chapter 3, is about 5 nm. Without light diffraction, all the optical proximity effects mentioned in Chapter 2 do not exist. Therefore, pattern in the stamp is straight design pattern without any needs of complicated optical proximity correction (OPC) features.
- The line edge roughness (LER), which has become a big problem in optical lithography at sub-50 nm resolution, does not exist in SFIL.
- The template can be made into three-dimensional (3D) shape through multi-level or grey scale lithography process. Replication can be done no matter



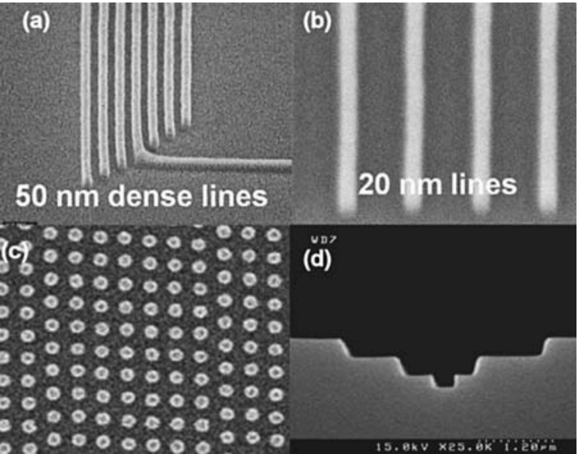


Fig. 5.21 (**a**) 50-nm dense lines, (**b**) 20-nm lines with fine line edge, (**c**) 60-nm dense dots, and (**d**) imprinted 3D structure (Reprint from [53] with permission)

what the shape of template is. Three-dimensional pattern structure can be made easily in one step of imprint. Figure 5.21 shows some scanning electron microscope (SEM) pictures of patterns produced by SFIL [53].

- There is no complex optical imaging system such as the 193 immersion optical lithography or light source such as in EUV lithography. The cost and cost of ownership (CoO) for SFIL is only half of current optical lithography [56].

Although SFIL has many advantages over optical lithography, IC manufacturers have been reluctant to accept it into production lines because of the inertia of using the existing technology. With the optical lithography continuously pushing the resolution limit, and 32-nm IC manufacturing has already been demonstrated using 193 nm optical lithography, SFIL has to aim at entering mainstream IC manufacturing at 22-nm node and below.

Although thermal NIL, UV-NIL, and SFIL have all demonstrated around 10-nm resolution patterning capability, it is a completely different game when moving the technology from laboratory to a volume manufacturing environment. To be accepted as a volume manufacturing technology, SFIL has to demonstrate major progresses in four areas: reliable fabrication of high-resolution templates, overlay accuracy, minimum defects level and throughput comparable to current optical lithography. These are not easy tasks when the minimum imprinted features are at 32 nm and below. However, the difficulties are mostly engineering rather than fundamental. There is a great hope for SFIL to enter IC manufacturing at 22-nm technology node. In addition, SFIL can have a major role to play in other application areas, such as fabrication of magnetic storage medium and photonic array devices. These applications require high-density periodic features at sub-100 nm dimension, which are most suitable to pattern by NIL.

5.4.4 Alignment through Transparent Stamps

Apart from room temperature and low pressure processes, another key advantage of UV-NIL is high alignment accuracy. As described in Section 5.2.4, thermal press imprint only has alignment accuracy of submicron because the stamp is opaque. Now UV-NIL uses transparent stamps. Alignment marks can be seen through the stamp. The alignment scheme of UV-NIL is similar to that of contact optical lithography in a maskaligner. However, it potentially has much higher alignment accuracy than either contact or projection optical lithography. Figure 5.22 schematically compares the alignment schemes for projection and contact optical lithography with that of SFIL. In an optical projection system (Fig. 5.22a), alignment is performed in several steps: (1) align the mask relative to the mask stage; (2) align the wafer relative to the wafer stage, and (3) align the mask and the wafer relative to each other using multiple reference mirrors. Thermal, vibrational, and mechanical noises can all contribute to

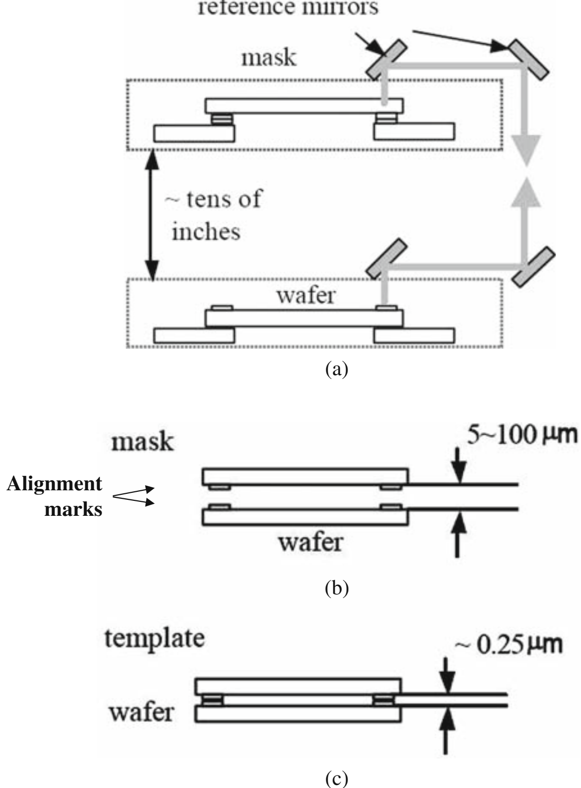


Fig. 5.22 Alignment schemes: (**a**) projection optical lithography, (**b**) contact optical lithography, and (**c**) SFIL (Reprint from [57] with permission)

alignment errors. In a contact optical printing system (Fig. 5.22b), mask and substrate have to keep a large gap in order for the mask to have free movement during alignment. The large gap deteriorates the imaging of alignment mark on the substrate. UV-NIL, in particular the SFIL, can allow much small gap (<0.25 μm) between the stamp and the substrate because of the much small size of stamp (∼ a square inch). There is a thin layer of liquid between the gap, which acts both as lubricant and damping medium. Imaging of alignment mark through such small gap is much clear and sharp using a conventional optical microscope.

However, conventional box-in-box or cross-in-box alignment marks do not have enough accuracy for sub-100 nm imprinting because of the resolution limit of optical microscope. A much high-resolution alignment can be achieved through interferometric alignment [58]. The interferometric alignment relies on gratings at both the stamp side and the substrate side to generate an interference pattern, called Moiré patterns. The Moiré pattern is formed when two grids are overlaid at an angle or have slightly different periods. If the grating period on substrate is P_1 and that on stamp is P_2, then the period length of Moiré pattern can be calculated by Eq. (5.2) [59]:

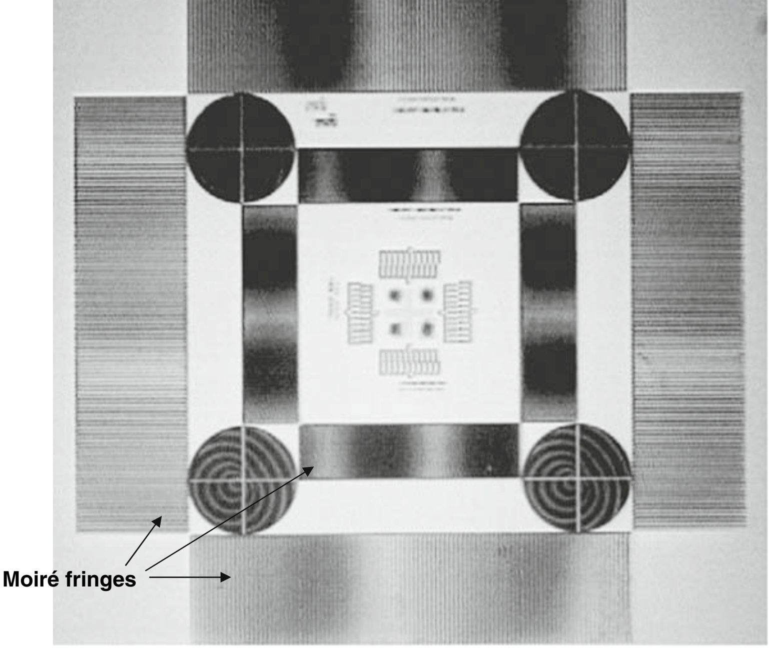


Fig. 5.23 Video image taken during alignment, showing interference dark fringes of Moiré pattern (Reprint from [59] with permission)

$$P_M = \frac{P_2}{1 - P_2/P_1}. \quad (5.2)$$

The interference pattern has much larger period length than the alignment gratings on both the stamp and the substrate. For example, if $P_2 = 1.96$ μm and $P_1 = 2$ μm, then the Moiré period length is 98 μm. Figure 5.23 is a video image taken during alignment of stamp and substrate. The formation of Moiré pattern is seen as dark fringes. If there is a misalignment Δ between stamp and substrate, then the Moiré period will shift a distance of $d = \Delta(P_M/P_2)$. A 150-nm misalignment between stamp and substrate will show up in Moiré pattern by a 7.5-μm shift, which is much easy to see from an optical microscope. Using this technique, sub-20 nm alignment accuracy has been achieved [60]. Theoretically, the method has the ultimate alignment accuracy of $\sim$0.5 nm [58]. Although Morié fringes can achieve sub-100 nm alignment accuracy, it is normally used in fine-tuning step. Conventional box-in-box alignment technique is still used at coarse alignment step.

5.4.5 Combined Nanoimprint and Photolithography

UV-NIL provides another advantage, that is, the polymer can be washed away if the UV irradiation is blocked from curing or exposure action on it, which is the basis of the so-called combined nanoimprint and photolithography (CNP) [61]. The CNP scheme is best explained in Fig. 5.24. It uses a hybrid stamp which has both the etched cavities for imprinting and opaque patterns, made by metal deposition and patterning, for blocking the UV light. The polymer resist can be a UV-curable polymer or a negative-tone photoresist. The imprinting process is the same as aforementioned UV-NIL process. When the hybrid stamp is pressed into the polymer resist and UV light shines through the transparent quartz stamp, part of the UV irradiation is blocked by the opaque chrome features on the stamp. After lifting off the stamp, the substrate is dipped into a developer solution. In those areas without UV irradiation, the polymer resist will be washed away, leaving both the imprinted features and the developed features. It is apparent that if a photoresist is used the photoresist has to be negative tone, so that the unexposed photoresist can be dissolved in a developer. One of the best negative-tone photoresists for this purpose is SU-8.

The reason for developing this hybrid lithography approach was to enable patterning both small and large features in one step. As mentioned in Section 5.2.2, the biggest difficulty for NIL is to pattern features with large dimensional differences. While small feature cavities in a stamp can be filled quickly, large features require displacement of more volume of polymers and longer flow distance. It is difficult to displace large amount of polymer, and defects are often created due to partial filling of large stamp

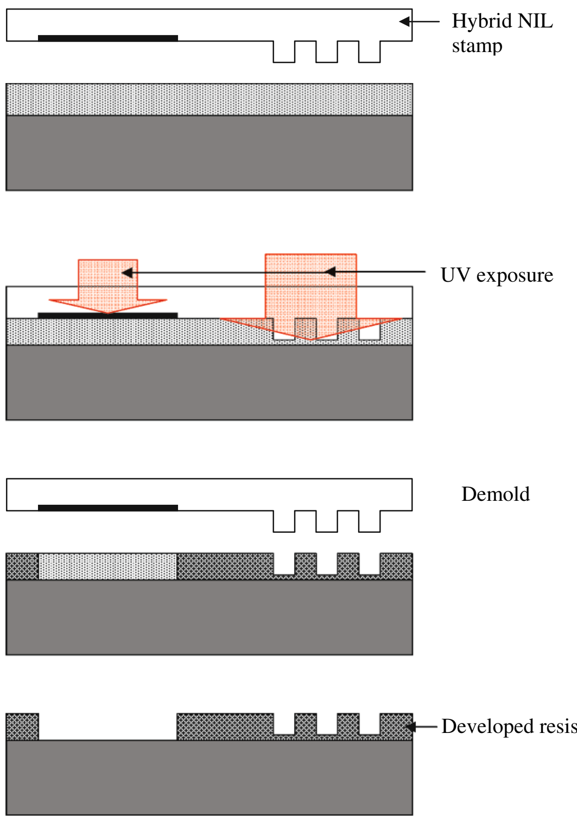


Fig. 5.24 Schematic of combined nanoimprint and photolithography (CNP)

cavities. The CNP has nicely solved the problem. Both 350-nm dense lines and 200-µm square patterns have been made in SU-8 by the CNP technique [61]. Other benefits the CNP can bring include reduced imprinting pressure because only small features need to be imprinted, which generally requires lower pressure and filling of stamp cavities is quicker; the residual layer thickness becomes more uniformly distributed, as larger imprinted features tend to have thicker residual layer due to uneven distribution of pressure in large and small imprinting features. In general, the imprinting time can be reduced because distribution and displacement of polymer only take place at small stamp cavities. The metal opaque features can be used as alignment marks which gives much higher image contrast than polymer or quartz alignment marks [50].

Based on the same concept, the opaque layer can be made directly onto the protrusion features of a stamp, as shown in Fig. 5.25 [62]. When imprinting is carried out and UV light shines through the stamp, the UV irradiation is blocked at the protrusion features. No curing of polymer takes place at these

areas. The subsequent removal of residual layer can be done simply by wet development in a developer solution without RIE. In this way, it is also possible to produce high-aspect-ratio imprinted patterns because the residual layer does not have to be very thin. The wet development only dissolves the unexposed polymer and does not affect the imprinted pattern structures. In contrast, conventional UV-NIL has to keep a thin residual layer, because removal of residual layer by RIE is indiscriminative, whereby RIE will etch both the residual layer and the imprinted polymer structures.

However, the CNP needs more complicated stamp fabrication process. For the approach illustrated in Fig. 5.24, the opaque patterns are made by an extra photolithography and metal lift-off after RIE of nanoimprinting patterns. In order to ensure the metal patterns are opaque, the metal layer has to have enough thickness. For example, conventional photomasks have a chrome layer at least 100-nm thick to be opaque to UV illumination. In order that the thick chrome patterns do not protrude out like an imprinting structure, the area of opaque patterns has to etch into quartz by RIE to form a shallow cavity. The metal coating, either chrome or nickel, is then filled in the cavity to form the metal opaque pattern [61]. For the approach illustrated in Fig. 5.25, thick metal layer is used as the etch mask to produce the

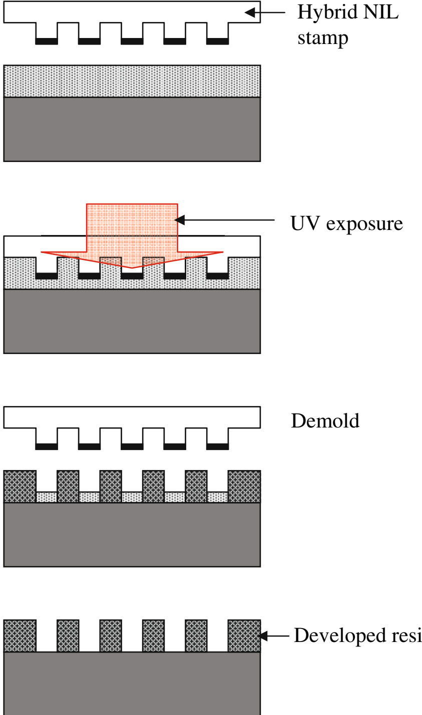


Fig. 5.25 CNP to facilitate removal of residual polymer layer

quartz protruding features. However, if such thick metal layer is used, high-resolution nanoimprints, especially below 100 nm, are difficult to achieve, because sub-100 nm stamp features are very difficult to make with such thick metal layer as quartz etching mask, as pointed out in Section 5.4.1.

5.5 Reverse Nanoimprint

So far the nanoimprint processes introduced are all using a structured stamp to press into a polymer-coated flat substrate. The imprinting process can also be reversed, that is, to have structured polymer sheet first, and then transfer it to a flat substrate [63]. The so-called reverse nanoimprint process is schematically illustrated in Fig. 5.26. Instead of spin-coating polymer onto a substrate, the polymer is spin-coated onto a structured stamp to fill up all the stamp cavities (Fig. 5.26a). A short time soft bake is normally carried out to evaporate solvent in the polymer. The molded polymer structure can be transferred to a flat substrate either by thermal press (Fig. 5.26b1) if the polymer is PS or PMMA [61], or by UV curing (Fig. 5.26b2) if the polymer is UV curable, such as negative-tone photoresists, and the mold is transparent [64]. The UV curing can also be performed from substrate side if the substrate is transparent and the mold is opaque [65]. A pressure is needed to achieve intimate contact between the polymer and substrate surface. After demolding, the polymer structures molded from the stamp is transferred onto the substrate (Fig. 5.26c), which is equivalent to that obtained by normal imprint.

The key condition to the success of reverse nanoimprint is to make sure that the surface energy of substrate is greater than that of the mold, so that the structured polymer can easily come off the mold but firmly stick to the

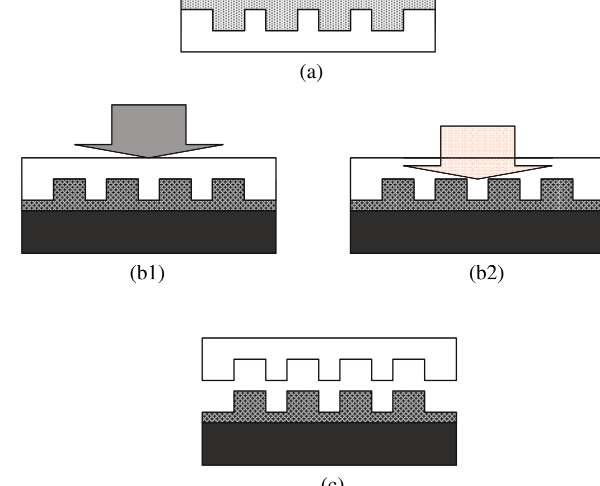


Fig. 5.26 Reverse nanoimprinting process: (**a**) spin-coating stamp with polymer, (**b1**) reverse imprinting by thermal press, (**b2**) reverse imprinting by UV curing, and (**c**) demolding

substrate. Both the mold surface and substrate surface need to be treated. There are different ways to alter the surface energy of mold and substrate, depending on the material property. For example, a quartz stamp for transferring SU-8 polymer structure to a silicon substrate can be treated with phenethyltrichlorosilane (PETS). The quartz stamp is first cleaned in $H_2SO_4 + H_2O_2$ (1:1) for 10 min, and then exposed to an oxygen plasma at 250 mTorr and 200 W for 3 min. After cleaning, the stamp is immersed in 2% PETS in n-heptane for 5 min and baked at 100°C for 10 min. After the treatment, the surface energy of the quartz stamp is 22 mJ m^{-2}, compared to 51 mJ m^{-2} without the treatment [66]. For the Si substrate, it has been found that the surface energy increases from 47 mJ m^{-2} to 73 mJ m^{-2} after oxygen plasma treatment, which is considerably higher than that of the quartz surface. The transfer of polymer imprint from quartz stamp to silicon substrate is ensured.

Because polymer is poured into mold cavities, instead of the mold pressing into polymer, high-aspect-ratio replication can be more easily achieved than conventional thermal press nanoimprint. Figure 5.27 shows replicated PMMA pillars with 17.5:1 aspect ratio by pouring the PMMA into a mold [67]. There are two difficulties facing the high-aspect-ratio replication: one is how to get the polymer into such deep and narrow cavities and second is how to pull the high-aspect-ratio polymer structure out of the mold. The first difficulty can be overcome by using a polymer liquid with very low surface tension. Polymers diluted in solvent such as toluene or chloroform are suitable for this purpose. For example, PMMA powder is dissolved in toluene in 1:40 weight ratio. A further vacuum baking can be performed to ensure the there are no air bubbles trapped inside the deep cavities. The demolding is less of a problem in reverse nanoimprint than in thermal press nanoimprint, because polymer molecules get into the mold cavity purely by gravity and capillary action. There is no external force applied to squeeze the polymer into cavity. Therefore, there is no shear stress created in the molded polymer structures. With proper antistiction

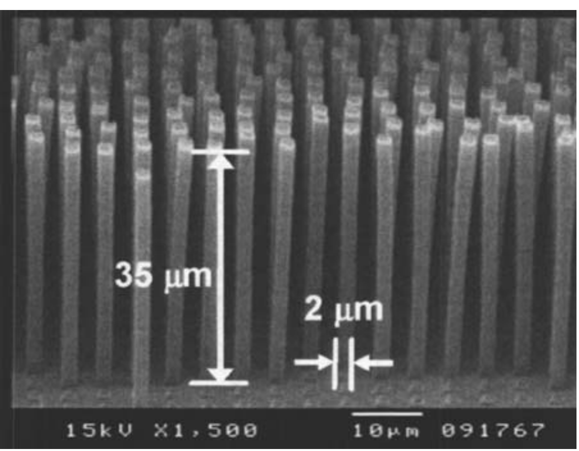


Fig. 5.27 High-aspect-ratio PMMA pillars replicated by cast molding (Reprint from [67] with permission)

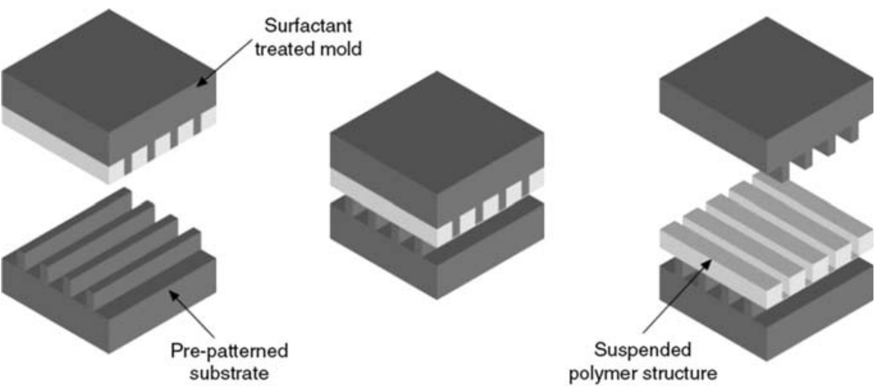


Fig. 5.28 Schematic of reverse imprint onto a pre-patterned substrate

treatment of mold surface, demolding of polymer structure should be easier, even though they are of high aspect ratio. Besides, UV curing process or drying of solvent can result in some shrinkage of polymer volume, which also facilitates the demolding process. Though the polymer structures shown in Fig. 5.27 are much big, the same process has been demonstrated to be able to mold sub-100 nm pattern structures in PMMA and polycarbonate (PC) [68].

Because patterns have already been formed in the polymer by casting into a mold, the imprinting process is only to press the polymer sticking onto a substrate. In this way, imprinted polymer structures can be transferred to a substrate which is either flat or already has surface topographical structures. The transfer process is schematically illustrated in Fig. 5.28. It is no different from the process shown in Fig. 5.26, except that the substrate now has topographical structures instead of being flat. An example of such transfer is to use a SiO_2 mold which has 350-nm-deep grating with a 700-nm period. To assist polymer transfer to the substrate, the mold is coated with 1H,1H,2H,2H-perfluorodecyl-trichlorosilane, which results in a very low-energy surface. The mold is then spin-coated with a polymer solution whose concentration is chosen so that the mold surface is planarized, having a surface profile of oscillating waves with the maximum peak height less than 100 nm. After spin-coating, the mold is soft-baked at 80°C for 5 min to remove residual solvent. The coated mold is then pressed against a pre-patterned SiO_2 substrate in a preheated hydraulic press. The typical imprinting pressure is 5 MPa. Cooling is initiated after 5 min and the pressure is sustained until the temperature has decreased to below the glass transition temperature (T_g) of the polymer. The mold is then separated from the patterned substrate, leaving polymer patterns on top of the protrusions on the substrate [69]. By repeating the process, multiple levels of polymer structures can be built up, as shown in Fig. 5.29, where three layers of PMMA patterns are stacked up.

One must be aware that there is always residual layer when the mold is coated with polymer to fill up the mold cavity. The sheet of polymer filling may or may

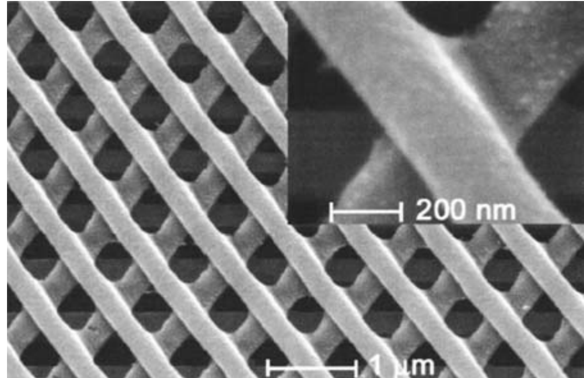


Fig. 5.29 Three layers of PMMA pattern structures are imprinted by reverse nanoimprint (Reprint from [69] with permission)

not be completely transferred onto the topographical substrate, depending on both the pattern density on the stamp and the property of polymer [64]. For dense pattern on a stamp (pattern protrusion area >25% of whole stamp area), the whole polymer sheet including the residual layer is most likely to be transferred onto the substrate, as schematically shown in Fig. 5.30a. This comes from the fact that dense pattern on stamp surface is more likely to have much better planarized polymer coating layer, resulting in a good contact between the residual layer and substrate surface. When the stamp is thermally pressed against the substrate, the residual layer can well stick to the substrate. On the other hand, sparsely patterned stamp (pattern protrusion area <25% of whole stamp area) will have poor planarization when coated with polymer. The polymer surface between two protrusions is in concave shape. There will be poor or no contact between this part of polymer and substrate. When the stamp is separated from substrate, this part of polymer will not stick to the substrate. Therefore, only the residual layer on top of the stamp protrusions will be transferred to the substrate, as schematically shown in Fig. 5.30b. However, the above rule is not absolute. If the polymer is more ductile, such as PC, the polymer in stamp cavities can still be

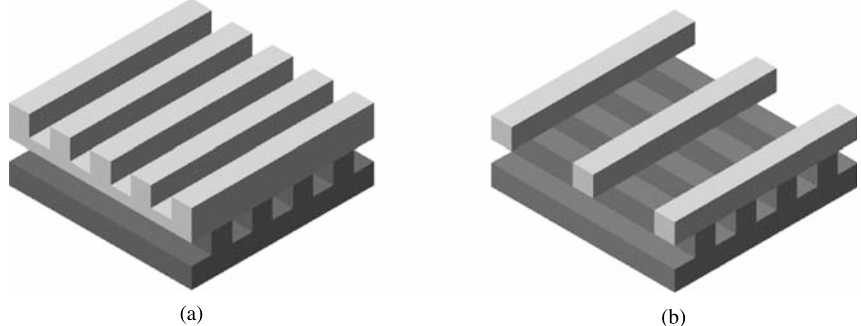


Fig. 5.30 (**a**) Reverse nanoimprint with high pattern density area. (**b**) Reverse nanoimprint with low pattern density area

transferred to substrate, even though there is a large span between two protrusions in a stamp. Likewise, for more brittle polymers such as PMMA, the polymer fillings can be taken away by stamp even for dense protrusions.

5.6 Soft Lithography

The previous section has shown that molded polymer pattern structures can be transferred from the stamp to a substrate by pressing the stamp in contact with the substrate. The same technique can be used to transfer liquid "ink" from the stamp to substrate. In fact, the concept of stamping ink onto a surface was proposed before the advent of nanoimprint. In 1993, Kumar and Whitesides reported using an elastomer stamp, made by casting PDMS into a master mold, to transfer alkanethiol ink to a substrate coated with gold thin film [70]. The alkanethiol ink pattern then acts as a gold etching mask to transfer the ink pattern into a gold thin film pattern on a substrate. The process is schematically illustrated in Fig. 5.31.

The PDMS (from Dow Corning with trade name Sylgard 184) is first mixed with curing agent (Sylgard curing agent 184) in 10:1 volume ratio. A vacuum deairing is normally performed to remove air bubbles introduced during mixing. The mixed liquid is then cast onto a patterned master surface. The curing can be done at room temperature (23°C) for 24 h, or 65°C for 4 h, or 100°C for 1 h, or 150°C for 15 min. After curing the rubber-like PDMS sheet is peeled off and soaked with thiol solution (inking). The stamp is then brought into contact with a substrate coated with thin gold film. Upon contact, the thiol ink is transferred from the stamp protrusion surface to the gold film. The thiol pattern serves as an etching mask during gold wet etch. In this way, the stamp pattern is transferred into a gold thin film pattern on the substrate. The initial work

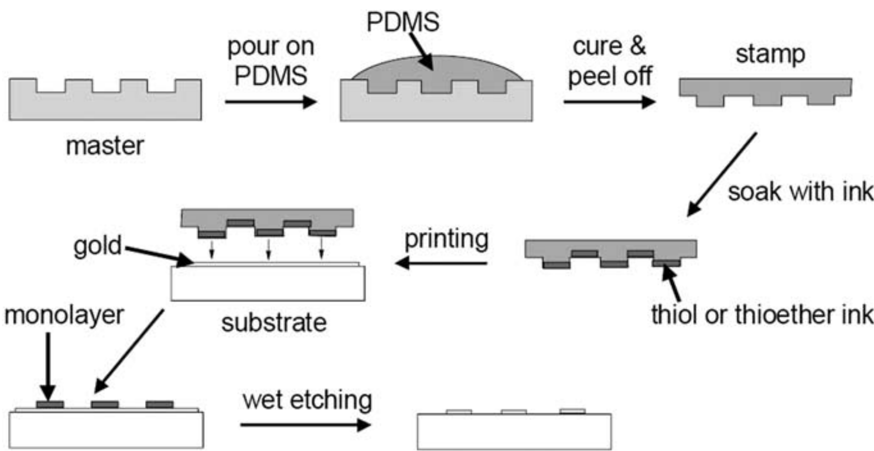


Fig. 5.31 Schematic of soft lithography process

reported in 1993 might not be perfect. The choice of PDMS for making the stamp might just be because it was cheap and easy to shape into a stamp. However, the concept has been taken up by many research groups around the world in the following decade. The word "soft lithography" was coined [71], and like the nanoimprint since it was proposed, many variations of soft lithographic technique have been developed, such as microcontact printing (μCP), replica molding (REM), microstransfer molding (μTM), micromolding in capillaries (MIMIC), and solvent-assisted micromolding (SAMIM) [72]. It has become an important replication technique parallel to the NIL technology.

5.6.1 Soft Stamps

Soft stamps are the key components in the whole array of soft lithography techniques. Why soft? It is because a soft stamp allows intimate contact between stamp surface and substrate surface even when the substrate surface is rough or non-flat. Unlike nanoimprint in which a stamp embosses into a polymer layer by external pressure, soft lithography transfers the ink pattern simply by surface contact. PDMS was the earliest and still is the primary choice of material. PDMS has many wonderful properties which are well suited for soft lithography [73]: (1) as an elastomer, PDMS is deformable and can easily make conformal contact with a nonplanar surface on micrometer scale. The elastic nature and low surface energy (~22 mJ m^{-2}) also allows it to be peeled off easily. On the other hand, the surface energy can be increased by oxygen plasma treatment so that surface binding becomes irreversible. (2) PDMS is chemically inert. It is resistant to many diluted acids and stable against many solvents and a good etch mask for RIE of silicon as well. (3) PDMS is optically transparent down to wavelength of 300 nm. Therefore, it is possible to use PDMS as a mold for UV curing imprint.

However, the very fact of softness of PDMS also causes several drawbacks in soft lithography, particularly, when trying to pattern sub-100 nm features. As schematically shown in Fig. 5.32, the low tensile modulus (~1.8 MPa) can lead to pattern structure collapse if h/a is too large, or to pattern sagging if w/a is too

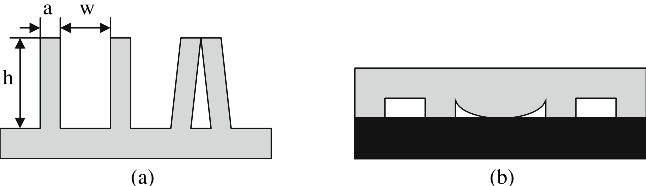


Fig. 5.32 Distortion of PDMS stamp: (**a**) pattern collapse due to high aspect ratio and close packed structures, and (**b**) pattern sagging due to large span and shallow cavity features

large [74]. One can imagine that for sub-100 nm stamp structures, the standing features are weaker than larger features and the depth in the stamp is shallower, all of which can lead to the aforementioned collapse and sagging. In fact, PDMS stamp structures tend to merge or collapse during inking and printing for structure dimension smaller than 1 μm [75].

The problem of soft PDMS stamp in printing sub-micrometer patterns is best demonstrated in Fig. 5.33, where the printed dense dot array from 250 nm down to 80 nm are compared using stamps with two different hardness [76]. One stamp was made of commercial PDMS (Sylgard 184) with modulus of 2 MPa and the printed dot array is shown in Fig. 5.33a. The other stamp was made of modified PDMS material with modulus of 9.7 MPa and the printed dot array is shown in Fig. 5.33b. The Sylgard 184 stamp can only print out 250-nm dot array, while the hard PDMS can print features down to 80 nm. Even for the largest 250-nm dots, soft stamp is printed out in round shape while hard stamp is printed in square. In general, commercial Sylgard 184 is only suitable for pattern structures above 500 nm [73]. In addition to the short range

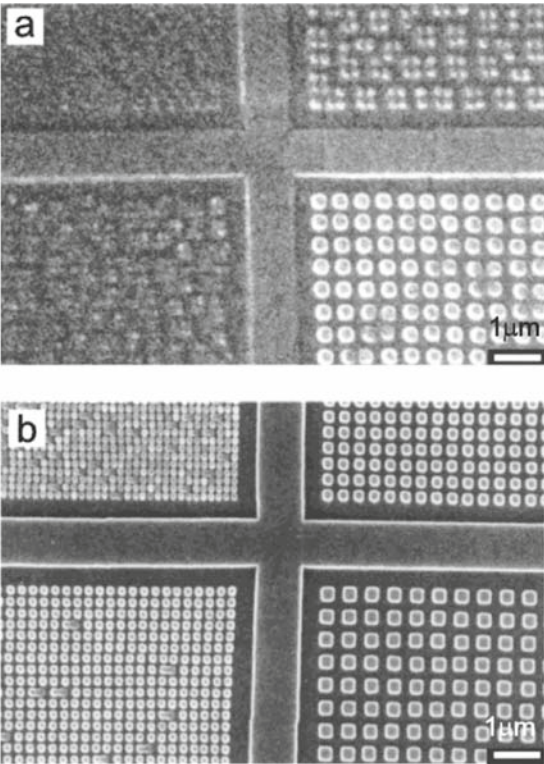


Fig. 5.33 Comparison of dot array pattern printed by (**a**) normal PDMS stamp and (**b**) hard PDMS stamp, dot sizes range from 250 nm (*lower right quadrant*) down to 80 nm (*upper left quadrant*) and structure height is 100 nm (Reprint from [76] with permission)

distortion, there is long range distortion in the form of pattern displacement error. Unsupported stamps made of Sylgard 184 showed nonuniform pattern displacements up to 20 μm between two successive prints over areas of 5 × 5 cm^2 [75]. The use of a rigid backplane can overcome such long-range distortion problem [77]. In order to maintain the flexibility of a soft stamp without the aforementioned drawbacks, the best option is to have a trilayer stamp which has a rigid glass backplane, a soft PDMS cushion, and hard PDMS pattern layer, as has already been schematically shown in Fig. 5.15. Some have investigated the tunability of modulus of polymer stamps. By tailoring the length of molecular chains in polyurethane acrylate (PUA), the tensile modulus of stamp can be tuned to either 320 MPa (hard stamp) or 19.8 MPa (soft stamp). Such a stamp can be either used for hard press imprint or soft contact imprint [78]. Apart from mechanical instability problem, there are other problems associated with PDMS soft stamps: PDMS shrinks about 1% after curing; it can swell when absorbing some solvents such as toluene and hexane [72]. All the above problems make it much hard for the soft lithography to achieve resolution as good as the nanoimprint techniques described in previous sections.

However, there is one possibility that can make the soft stamp capable of patterning sub-100 nm features. It relies on the concept of chromeless phase-shift lithography (CPL) which has been introduced in Chapter 2 (Section 2.5.2). A molded PDMS stamp with surface relief structures very much resembles a chromeless PSM, as long as the depth of surface relief structure is the half the wavelength of UV illuminating light. The optical property of PDMS is very much similar to that of quartz material. Once illuminated by a conventional UV light source, dark lines much narrower than the stamp feature dimension are produced at the edge of each surface relief feature due to phase cancellation of incoming light. Features of 30–50-nm dimension have been produced using PDMS stamps with stamp features of 400 nm and illuminated with 436-nm UV source [79]. By rotating the stamp 90° and having a second exposure on the same photoresist layer, dark dots can be produced. Similarly, ring-shaped patterns can be produced by exposure of stamp with circular posts.

5.6.2 Microcontact Printing

The soft lithography is basically a contact printing process. Microcontact printing is therefore an alternative name for soft lithography. Microcontact printing is possible because of two essential components: one is the PDMS stamp and another is the alkanethiol self-assembled monolayer (SAM) system. The previous section has emphasized on the soft aspect of contact printing. This section will focus on its printing characteristics. PDMS has an important property, apart from those mentioned previously, that is, the permeability to some

solvents. As a stamp, PDMS has to absorb inks first before it can stamp an ink pattern onto a substrate. The ink is normally alkanethiol dissolved in a solvent and PDMS can absorb the solvent like a sponge. In this case, aqueous ink (water-based) is not suitable because alkanethiol has low solubility in water and PDMS does not soak up water well (low permeability). The commonly used solvent is ethanol. Although ethanol is not the only solvent that can permeate PDMS, it causes less swelling in PDMS than others. The optimum concentration of thiols in ethanol is between 0.1 mM and 1 mM, depending on the resolution of stamp structures. Normally, low concentration is preferable for printing small features [72].

Inking of a PDMS stamp is normally done by placing a drop of thiol solution (about 1 μL per mm^2 area of stamp) onto the patterned stamp for 30 sec, as shown in Fig. 5.34a. The drop is then blown away with N_2 and the stamp is dried for 10 sec under a continuous stream of N_2. The inked stamp is then used to print within ~15 s after inking. Another "wet inking" method is to dip the PDMS stamp into thiol solution for a few minutes and blow dry under a N_2 stream, as shown in Fig. 5.34b. Although there was an attempt to attach a reservoir of thiol solution to a PDMS stamp to provide constant supply of thiol ink, it did not succeed because of the swelling of PDMS stamp which was constantly soaking up with the ethanol solvent [80]. It is obvious that both of the "wet inking" methods (Fig. 5.34a and b) have thiols homogeneously distributed over all the surface of stamp, including sidewalls and bottom of stamp cavities. When the PDMS stamp is pressed down to print, two things can happen in addition to printing a stamp pattern: one is the diffusion of thiols from cavity sidewalls to the substrate surface, which broaden the printed ink feature; another is the evaporation of thiol-containing solvent from the cavity onto the substrate area where an ink pattern is not supposed to be there. The evaporated traces of thiol can affect the contrast of printed pattern because at a later stage of etching, any traces of thiol can deter the etching of substrate material. The consequence of "wet inking" of stamp is the difficulty to achieve high-resolution printing features. To overcome the above problems, there is a third inking method, as shown in Fig. 5.34c. It uses an ink pad which is a piece of flat PDMS soaked up with thiol solution. The inking of stamp is done by direct contact of stamp with the inking pad. The amount of ink transferred onto the stamp surface is controlled by the concentration of thiol in the solvent. In this way, only the pattern protrusion surfaces have inks. The printed ink pattern will be exactly the same dimension as the pattern structures on the stamp. With the contact-inking method, 100-nm gold line features have been produced [80]. One thing should be noticed that a PDMS stamp with "wet inking" can be used repeatedly without further inking (~10 times per inking) because the absorbed ink in PDMS stamp can slowly diffuse out, whereas stamps inked by direct contact can only be used once per inking.

As mentioned earlier in this section, μCP cannot work without the alkanethiol SAM system. Two aspects of the alkanethiol SAM system make it unique in μCP: one is that the thiol molecules can self-assemble into a

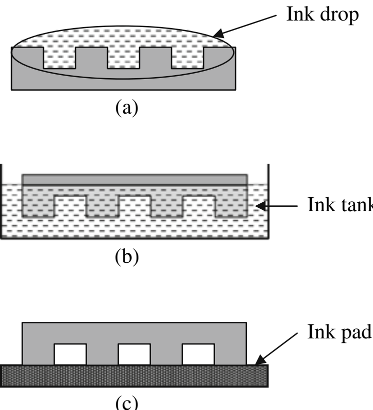


Fig. 5.34 (**a**) Wet inking by ink drop. (**b**) Wet inking by dipping into an ink tank. (**c**) Dry inking by contact with an ink pad

monolayer on the surface of coinage metals such as Au, Ag, Cu, and Pd. The monolayer-forming process is fast and takes place at room temperature. This is important for making a uniform film with low defect density. Although PDMS stamp is elastic and can form an intimate contact with substrate surface, it can never be guaranteed that the ink can be transferred onto substrate uniformly. Particularly, if the ink molecules congregate into clusters once on another material surface, pinhole type of defects can be easily formed, rendering the ink pattern useless as an etching mask. The self-assembling of thiol molecules on a gold surface is fast and automatic, only needing the inked stamp in contact with the gold surface for a short time between 10 s and a min [72]. Longer contact time between an inked stamp and substrate can further increase the density and coverage of SAM, but it also can induce diffusion of thiol molecules outside the boundary of stamp patterns, making the printed feature larger than intended. The diffusion limits the resolution of μCP technique to above 100 nm. Special ink such as immobilized enzyme can be used, which transfer the stamp pattern by catalytic reaction when in contact with acid-coated surface. With enzymatic ink, the resolution of μCP is only limited by the range of motion of the immobilized catalyst, which is about 2 nm [81].

In addition to the self-assembling property, another unique property of thiol-based SAM is its very high resistance to wet chemical etching. For example, gold surface protected by a monolayer of hexadecanethiol (HDT) is 10^6 times slower in cyanide (gold etchant) [82]. The two unique properties make the thiol SAM a perfect masking material. Unfortunately, the thiol SAM system only works with coinage metals, among them thin film gold is the most used material. Nevertheless, one can always use the patterned gold film as a mask to further transfer into other materials. For example, μCP has been used to pattern PMMA resist layer, in which a 20-nm Au film was coated on

top of PMMA resist layer. After contact printing by a PDMS stamp, the HDT-inked pattern was transferred into the Au film. Then the Au film was used as etch mask to pattern the PMMA resist layer by oxygen plasma etching [83]. The same concept can be applied to many other multilayer systems as long as the Au or Ag film can act as a masking material in subsequent pattern transfer processes.

5.6.3 Replication by Capillary Force

The unique property of PDMS also enables some other ways of replicating micro- and nanostructures. One of the earliest reported, apart from the μCP, was micro moulding in capillaries (MIMIC) [84]. The principle of MIMIC is quite simple, as schematically shown in Fig. 5.35. A PDMS stamp is brought to intimate contact with a flat substrate. A low-viscosity liquid polymer is fed into the channels formed by cavities in the stamp. By capillary action, the liquid polymer fills in these channel cavities (Fig. 5.35a and b). A UV curing or thermal curing solidifies the polymer inside the cavities. After peeling off the PDMS stamp, a polymer replication of the channel pattern is left on the

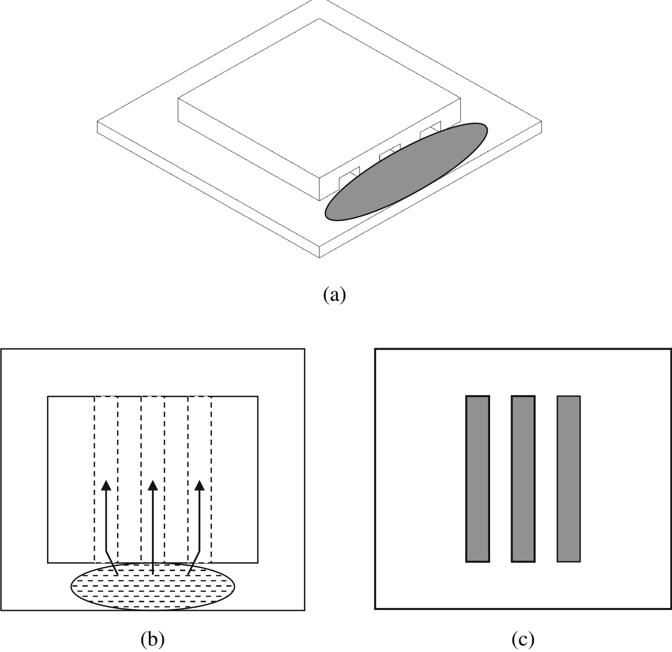


Fig. 5.35 Schematic of micro molding in capillary (MIMIC)

substrate (Fig. 5.35c). Instead of feeding liquid polymer from the side of stamp, the polymer can also be coated on a flat substrate and is heated above its T_g while a PDMS stamp seals from the top. When the polymer lowers its viscosity at high temperature, it can also fill up the stamp cavities with capillary action [85]. Alternatively, the PDMS stamp can be directly placed onto a freshly spin-coated polymer film while it is still wet. Liquid polymer gets into the stamp cavities by capillary force, if the PDMS surface is properly treated and can be wetted. The PDMS stamp can absorb all the solvent contained in the polymer film because of the solvent permeability of PDMS. For example, PDMS can absorb propylene glycol mono ether acetate (PGMEA) up to 27 wt.% of its weight and 30 wt.% of absorbed PGMEA will diffuse into air in 1 h, which creates a partial vacuum inside the PDMS cavity and helps to draw in liquid polymer [86]. Another example of filling mold cavity by capillary action is the SFIL technique which has been described in Section 5.4.3. Analysis has shown that the capillary pressure is greater than 1 bar when the stamp feature size is smaller than 440 nm. The smaller the stamp feature the higher the capillary pressure becomes [87].

Another form of replication using capillary force is the solvent-assisted microcontact molding (SAMIM) [88]. The SAMIM process is illustrated in Fig. 5.36. A polymer film is coated onto a substrate and baked. A patterned PDMS stamp is wetted with a matching solvent for the polymer, and then brought into contact with the polymer surface. The solvent on the stamp dissolves the polymer and this dissolved polymer is molded into the channels or cavities formed between the stamp and the substrate by capillary action. The matching solvent for PMMA is acetone and for photoresists it is ethanol. The molded polymer structure becomes solidified in a few minutes after evaporation of solvent, while the stamp is still in conformable contact with the substrate. There are two problems associated with the SAMIM technique. The first is that the surface of the patterned polymer is highly textured because of kinetic movement of polymer molecules during solvent dissolving process. The other is that the pattern formed by SAMIM gets distorted. In particular, the base of a rectangular stripe becomes wider than the top.

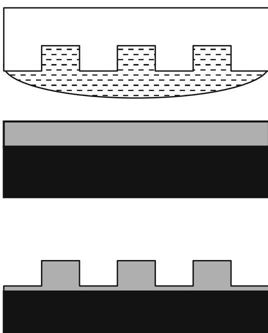


Fig. 5.36 Solvent-assisted microcontact molding (SAMIM)

Polydimethylsiloxane (PDMS) is not only permeable to solvent but also permeable to gas. The gas permeability coefficient of PDMS is higher than other polymers. Table 5.1 compares permeability coefficients of PDMS with two other common polymer materials for some typical gases.

Table 5.1 Gas permeability coefficients of PDMS, PS, and PMMS

	Permeability coefficient (Barrer)		
Polymers	He	N_2	O_2
PDMS	590	351	781
Polystyrene (PS)	22.4	0.52	2.9
PMMA	8.4	0.02	0.14

(1 Barrer $= 10^{-10}$ cm^3 (STP) cm cm^{-2} $cmHg^{-1}$ s^{-1}) [89]

It is apparent that PDMS has very high air ($N_2 + O_2$) permeability. This property has been utilized to replicate polymer pattern structure through PDMS molds in a so-called micro-aspiration-assisted lithography (MAAL) technique [90]. The MAAL is illustrated in Fig. 5.37. It uses a double-layer mold structure. The top layer consists of thick PDMS (5 mm) with a matrix of channels leading to an outlet which is connected to a vacuum pump. The bottom layer is the actual PDMS mold consisting of various pattern cavities and is separated from the top layer by a thin PDMS membrane (20 μm). Here is the working principle of MAAL: when the composite mold is brought in contact with the polymer-coated flat substrate (no pre-baking of spin-coated polymer), it seals off the cavities with the liquid polymer layer; a low vacuum (870 mbar) is created by evacuating air from the channels in the top PDMS layer; the pressure difference between the sealed mold cavities and the channels helps air permeating ("aspirating") through separation PDMS membrane, resulting in a vacuum in the mold cavities; the liquid polymer is then drawn into the cavities by the ambient atmospheric pressure. After filling the mold cavities, a thermal or UV curing process can solidify the polymer and a replication of polymer pattern is formed after peeling of the PDMS mold.

The mold filling process is not assisted by capillary action, but by the air permeability of PDMS and the pressure difference created across the PDMS membrane. The time required to fill the mold cavities depends directly on the air permeability coefficient of PDMS and the pressure difference. If the polymer cannot wet the PDMS mold surface, or the PDMS surface is hydrophobic to the polymer liquid, then the pressure difference has to be greater than the negative capillary force which resists the polymer flowing into the PDMS mold cavity. The polymer filling speed also depends on the viscosity of polymer and geometry of the mold cavities. Nevertheless, MAAL is a much faster process than previously described capillary force lithography. For the same mold area, MAAL only needs 2 min to fill up the entire mold cavity while capillary force filling needs 20 min, provided the polymer is a wetting material to PDMS [90]. Then for nonwetting polymer, MAAL also works. Despite its advantages, MAAL may have difficulty to fill mold cavity of small geometry if the PDMS

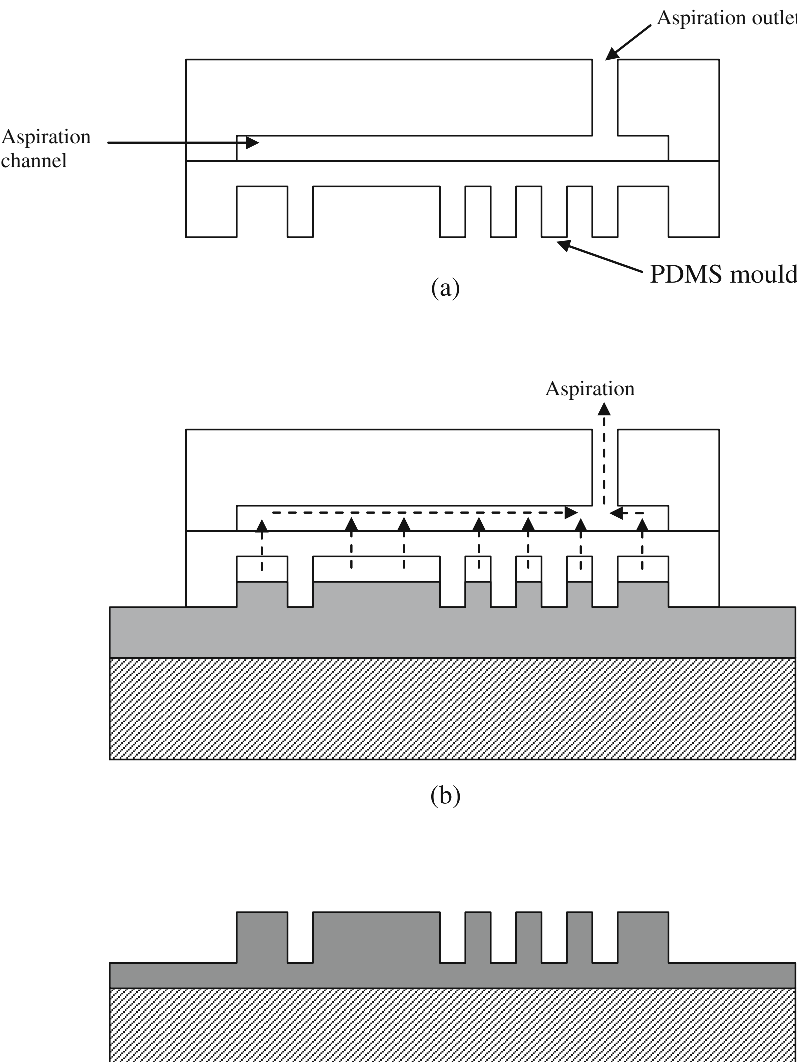


Fig. 5.37 Schematic of micro-aspiration assisted lithography (MAAL) process: (**a**) composite PDMS mold, (**b**) aspiration filling mold vertically, and (**c**) Molded structure after UV curing and demolding

surface appears to be hydrophobic to the filling polymer, because the smaller the cavity geometry the greater the negative capillary force. If the PDMS surface is hydrophilic to the filling polymer, filling can be accomplished even without the vacuum assistance.

References

1. Chou, S.Y., P.R. Krauss, and P.J. Renstrom, *Imprint of sub-25 nm vias and trenches in polymers*. Appl. Phys. Lett., 1995. **67**(21): pp. 3114–3116.
2. Chou, S.Y., et al., *Sub-10 nm imprint lithography and applications*. J. Vac. Sci. Technol., 1997. **B15**(6): p. 2897.
3. *International Technology Roadmap for Semiconductors*. [cited; Available from: http://www.itrs.net/].
4. Torres, C.M.S., *Alternative lithography - Unleashing the Potentials of Nanotechnology*. 2003, New York: Kluwer Academic/Plenum Publisher.
5. Heidari, B., et al., *Large scale nanolithography using nanoimprint lithography*. J. Vac. Sci. Technol., 1999. **B17**(6): p. 2961.
6. Haatainen, T., et al., *Nickel stamp fabrication using step & stamp imprint lithography*. Microelectron. Eng., 2006. **83**: pp. 948–954.
7. Haffner, M., et al., *Simple high resolution nanoimprint-lithography*. Microelectron. Eng., 2007. **84**: pp. 937–939.
8. Bird, R.B., R.C. Amstrong, and O. Hassager, *Fluid mechanics*, in *Dynamics of Polymeric Liquids*, 1977, John Wiley & Sons.
9. Schift, H. and L. Heyderman, *Nanorheology: Squeeze flow in hot embossing of thin films*, in *Alternative Lithography*, C.M.S. Torres, Editor. 2003, Kluwer Academic.
10. Halary, J.L., et al., J. Polym. Sci. (B): Polym. Phys., 1991. **29**: p. 933.
11. Hoffmann, T., *Viscoelastic properties of polymers: Relevance for hot embossing lithography*, in *Alternative Lithography*, C.M.S. Torres, Editor. 2003, Kluwer Academic.
12. Scheer, H.C. and A.E. Al, *Problems of the nanoimprinting technique for nanometer scale pattern definition*. J. Vac. Sci. Technol., 1998. **B16**(6): p. 3917.
13. Landis, S., et al., *Stamp design effect on 100 nm feature size for 8 inch nanoImprint lithography*. Nanotechnology, 2006. **17**: pp. 2701–2709.
14. Cui, B. and T. Veres, *Pattern replication of 100 nm to millimeter-scale features by thermal nanoimprint lithography*. Microelectron. Eng., 2006. **83**: pp. 902–905.
15. Schulz, H., et al., *Impact of molecular weight of polymers and shear rate effects for nanoimprint lithography*. Microelectron. Eng., 2006. **83**: pp. 259–280.
16. *Nanonex Corp*. [cited; Available from: http://www.nanonex.com/].
17. *Microresist GmbH*. [cited; Available from: http://www.microresist.de/].
18. Bogdanski, N., et al., *Structure size dependent recovery of thin polystyrene layers in thermal imprint lithography*. Microelectron. Eng., 2007. **84**: pp. 860–863.
19. Workum, K.V. and J.J.D. Pablo, Nano Lett., 2003. **3**: p. 1405.
20. Ro, H., et al., *Evidence for internal stresses induced by nanoimprint lithography*. J. Vac. Sci. Technol., 2006. **B24**(6): p. 2973.
21. Ramachandran, S., et al., *Deposition and patterning of diamondlike carbon as antiwear nanoimprint templates*. J. Vac. Sci. Technol., 2006. **B24**(6): p. 2993
22. Chou, S.Y., P.R. Krauss, and P.J. Renstrom, *Nanoimprint lithography*. J. Vac. Sci. Technol., 1996. **B14**(6): p. 4129.
23. Konishi, T., et al., *Multi-layered resist process in nanoimprint lithography for high aspect ratio pattern*. Microelectron. Eng., 2006. **83**: pp. 869–872.
24. Zhang, W. and S.Y. Chou, *Multilevel imprinting lithography with submicron alignment over 4 in. Si wafers*. Appl. Phys. Lett., 2001. **79**(6): p. 845.
25. Tan, H., et al., *Current status of Nanonex nanoimprint solutions*. SPIE, 2004. **5374**: pp. 213–221.
26. *Suss MicroTech*. [cited; Available from: http://www.suss.com/].
27. Ahopelto, J. and T. Haatanien, *Step and stamp imprint lithography*, in *Alternative Lithography*, C.M.S. Torres, Editor. 2003, Kluwer Academic.
28. Chen, Y., et al., *A study of pattern placement error by thermal expansions in nanoimprint lithography*. J. Microlitho. Microfab. Microsys., 2006. **5** (1).

29. Lebib, A., et al., *Room temperature and low pressure nanoimprint lithography*. Microelectron. Eng., 2002. **61–62**: p. 371.
30. Nakamatsu, K. and S. Matsui, *Room-temperature nanoimprint and nanocontact technologies*, in *Nanomanufacturing Handbook*, A. Busnaina, Editor. 2007, CRC Press.
31. Tao, J., et al., *Room temperature nanoimprint lithography using a bilayer of HSQ/PMMA resist stack*. Microelectron. Eng., 2005. **78–79**: pp. 665–669.
32. Matsui, S., et al., J. Vac. Sci. Technol., 2001. **B19**: p. 2801.
33. Lu, Y., et al., *Patterning layered polymeric multilayer films by room-temperature nanoimprint lithography*. Macromol. Rapid Comm., 2006. **27**(7): pp. 505–510.
34. Haisma, J., et al., *Mold-assisted nanolithography: A process for reliable pattern replication*. J. Vac. Sci. Technol., 1996. **B14**: p. 4124.
35. Voisin, P., et al., *High-resolution fused silica mold fabrication for UV-nanoimprint*. Microelectron. Eng., 2007. **84**: pp. 916–920.
36. Min, J., et al., *Effect of sidewall properties on the bottom microtrench during SiO_2 etching in a CF4 plasma*. J. Vac. Sci. Technol., 2005. **B23**(2): p. 425.
37. Cui, Z., *Etching technology*, in *Micro-Nanofabrication Technologies and Applications,* 2006, Springer.
38. Dauksher, W.J., et al., *Step and flash imprint lithography template characterization from an etch perspective*. J. Vac. Sci. Technol., 2003. **B21**(6): p. 2771.
39. Plachetka, U., et al., *Wafer scale patterning by soft UV-nanoimprint lithography*. Microelectron. Eng., 2004. **73–74**: pp. 167–171.
40. Schmid, G.M., et al., *Template fabrication for the 32 nm node and beyond*. Microelectron. Eng., 2007. **84**: pp. 853–859.
41. Bender, M., et al., *Status and prospects of UV-nanoimprint technology*. Microelectron. Eng., 2006. **83**: pp. 827–830.
42. Plachetka, U., et al., *Comparison of multilayer stamp concepts in UV-NIL*. Microelectron. Eng., 2006. **83**: pp. 944–947.
43. Kawaguchi, Y., F. Nonaka, and Y. Sanada, *Fluorinated materials for UV nanoimprint lithography*. Microelectron. Eng., 2007. **84**: pp. 973–976.
44. Beck, M. and B. Heidari, *Nanoimprint lithography for high volume HDI manufacturing*. OnBoard Technology, 2006. **September** p. 52.
45. Wang, X., et al., *High density patterns fabricated in SU-8 by UV curing nanoimprint*. Microelectron. Eng., 2007. **84**(872–876).
46. Schmitt, H. and C. Lehrer, *UV polymers for nanoimprint lithography*, presentation at *2nd FORNEL Workshop on Nanoelectronics* March 2006.
47. Bender, M., et al., *Multiple imprinting in UV-based nanoimprint lithography: Related material issues*. Microelectron. Eng., 2002. **61–62**: pp. 407–413.
48. Vogler, M., et al., *Development of a novel, low-viscosity UV-curable polymer system for UV-nanoimprint lithography*. Microelectron. Eng., 2007. **84**: pp. 984–988.
49. Voisin, P., et al., *Characterisation of ultraviolet nanoimprint dedicated resists*. Microelectron. Eng., 2007. **84**: pp. 967–972.
50. Guo, L.J., *Nanoimprint lithography: Methods and material requirements*. Adv. Mater., 2007. **19**: pp. 495–513.
51. Le, N.V., et al., *Selective dry etch process for step and flash imprint lithography*. Microelectron. Eng., 2005. **78–79**: pp. 464–473.
52. Colburn, M., et al., *Step and flash imprint lithography: A new approach to high-resolution patterning*. Proc. SPIE, 1999. **3676**: p. 379.
53. Resnick, D.J., S.V. Sreenivasan, and C.G. Willson, *Step & flash imprint lithography*. Mater. Today, 2005. **8**(2): pp. 34–42.
54. *Molecule Imprint Corp*. [cited; Available from: http://www.molecularimprints.com/].
55. Melliar-Smith, M., *Lithography beyond 32 nm – a role for imprint?*, Available at http://www.molecularimprints.com/
56. Murthy, S., et al., *S-FIL technology: cost of ownership case study*. Proc. SPIE, 2005. **5751**: pp. 964–975.

57. Choi, B.J., et al., *Layer-to-layer alignment for step and flash imprint lithography*. Proc. SPIE, 2001. **4343**: pp. 436–442.
58. Moel, A., et al., *Novel on-axi interferometric alignment method with sub-10 nm precision*. J. Vac. Sci. Technol., 1993. **B11**(6): p. 2191.
59. Muhlberger, M., et al., *A Moiré method for high accuracy alignment in nanoimprint lithography*. Microelectron. Eng., 2007. **84**: pp. 925–927.
60. Li, W.W.N. and S.Y. Chou, *Sub-20-nm alignment in nanoimprint lithography using Moiré Fringe*. Nano Lett., 2006. **6**(11): pp. 2626–2629.
61. Cheng, X. and L.J. Guo, *One-step lithography for various size patterns with a hybrid mask-mold*. Microelectron. Eng., 2004. **71**: pp. 288–293.
62. Guo, L.J., *Recent progress in nanoimprint technology and its applications*. J. Phys. D: Appl. Phys., 2004. **37**: pp. R123–R141.
63. Huang, X.D., et al., *Reversal imprinting by transferring polymer from mold to substrate*. J. Vac. Sci. Technol., 2002. **B 20**: p. 2872
64. Kehagias, N., et al., *Three-dimensional polymer structures fabricated by reversal ultraviolet-curing imprint lithography*. J. Vac. Sci. Technol., 2005. **B 23**(6): p. 2954.
65. Sogo, K., et al., *Reproduction of fine structures by nanocasting lithography*. Microelectron. Eng., 2007. **84**: pp. 909–911.
66. Yang, B. and S.W. Pang, *Multiple level nanochannels fabricated using reversal UV nanoimprint*. J. Vac. Sci. Technol., 2006. **B 24**(6): p. 2984.
67. Yoshikawa, T., et al., *Fabrication of 1/4 wave plate by nanocasting lithography*. J. Vac. Sci. Technol., 2005. **B 23**(6): p. 2939.
68. Hirai, Y., et al., *Fine pattern transfer by nanocasting lithography*. Microelectron. Eng., 2005. **78–79**: p. 641.
69. Bao, L.R., et al., *Nanoimprinting over topography and multilayers three-dimensional printing*. J. Vac. Sci. Technol., 2002. **20**(6): p. 2881.
70. Kumar, A. and G.M. Whitesides, *Features of gold having micrometer to centimeter dimensions can be formed through a combination of stamping with an elastomeric stamp and an alkanethiol "ink" followed by chemical etching*. Appl. Phys. Lett., 1993. **63**(14): p. 2002.
71. Zhao, X.M., Y.N. Xia, and G.M. Whitesides, *Soft lithographic methods for nano-fabrication*. J. Mater. Chem., 1997. **7**(7): pp. 1069–1074.
72. Xia, Y. and G.M. Whitesides, *Soft lithography*. Angew. Chem. Int. Ed., 1998. **37**: pp. 550–575.
73. Tormen, M., *Microcontact printing techniques*, in *Alternative Lithography*, C.M.S. Torres, Editor. 2003, Kluwer Academic.
74. Hui, C.Y., et al., *Constraints on microcontact printing imposed by stamp deformation*. Langmuir, 2002. **18**: pp. 1394–1407.
75. Delamarche, et al., *Stability of molded polydimethylsiloxane microstructures*. Adv. Mater., 1997. **9**(9): pp. 741–746.
76. Schmid, H. and B. Michel, *Siloxane polymers for high-resolution, high-accuracy soft lithography*. Macromolecules, 2000. **33**: pp. 3042–3049.
77. Tormen, M., et al., *Sub-mm thick rubber-elastic stamp on rigid support for high reliability microcontact printing*. Microelectron. Eng., 2002. **61–62**: pp. 469–473.
78. Yoo, P.J., et al., *Unconventional patterning with a modulus-tunable mold: From imprinting to microcontact printing*. Chem. Mater., 2004. **16**: pp. 5000–5005.
79. Odom, T.W., et al., *Generation of 30–50 nm structures using easily fabricated composite PDMS masks*. J. Am. Chem. Soc., 2002. **124**: pp. 12112–12113.
80. Libioulle, L., et al., *Contact-inking stamps for microcontact printing of alkanethiols on gold*. Langmuir, 1999. **15**: pp. 300–304.
81. Snyder, P.W., et al., *Biocatalytic microcontact printing*. J. Org. Chem., 2007. **72**: pp. 7459–7461.
82. Biebuyck, H.A., et al., *Lithography beyond light: Microcontact printing with monolayer resists*. IBM J Res. Dev., 1997. **41**(1/2): p. 159.

83. Chen, Y., et al., *Microcontact printing and pattern transfer with a tri-layer processing*. Microelectron. Eng., 2000. **53**: pp. 253–256.
84. Kim, E., Y. Xia, and G.M. Whitesides, *Polymer microstructures formed by moulding in capillaries*. Nature, 1995. **376**: p. 581.
85. Suh, K.Y. and H.H. Lee, *Capillary force lithography: Large-area patterning, self-organisation and anisotropic dewetting*. Adv. Funct. Mater., 2002. **12**(6–7): p. 406.
86. Kim, Y.S., K.Y. Suh, and H.H. Lee, *Fabrication of three-dimensional microstructures by soft molding*. Appl. Phys. Lett., 2001. **79**(14): p. 2285.
87. Yoon, H., et al., *Capillary force lithography with impermeable molds*. Appl. Phys. Lett., 2006. **88**: p. 254104
88. Kim, E., et al., *Solvent-assisted microcontact molding: A convenient method for fabricating three-dimensional structures on surfaces of polymers*. Adv. Mater., 1997. **9**: p. 651.
89. Eddings, M.A. and B.K. Gale, *A PDMS-based gas permeation pump for on-chip fluid handling in microfluidic devices*. J. Micromech. Microeng., 2006. **16**: pp. 2396–2402.
90. Berre, M.L., et al., *Micro-aspiration assisted lithography*. Microelectron. Eng., 2007. **84**: pp. 864–867.

Chapter 6
Nanoscale Pattern Transfer

6.1 Introduction

The lithography process is not finished until a functional pattern structure has been made on or into a substrate material. There are of course many ways of making a pattern directly out of a material. Some of them have already been introduced in previous chapters. For example, focused ion beam (FIB) can directly sputter etch a nanoscale pattern into a substrate material, or make a structure on top of a substrate by gas-assisted deposition of a material. Electron beam (e-beam), though not able to sputter etch, can also deposit materials with gases. Both of the techniques do not need a mask, because the charged particle beams can be focused and scanned into any patterns, as described in Chapter 3. The same is true for patterning by a scanning probe. Etching or deposition of materials can be done by a probe interacting with a substrate, which has been explained in detail in Chapter 4. The present chapter focuses on the processes whereby a mask pattern is transferred into another material. The mask is in most cases a polymer resist which has been patterned by a lithographic technique, such as photons, charged particle beams, or nanoimprinting stamps.

The technologies introduced so far in this book have been mostly lithographic patterning of polymer resists. Generally, the polymer nanopatterns have no other functions except serving as a mask in subsequent pattern transfer processes. There are many ways of transferring a polymer resist pattern into a substrate. These processing techniques can be grouped into either additive processes or subtractive processes; that is, adding a new material through, or removing material from, the openings of polymer resist mask. Figure 6.1 illustrates the four most commonly used pattern transfer processes. Figure 6.1a is lift-off process by which another material is deposited through a resist pattern opening. Figure 6.1b is another additive process which adds a material by electroplating inside the resist pattern opening. Figure 6.1c is a subtractive process by which a thin film deposited on a substrate is patterned by etching through the resist mask opening, whereas Fig. 6.1d shows the resist pattern is transferred directly into a substrate by etching through the resist mask openings.

Ideally, a mask pattern should be faithfully transferred into another material without any distortions. In practice, this is a difficult task, because each pattern

Z. Cui, *Nanofabrication*, DOI: 10.1007/978-0-387-75577-9_6,

© Springer Science+Business Media, LLC 2008

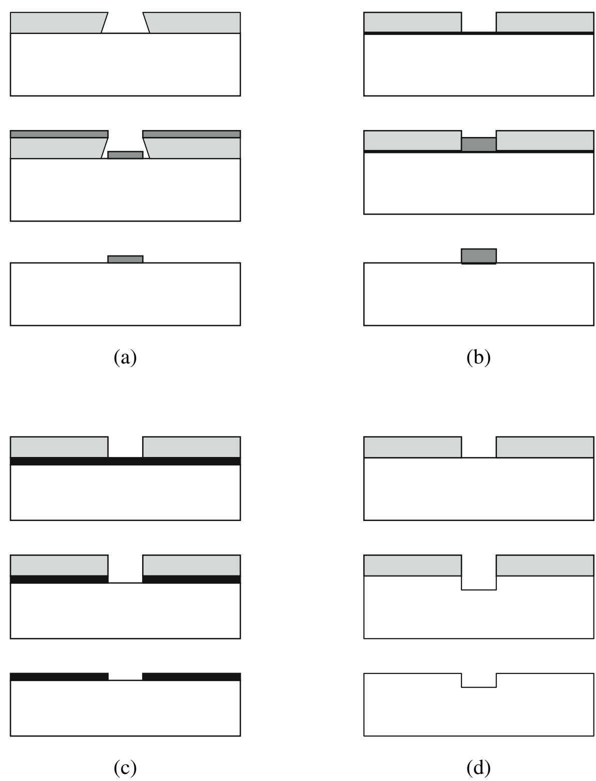


Fig. 6.1 Schematic of pattern transfer processes: (**a**) lift-off, (**b**) electroplating, (**c**) etching into a thin film, and (**d**) etching into substrate

transfer technique has its own capabilities and limitations, particularly when dealing with sub-100 nm pattern transfer. Some of the pattern transfer techniques are inherently not able to replicate the dimension of mask patterns. Even those techniques which are capable of nanoscale pattern transfer are still highly dependent on various process conditions to achieve desired accuracy and structure profile. The aim of this chapter is not to introduce every pattern transfer technique, but only those having the capability of transferring polymer resist patterns in sub-100 nm dimension. The principles of these techniques are described and particular attentions are drawn to the conditions at which nanoscale pattern transfer becomes possible.

6.2 Additive Pattern Transfer

6.2.1 Thin Film Deposition

The additive pattern transfer process involves depositing thin films into the opening of a polymer resist mask. Although there are many different thin-film deposition techniques, not all of them are suitable for additive pattern transfer.

A thin film can be deposited by physical vapor deposition (PVD), chemical vapor deposition (CVD), molecule beam epitaxy (MBE), oxidation, spin or spray coating and plating [1]. Physical vapor deposition and CVD are the two most commonly used processes. Physical vapor deposition includes thermal evaporation deposition and plasma sputtering deposition. The thermal evaporation can further be divided into resistive heating evaporation and e-beam heating evaporation. The plasma sputtering is differentiated according to how the plasma is generated. It can be a DC plasma sputtering, magnetron plasma sputtering, or radio frequency (RF) plasma sputtering. Chemical vapor deposition includes low-pressure CVD (LPCVD), plasma-enhanced CVD (PECVD), atmospheric pressure CVD (APCVD), metal organic CVD (MOCVD). The main features of some PVD and CVD thin-film deposition techniques are summarized in Table 6.1.

One of the immediate concerns, when thin-film deposition process is used for additive pattern transfer, is the substrate temperature. As described earlier, polymer resist which is the masking material for the pattern transfer cannot withstand temperature higher than its glass transition temperature (T_g). As it is known from the previous chapters, at above T_g temperature polymers become softened and flowable. Distortions become inevitable particularly for small-dimension and high-aspect-ratio patterns. The requirement for low substrate temperature leaves only the PVD as the suitable thin-film deposition technique for additive pattern transfer. Another requirement is the directionality of deposition. As shown in Fig. 6.1a, additive pattern transfer by lift-off needs the thin-film deposition to have poor step coverage, so that the film deposited onto the resist masking layer can be easily removed by dissolving the polymer resist. Good step coverage will connect the film between top and bottom of polymer resist patterns, making the lift-off difficult if not impossible. The poor step coverage is ensured by good directionality of deposition. As seen in Table 6.1, this leaves only the thermal evaporation deposition as the most suitable thin-film deposition technique for lift-off pattern transfer. Thin films deposited by CVD techniques are primarily for subtractive pattern transfer process; that is, the films have already been deposited on a substrate before the polymer resist layer is lithographically patterned.

Because of its high directionality, the deposition rate of thin film by thermal evaporation depends on the distance and relative position between substrate and evaporation source. It follows the Knudsen's cosine law, as expressed in Eq. 6.1:

$$R \propto \frac{\cos\theta\cos\phi}{r^2}(P_s(T) - P_c). \quad (6.1)$$

In the above equation, R is the deposition rate which is the function of evaporation source to sample distance r, the angle (θ) of sample to the axis of evaporation source and the orientation of sample surface ϕ. The source position relative to a sample is schematically indicated in Fig. 6.2a. The rate also depends on the vapor pressure of source material P_s which is the function of heating temperature and the pressure in the evaporation chamber P_c. It is

Table 6.1 Characteristics of some major thin-film deposition techniques

Techniques	Materials	Film uniformity	Film density	Film grain size	Deposition rate	Substrate temperature	Directionality
Resist heating evaporation	Metals of other low-melting-point materials	Poor	Poor	10–100 nm	1–20 A s^{-1}	50–100°C	Good
E-beam evaporation	High-melting-point metals and dielectrics	Poor	Poor	10–100 nm	10–100 A s^{-1}	50–100°C	Good
Plasma sputtering	Metals and dielectrics	Very good	Good	~10 nm	Metal: ~100 A s^{-1} Dielectric: 1–10 A s^{-1}	~200°C	Moderate
PECVD	Dielectrics	Good	Good	10–100 nm	10–100 A s^{-1}	200–300°C	Moderate
LPCVD	Dielectrics	Very good	Excellent	1–10 nm	10–100 A s^{-1}	600–1200°C	Poor

apparent that a flat sample placed at a certain distance from the evaporation source will have nonuniform film thickness at different locations of the sample. Assuming the film thickness at the center of sample (on the axis of evaporation source) is t_0, the thickness at an off-center location with distance x is t_x, the height of sample above the source is h, and the size of evaporation source is assumed to be much smaller than the source to sample distance (h), then t_x can be worked out from Eq. 6.2, according to the Knudsen's cosine law [2].

$$t_x = \frac{t_0}{\left[\left(1 + \frac{x}{h}\right)^2\right]^2} \tag{6.2}$$

The Eq. 6.2 indicates that film thickness at an off-center location on a flat sample is always less than at a close-to-center location. Such difference can only be reduced if the sample is further away from the evaporation source. In order to achieve uniform deposition, the sample size should be small and the evaporation chamber should be big. All of these measures are at a price of low deposition rate and less efficient use of source material.

It is possible to obtain uniform film thickness, provided that the sample surface is spherical, as illustrated in Fig. 6.2b. In practice, samples which are usually small in size can be placed on spherically shaped sample holder. At any point of the spherical surface, the angles θ and ϕ are equal. Assuming the sphere has radius of r_0, then the distance between any point on the spherical surface and the source can be expressed as $r = 2r_0 \cos\theta = 2r_0 \cos\phi$. Replacing the distance r in Eq. 6.1, the deposition rate at the spherical surface becomes Eq. 6.3:

$$R \propto \frac{P_\mathrm{e} - P_\mathrm{c}}{4r_0^2}. \tag{6.3}$$

Now the deposition rate only depends on the radius of the sphere. The film thickness should be uniform as long as the sample is placed at a position which

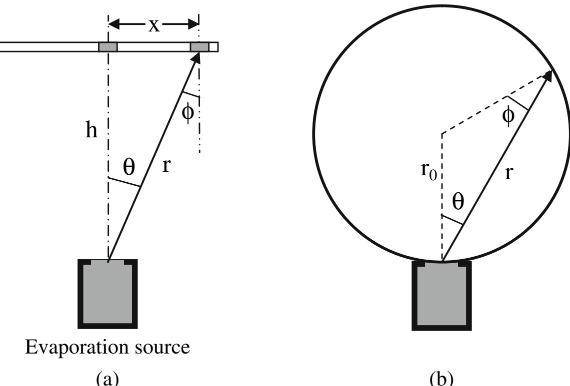


Fig. 6.2 (**a**) Thin-film deposition on a flat sample by thermal evaporation. (**b**) Spherical arrangement for uniform deposition by thermal evaporation

coincides with the surface of a sphere and the evaporation source is located at the edge of this sphere.

Another widely used thin-film deposition technique in additive pattern transfer is the plasma sputtering deposition. The sputter coating system is normally in diode configuration, as schematically shown in Fig. 6.3. Two parallel electrodes are mounted in a vacuum chamber at a distance. One is the stage on which a sample is placed and another is the target from which comes the deposition material. A DC voltage of several kilovolts is applied between the electrodes, with the sample stage as anode and target as cathode. Inert gas, normally argon, is introduced into the vacuum chamber. At high electrical field, free electrons are accelerated and collide with Ar atoms which become ionized. More free electrons are available in the chamber because of the ionization. These electrons acquire energy from the electrical field between the diode electrodes, which causes more Ar atoms to be ionized. The Ar^+ ions are accelerated toward the target, sputtering off target atoms. These free atoms with initial kinetic energy fly everywhere and deposit on the sample surface as well as other surrounding surfaces. Plasma sputtering in the DC diode configuration is characterized by slow deposition rates, high voltages, and low currents. It is also inadequate for the deposition of dielectric films because of charging, arcing, and very low deposition rates. The improvement was to replace the DC power supply, shown in Fig. 6.3, with a RF power supply, generally operating at a frequency of 13.56 MHz. This change eliminates the charging and arcing problems when sputtering dielectrics and also leads to slightly higher deposition rates. The oscillating RF potential applied to anode and cathode results in a modification to the electron motions, which provides better energy coupling to the electrons and produces higher plasma density [3].

Further evolution from the RF sputtering system was the magnetron cathode system. This type of tool uses magnetic confinement of electrons in the plasma, which results in a much higher plasma density than in either the RF or DC diode systems. The high plasma density reduces the discharge impedance and results in high current and low voltage discharge. In a magnetron, the electric field is always oriented normal to the surface of the cathode. The transverse magnetic field is configured so that the $E \times B$ drift paths form closed loops, in which the trapped and drifting electrons are constrained to circulate

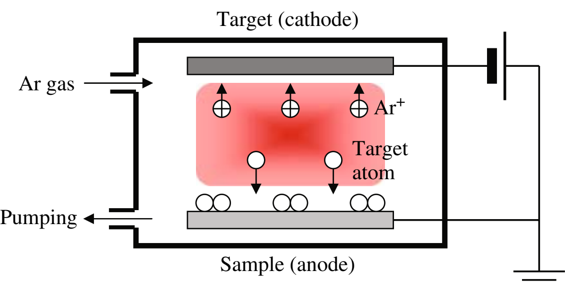


Fig. 6.3 Schematic of plasma sputtering deposition process

many times around the cathode face, which greatly enhances the Ar ionization rate hence the deposition rate [4].

The deposition by plasma sputtering is also characterized by the Knudsen's cosine law. However, various departures from the cosine distribution can happen, depending on the ion energy and the target material structure. As a function of ion energy, the sputtering rate of low-energy ions is higher at larger emission angle and the sputtering rate of high-energy ions is higher at smaller emission angle. The sputtering rate is also dependent on specific crystalline orientations of the target. The size of sputtering target is much larger than the size of an evaporation source. The sputter-off target atoms in the chamber space are frequently scattered by gas atoms before reaching the sample surface. All the above factors imply that the movement of atoms has a very broad, almost isotropic nature. This is the key reason why plasma sputtering deposition is less directional than thermal evaporation deposition.

There are ways to improve partially the directionality of sputtering deposition. Two techniques have been developed to filter out the angular distribution of the sputtered atoms to achieve a more normal-incident deposition process; one is long-throw deposition and the other is collimated deposition [4].

6.2.1.1 Long-Throw Deposition

Long-throw deposition is simply to increase the distance between the target and the sample holder plate. It is a mimic to the thermal evaporation deposition whereby the emission source is far apart from the sample. The reason for long-throw is to decrease the pressure in the plasma chamber. It is known that to sustain a plasma the electrode spacing L and chamber pressure P have to satisfy the condition of $L{\bullet}P > 0.5$ (cm•Torr). Then to maintain plasma while keeping low pressure, the electrode spacing has to increase. The above condition is for DC plasma. For RF and magnetron plasma, electrons acquire extra energy from extended circling. The plasma can be sustained at much lower pressure than in a DC system. At 25-cm throw distance, the working pressure in a magnetron system can be in the low 0.1-mTorr range. As the mean free path of an atom in ambient gas can be expressed as Eq. 6.4, low pressure means long mean free path for target atoms.

$$\text{Mean free path} \approx \frac{5 \times 10^{-3}}{P(\text{Torr})}(\text{cm}). \tag{6.4}$$

With the throw distance equal to the cathode diameter, the ballistic, unscattered atoms will have the maximum angular divergence of the about ±45°. Atoms traveling at higher angles will be unable to reach the sample, which effectively increases the directionality of deposition. The drawback of long-throw deposition is some loss of deposition material because large angular atoms deposit onto the chamber walls instead of the sample.

6.2.1.2 Collimated Deposition

Directional deposition can also be obtained by inserting an array of collimating tubes between the magnetron cathode and the sample, as schematically shown in Fig. 6.4. The collimator functions much like the sidewalls of the chamber. It collects atoms which are traveling laterally to the cathode/sample plane. These atoms are deposited on the sidewalls of the collimator cells and hence not deposited onto the sample. Geometrically, the use of a collimator need not require any significant increase in throw distance over conventional sputter deposition, thus removing the very low-pressure constraint of long-throw sputtering. The aspect ratio of the collimator can be readily changed without significantly altering its overall dimensions. For practical manufacturing applications, collimator cell diameters are of the order of 1 cm, and therefore the height of a 2:1 collimator is only 2 cm. The overall throw distance increase is then of the order of 3–4 cm. The directionality of deposition can be increased by increasing the aspect ratio of collimating tube. However, the deposition rate will be significantly reduced due to collimation. Generally, the deposition rate is reduced by a factor of about 35% for each unit increase in the collimator aspect ratio. The reduction of deposition on the sample becomes the deposition onto the collimator tubes. Eventually this leads to either closure of the collimator holes, or peeling and flaking of thick deposits, neither of which is desirable. In practice, it is necessary to change the collimator about 2–3 times more often than changing of cathode, which unfortunately can add significantly to the overall cost.

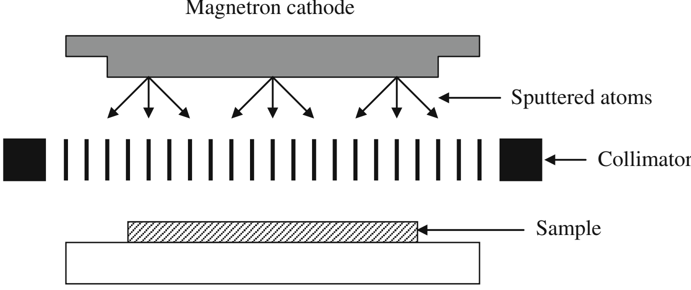


Fig. 6.4 Insertion of collimators to improve directionality in sputtering deposition

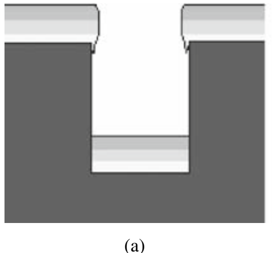
6.2.2 Pattern Transfer by Lift-off

The concept of patterning high-resolution metallic structures by lift-off was proposed nearly 30 years ago [5], and soon afterwards sub-20 nm metal lines were made by the technique using patterned poly(methyl methacrylate) PMMA resist [6]. Lift-off has now become the primary technique for patterning metallic thin films, because for many metallic materials there are no other means to pattern them except the lift-off. Lift-off process is illustrated in Fig. 6.1a. A polymer resist layer is patterned first by optical or e-beam lithography. Metallic

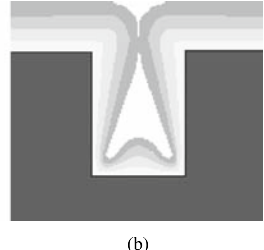
thin film is then deposited onto the patterned resist layer. A wet chemical solution, normally acetone, dissolves the resist layer, which also lifts off the metallic thin film on top of the resist layer from the substrate. Only the metallic film deposited through the resist pattern opening onto the substrate remains. In this way, the resist pattern is transferred onto the substrate as a metallic pattern of reverse polarity. The resolution capability of lift-off process is in principle the same as the resist patterning process, whether the patterning is done by optical lithography, e-beam lithography, or nanoimprinting lithography. Sub-10 nm lift-off metal patterns have been demonstrated [7]. From thin-film deposition point of view, the only limitation on the ultimate resolution of lift-off patterns is the grain size of deposited thin films. A continuous line pattern may be broken if the grain size is comparable to the linewidth. As seen in Table 6.1, each deposition technique has its own range of grain size. Pattern transfer at sub-100 nm feature dimension requires small grain size which depends on the deposition conditions. Generally, high deposition rate, whether by thermal evaporation or plasma sputtering, will result in small grain size in the deposited film.

The key to the successful lift-off is to make sure that the liquid solvent can penetrate into the resist layer and dissolve it. This requires that the metal film on top of the resist layer disconnects with the metal film inside the resist pattern opening. The discontinuity of film is only possible if (1) the deposited film thickness is much less than the resist layer thickness, and (2) the deposited film has poor step coverage over the patterned resist layer. The first condition can be easily met by using thick resist layer. For successful lift-off, the resist layer should be at least three times thicker than the deposited thin film. The second condition depends on the directionality of thin-film deposition process. As described in the previous section, thermal evaporation deposition has the best directionality than any other deposition method. Figure 6.5 compares the deposition profiles for thermal evaporation (Fig. 6.5a) and sputtering deposition (Fig. 6.5b). For this reason, whenever possible lift-off should always be done through thermal evaporation deposition of thin film.

However, in term of film quality, thermally evaporated film is the poorest among other methods. More importantly, thermal evaporation is not best suited for depositing alloy materials, because each constituent in an alloy may have different vapor pressure hence different evaporation rate. The composition of deposited film may deviate from the stoichiometry of original alloy in the evaporation source. Plasma sputtering deposition, on the other hand, has



(a)



(b)

Fig. 6.5 Profiles of thin-film coverage over a patterned resist structure with (**a**) thermal evaporation deposition and (**b**) plasma sputtering deposition

more or less the same sputtering rate for different constituents in an alloy. Although different materials have different ion sputtering yield [8], the sputtering only takes place at the target surface, which would quickly deplete the atoms of constituents with high sputtering yield, which leads to decrease of its sputtering rate until all the atoms of other constituents in the surface layer are sputtered off and a new fresh layer is revealed. In this way, the overall sputtering rates for all the constituents in an alloy become similar. Although thin film deposited by plasma sputtering is better, its poor directionality makes the lift-off much more difficult than using thermal evaporation deposition. In order to ensure the discontinuity of thin film over patterned polymer resist topography, it is highly desirable to create an undercut profile in the patterned resist structure, as is schematically drawn in Fig. 6.1a. With an undercut profile, in addition to a thicker resist layer, the poor directionality of thin-film deposition can be offset, making lift-off possible even with plasma sputtering deposition.

It is difficult to produce an undercut resist profile in optical lithography, because the absorption of photons in photoresist is always higher at top part of the photoresist layer, resulting in an overcut profile after development. In e-beam lithography, undercut resist profile is possible. As described in Chapter 3, electrons scatter in a resist layer and electron energy deposition becomes broader as electrons go deeper in resist. This feature can be confirmed by Monte Carlo simulation of electron scattering and resist exposure. Figure 6.6a is the computer-simulated resist profile exposed with 20-keV e-beam. The resist is a 0.5-μm thick PMMA and the pattern is single 50-nm line. The pear-shaped profile is the result of electron scattering. Three resist profiles are shown, representing three different exposure doses. The higher the exposure dose the broader the undercut range. However, e-beam lithography at low energy does not have high resolution, particularly for dense patterns, as demonstrated by Fig. 3.11 in Chapter 3. When patterning thick resist at high resolution, especially for sub-100 nm features, high beam energy is preferred. However, high beam energy makes it difficult to produce an undercut profile in a single-layer resist, as is seen in Fig. 6.6b. Increasing exposure dose makes little difference in increasing the undercut range. In this case, if a bilayer resist system is employed whereby the underlayer is of much higher sensitivity than the top layer, an undercut profile can be produced even at high beam energy exposure. Figure 6.6c shows the profile of bilayer resist, which is simulated with 100-keV beam energy. The top layer is 150-nm thick and the sensitivity of bottom layer is assumed to be 50 times higher than the top layer. The development rate for top and bottom layers is assumed the same. An undercut profile has been created. The commonly used bilayer resist system for lift-off is the combination of PMMA and its copolymer [9]. High–molecular-weight PMMA, which is of low sensitivity to electron exposure, forms the top layer. Low-molecular-weight copolymer P(MMA-MAA), which is of high sensitivity, forms the bottom layer. The PMMA and copolymer are both sensitive to electron exposure. They are both developed in the same developer solution. The undercut profile depends only on their difference in sensitivity.

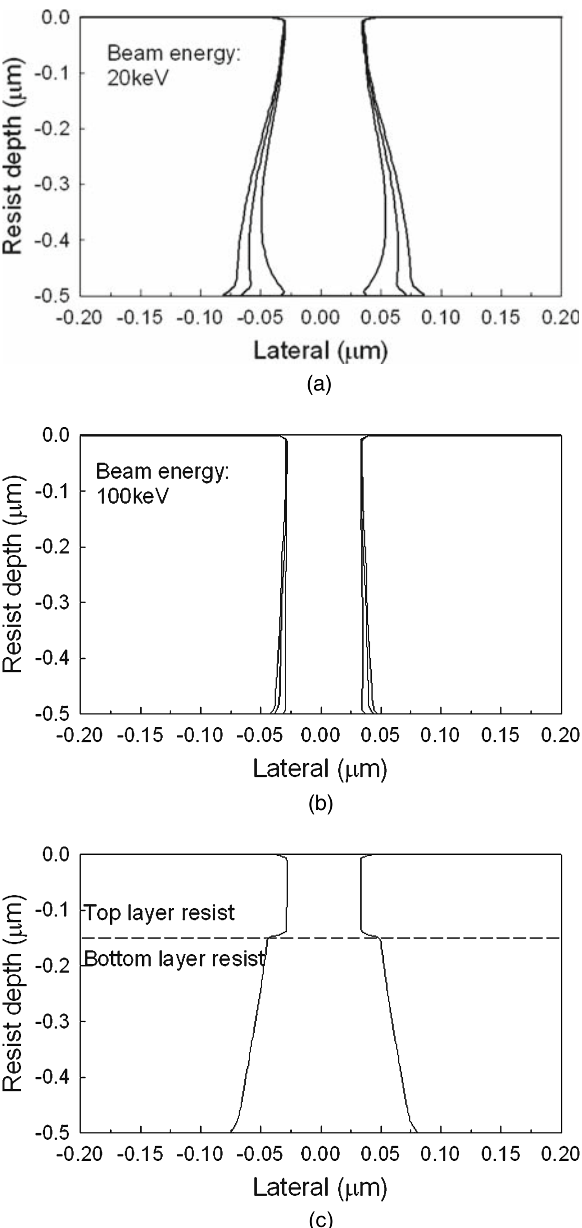


Fig. 6.6 Computer-simulated undercut profiles by e-beam exposure: (**a**) Single-layer PMMA at 20-keV beam energy, three profiles at exposure dose of 100 μC cm^{-2}, 200 μC cm^{-2}, and 300 μC cm^{-2}; (**b**) single-layer PMMA at 100-keV beam energy, three profiles exposure dose of 300 μC cm^{-2}, 400 μC cm^{-2}, and 500 μC cm^{-2}; (**c**) bilayer PMMA at 100-keV beam energy and exposure dose of 300 μC cm^{-2}

Ideally, the length of undercut should be independently controllable. Therefore, the bottom layer does not have to be an e-beam-sensitive polymer. As long as the top layer is sensitive to e-beam, resist patterning can be performed only on the top layer. Once the top layer is exposed and developed, the bottom polymer can be dissolved by a different solvent through the opening of top layer. By controlling the extent of dissolution for the bottom-layer polymer, the length of undercut can be controlled. The process is illustrated in Fig. 6.7. Since the bottom layer is not an e-beam resist, its profile has nothing to do with the e-beam exposure dose and development but is only controlled by an independent dissolving process. One of such polymers developed in recent years especially for lift-off process is the lift-off resist (LOR) [9]. Using LOR as the bottom layer in a bilayer resist stack can produce precisely controlled undercut profile as seen in Fig. 6.8a, where three undercut profiles are shown. The length of undercut is controlled only by the dissolution time. Figure 6.8b shows deposited metal thin film over the patterned bilayer resist stack before lift-off [10]. With the proper undercut profile, the metal film on top of the resist stack and at the substrate surface are well disconnected. The same concept can be applied to various other combinations of resists to allow undercut to be independently created. A recent work demonstrated a PMMA/aluminum (Al) bilayer for metal lift-off patterning, with PMMA on top as an imaging layer [11]. The Al underlayer works in the same way as the aforementioned LOR. Aluminum can be dissolved in an alkaline solution, which does not affect the PMMA pattern. The added benefit of using Al underlayer is the ability to carry out e-beam lithography on an insulating substrate, because the Al coating can effectively prevent electron charging of the insulating substrate. Silver lift-off pattern features with minimum 40-nm linewidth have been produced on a quartz substrate by this process.

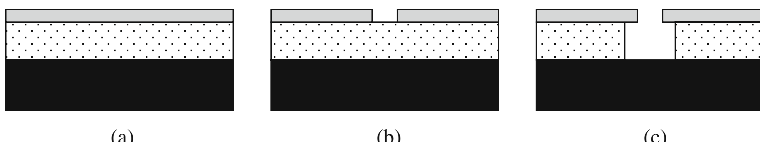
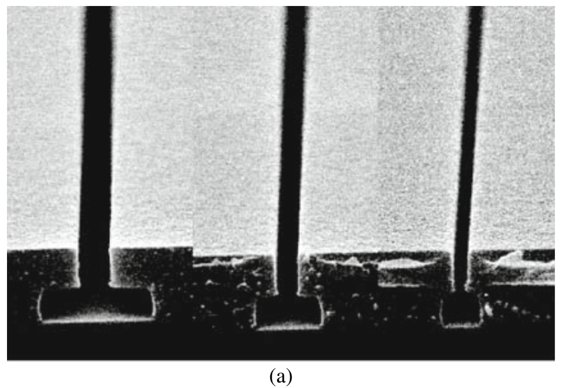


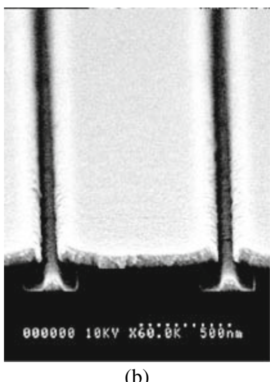
Fig. 6.7 Schematic of bilayer resist system with independent control of undercut: (**a**) bilayer resist stack, (**b**) lithography opening of top layer (imaging layer), (**c**) independent development of bottom layer to control the undercut length

Although lift-off combined with e-beam lithography has been widely used for patterning sub-100 nm metallic structures, one must be aware that with undercut profile the density of resist pattern is compromised, because there have to be adequate space between adjacent features to accommodate the undercut at the edges of a pattern feature.

It is always difficult to have a completely clean lift-off of metal film, particularly when the pattern area is small and most of the metal film covering resist layer has to be removed and cleaned. The solvent has to penetrate the



(a)



(b)

Fig. 6.8 (**a**) Undercut profiles in PMMA/LOR bilayer resist stack, (**b**) metal film deposited on the patterned bilayer resist stack before lift-off

metal layer around the gaps at resist pattern edges and dissolve the resist underneath. Once the resist is dissolved, the bulk of metal film must be cleaned away from the sample surface with the assistance of a pressure spray of solvent. The process is slow and sometimes the metal film can redeposit onto the sample surface once the resist layer is gone. If the metal film is deposited by plasma sputtering which is known to have poor directionality, the film may hang over the pattern edges here and there, resulting in many defects on the sample surface after lift-off. A new technique that has been developed in recent year is to use a stream of solid CO_2 particles (dry ice), or called carbon dioxide snow, to remove the metal film on top of resist layer prior to a solvent stripping off the resist [12]. The principle is to create a thermal expansion mismatch between the metal film and the resist layer, because solid CO_2 is formed at temperature below –60°C. This can cause the metal film to crack and delaminate. The loosened metal film is then cleared away by a CO_2 snow jet with high speed up to 40 m s^{-1}. Then the resist layer is stripped off in a solvent. The sample backside is usually heated to + 60°C during the process to promote the thermal mismatch and delamination of metal film. Because the metal film is cleared away before the resist layer is dissolved in solvent, the sample surface is well protected by the resist layer, which can prevent any metal film patches from redepositing onto the sample surface.

Most of nanoscale metal pattern lift-off was done by e-beam lithography of positive resist, notably the PMMA resist of either single layer or multilayer. However, there are applications requiring negative-tone lift-off, whereby metal thin film is required to cover majority of a substrate surface. In this case, negative-tone resist process is preferred because lift-off by positive-tone resist such as PMMA will require exposure of large areas and consume long e-beam lithography time. However, negative resist is known to produce overcut profiles, which make lift-off very difficult. Although some negative photoresists

have been used for negative-tone lift-off process and undercut profiles can be produced by overdevelopment of the resist structure [13], the process is, however, only suitable for much larger pattern dimension. For high-resolution negative-tone lift-off, the bilayer concept is still applicable. A hydrogen silsesquioxane (HSQ)/PMMA bilayer resists process has been developed [14], as illustrated in Fig. 6.9.

Hydrogen silsesquioxane is spin-coated on top of PMMA, and e-beam lithography is performed on the HSQ layer. After development HSQ negative-tone resist patterns are formed. The PMMA will not be affected during the development of HSQ top layer even though it has been exposed by e-beam irradiation through the HSQ layer. The HSQ patterns are then used as an etching mask to transfer into the PMMA underlayer. By controlling the etching parameters, undercut profile of the underlayer PMMA resist can be

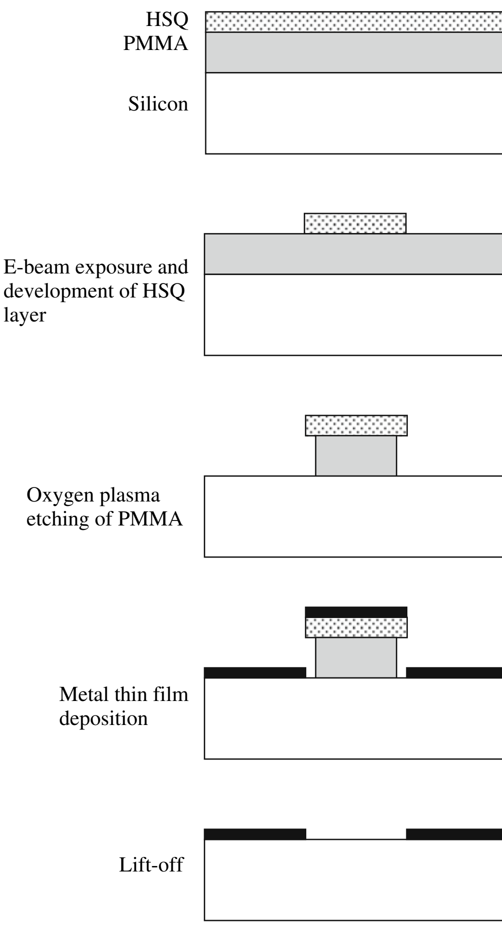


Fig. 6.9 Schematic of negative-tone lift-off process using HSQ/PMMA bilayer resist

controlled independently. Figure 6.10 shows the scanning electron microscope (SEM) image of HSQ/PMMA resist profile after etching of PMMA underlayer. The negative-tone undercut profile is clearly demonstrated. Metal lift-off can be easily done with resist pattern structure. Figure 6.11a shows chrome metal pattern in which each metal line is separated by 50-nm gap which is formed by lift-off of HSQ/PMMA resist lines. Figure 6.11b shows a single 100-nm diameter hole in a chrome metal film formed from lift-off of HSQ/PMMA resist dot [14].

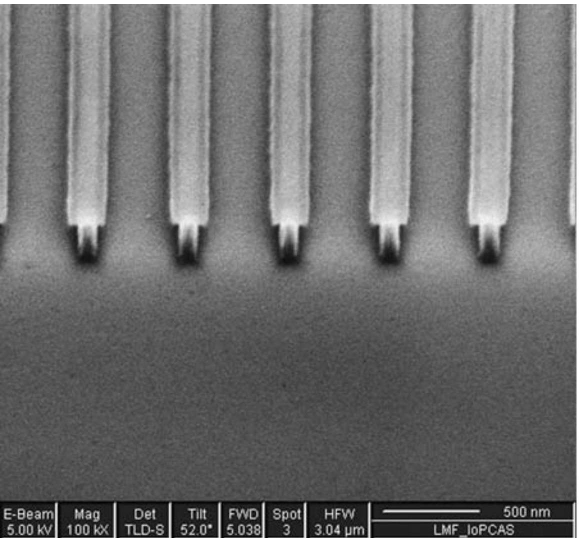
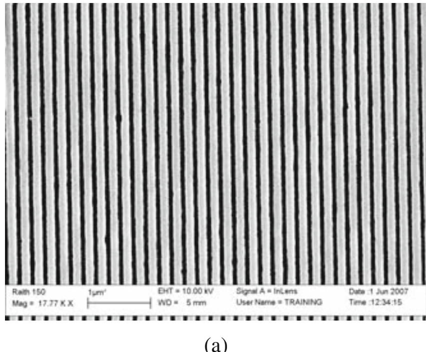
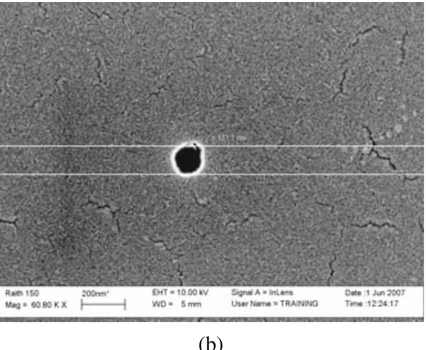


Fig. 6.10 Undercut profile formed from HSQ/PMMA bilayer resist patterning



(a)



(b)

Fig. 6.11 Negative-tone lift-off patterns: (**a**) 50-nm gaps between chrome, (**b**) 100-nm single hole in chrome film

6.2.3 Pattern Transfer by Plating

Despite its many advantages in creating metallic patterns, lift-off technique is not suitable to produce high-aspect-ratio pattern structures. One reason is that the deposited film thickness has to be at least less than one-third of resist depth in order to ensure discontinuity of film, hence a successful lift-off. The other reason is that gradual accumulation of deposition material at top of resist pattern edge can close up the opening if the resist structure is too narrow and too deep, as illustrated in Fig. 6.5b. High-aspect-ratio metallic micro- or nanostructures are produced by plating, either electroplating or electroless plating.

Plating is a well-established industrial manufacturing process. Building microstructures by plating started in 1970s when magnetic storage devices were made by plating copper and permalloy through molds [15]. The molds were photolithography-defined resist patterns. For electroplating, a seed layer is needed and electroplating of metallic patterns is done by patterning a resist layer on top of the seed layer and then filling up the patterned resist mold with metal, as illustrated in Fig. 6.1b. Electroless plating is mainly used for conformable coating of a surface. Only a limited number of metals and alloys can be deposited by electroless plating and the deposition rate is much lower than electroplating [16]. In contrast, many metals and alloys can be electroplated. The major applications of electroplating in microstructure patterning are in thin-film head manufacturing for magnetic storage devices [17] and copper filling of via hole for integrated circuits (ICs) [18]. Many other areas are also using electroplating as an effective method to build high-aspect-ratio metallic microstructures. Notably, electroplating is one of the main fabrication techniques to make micro-electro-mechanical systems (MEMSs). Electroplating, or electroforming, is one of the key components in the LIGA technique which is a German acronym of "Lithographie (LI) Galvanoformung (G) Abformung (A)", meaning lithography + electroforming + molding [19]. By hard X-ray exposure of very thick polymer resist, microstructures with high aspect ratio have been made by electroplating nickel into the resist mould [16]. Thick photoresist can also be patterned by UV light to form the mould. Though the aspect ratio is not as high as those made by X-ray LIGA, the UV-based process is a low-cost version of LIGA and particularly useful for making micromechanical structures [20]. Another main application area where electroplating plays a major role is making various magnetic microdevices by electroplating of magnetic materials, notably the permalloy. Microstructures made of permalloy can be actuated and manipulated by external magnetic fields, for example, magnetic micromirrors which can be deflected by applying a magnetic field externally [21]. The magnetic force exerting on the magnetic microstructures is proportional to the volume of magnetic material. Electroplating is the effective way to build-up bulk magnetic materials.

The aforementioned applications using electroplating are mainly for making microscale structures with dimensions larger than a micrometer or a few hundreds of micrometers. Electroplating is being used for making sub-100 nm structures primarily in two areas: first area is in making masks and zone plate lens for X-ray lithography or X-ray imaging; second area is in making stamps for nanoimprinting. For X-ray lithography, heavy metal structures act as absorbers in an X-ray mask. They have to be thick enough to absorb all the X-ray radiation, as have been described in Chapter 2 of this book. As X-ray is 1:1 proximity imaging, to pattern sub-100 nm structures by X-ray lithography, the absorber pattern on the mask has to be sub-100 nm as well. Micro zone plate (MZP) is a common type of imaging lens to focus X-ray radiation. The resolution of a MZP lens can be expressed as Eq. 6.5 [22], which is similar to that of optical imaging (Eq. 2.3 in Chapter 2):

$$R_x = k_1 \frac{\lambda}{\mathrm{NA_{MZP}}}, \quad \text{where } \mathrm{NA_{MZP}} = \frac{\lambda}{2\Delta r_{\mathrm{MZP}}}. \tag{6.5}$$

λ is the X-ray wavelength, $\mathrm{NA_{MZP}}$ is the numerical aperture of the zone lens and Δr_{MZP} is the outmost zone width. k_1 is illumination-dependent constant. Then the resolution becomes $2k_1\Delta r_{\mathrm{MZP}}$. With an optimized X-ray illumination system, the k_1 factor can be less than 0.5 [23]. It is apparent that the imaging resolution is solely dependent on the zone width. The smaller the zone width is, the higher the imaging resolution. Making high-resolution MZP lens is a great challenge and electroplating is the only method to achieve high-aspect-ratio (requirement of opacity to X-ray) heavy metal structure.

The material for high-aspect-ratio electroplating at sub-100 nm dimension is mainly gold. Electroplated gold is in the form of polycrystalline structure which can have many grain boundaries and imperfections. The polycrystalline grains may grow faster in a particular orientation. For example, gold plating is faster in the <111> direction [24]. The grains usually grow perpendicular to the surface. Therefore, surface roughness is the common feature of electroplated patterns. Figure 6.12 shows a part of MZP lens which is made from electroplated gold. The outmost linewidth is 25 nm. Polycrystalline grains are visible from the surface of MZP. If the resist pattern dimension is comparable to the grain size, voids or even break can happen to the electroplated lines if the height of feature is much greater than its width [24]. Figure 6.13 shows a MZP with outmost nominal linewidth of 15 nm. As seen from the inset SEM image, the gold lines are wider than its nominal 15 nm width, which is the result of proximity effect in e-beam lithography. In addition, some of the gold lines are visibly broken, which is most likely due to gain size effect in gold electroplating.

Another application requiring nanoscale electroplating is to make nanoimprinting stamps. The fabrication process has been described in Chapter 5. It starts with e-beam lithography of PMMA resist on standard silicon (Si) wafer.

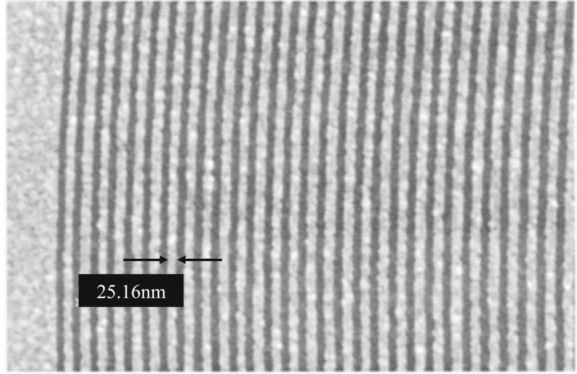


Fig. 6.12 Zone plate with 25-nm electroplated gold lines in the outmost zone (Reprint from [25] with permission)

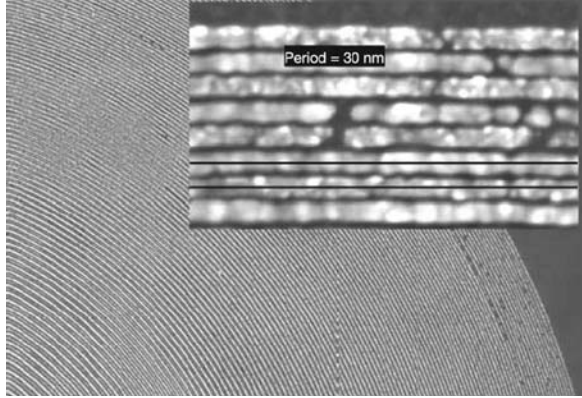


Fig. 6.13 Zone plate with 15-nm electroplated gold lines in the outmost zone (inset image showing detail of 15-nm gold lines) (Reprint from [23] with permission)

After development of exposed PMMA, a thin metal layer (10–20 nm) of either gold or nickel is coated to cover the whole surface of PMMA layer. Nickel is the metal for making the stamps, because electroplating of nickel is a very mature technology and the electroplating process is nontoxic and low cost. The material has sufficient hardness to serve as a nanoimprint stamp. After filling up the PMMA pattern cavities by electroplating of nickel, the electroplating continues until building up the total thickness of nickel sheet to 100–200 μm to make it mechanically robust. After removing the PMMA mold and silicon-supporting substrate, the nickel sheet is mounted onto a metal holder plate to become a nanoimprint stamp. A large size of stamp can be made by step-and-repeat of stamping a small stamp over a large size of substrate, which is similar to the working of optical stepper. Full 4″-size nickel stamp has been made by the step-and-stamp nanoimprinting with a 3×3 mm^2 silicon master stamp [26].

A critical issue in pattern transfer by electroplating is the uniformity of electroplated structures. As the deposition rate in electroplating is highly dependent on plating current, the uniformity of deposition thickness depends on the uniformity of current density distribution. The current density is always

higher at the location where there are asperities such as the boundary of a substrate or sharp edge of a pattern structure. The consequence of such nonuniform distribution of current density is the nonuniformity of the thickness of electroplated features. Generally, large electroplated features may be thicker at the edges. Isolated narrow lines are thicker than densely packed narrow lines [16]. For a large substrate with many different sizes of features, it often happens that some features have reached the correct thickness while others may have been overplated. Such a phenomenon is less an issue for electroplating of nanoimprint stamps, because the stamp is made by overplating nickel on top of a patterned nanoscale mold. Unlike a zone plate which requires high-aspect-ratio metallic structures, the sub-100 nm pattern structures in a nanoimprint stamp are always of low aspect ratio.

6.2.4 Damascene Process

An important additive pattern transfer process in IC manufacturing is the damascene process. The name “damascene” came from its similarity to ancient jewellery-making process which inlays metal wires and patterns. Integrated circuit manufacturing involves multilevel interconnecting transistors to make a circuit functioning. Interconnects were done by depositing a metal film (Al or Cu) and patterning the metal film into connecting lines. Between different levels of interconnects, via holes are etched and filled with metal to connect the top and bottom levels of metal lines. The process is schematically shown in Fig. 6.14, which is a dual-damascene process. In single-damascene process, via hole filling and metal line patterning are done separately. The dual-damascene process uses a single metal deposition process to simultaneously form the metal lines and metal vias [27].

The single-damascene process is suitable for Al interconnects because Al can be easily etched by reactive-ion etching (RIE). With interconnects moving from Al to Cu, the poor dry etchability of Cu made the single-damascene process unattainable. In the dual-damascene process, Cu filling can be done by PVD, notably the plasma sputtering deposition, or by CVD which has been a well-established process for filling Al vias, or by electroplating. Chemical vapor deposition process is slow in deposition rate. Plasma sputtering deposition is difficult to fill deep via holes for the reason as illustrated in Fig. 6.5b; wide angular spread of depositing atoms can close up the top opening of deep holes, leaving voids buried underneath. Electroplating copper is a low-cost and fast process but requires deposition of a seed layer first. To deposit seed layer into deep holes by plasma sputtering is also a challenge. Instead, a CVD copper deposition is an attractive process for seed layer deposition [18]. Using the dual-damascene process, multilevel interconnects can be achieved. Figure 6.15 demonstrates the multilevel copper interconnects which are revealed after the low-k dielectric materials have been etched away.

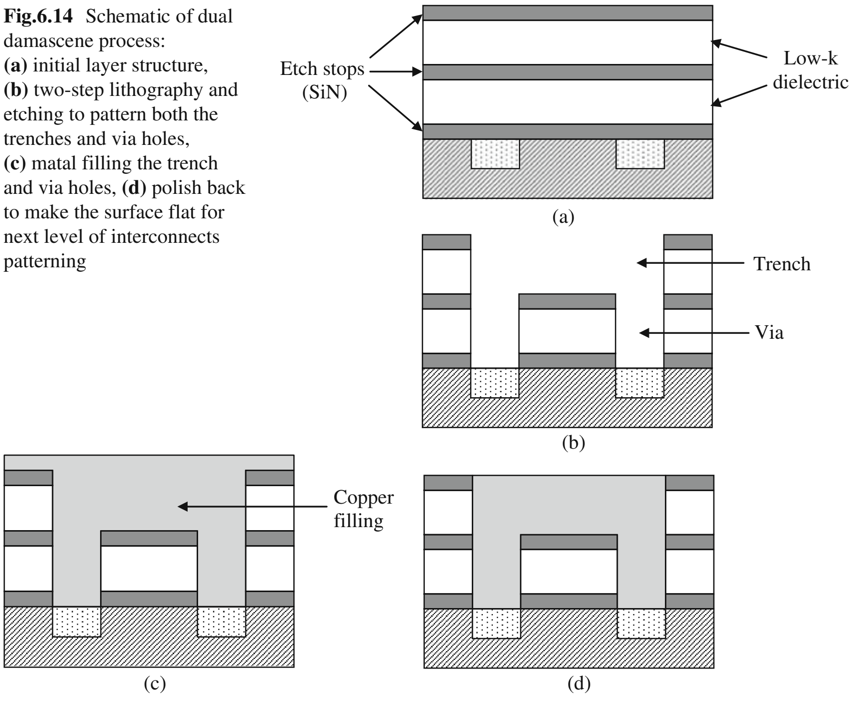


Fig.6.14 Schematic of dual damascene process: **(a)** initial layer structure, **(b)** two-step lithography and etching to pattern both the trenches and via holes, **(c)** matal filling the trench and via holes, **(d)** polish back to make the surface flat for next level of interconnects patterning

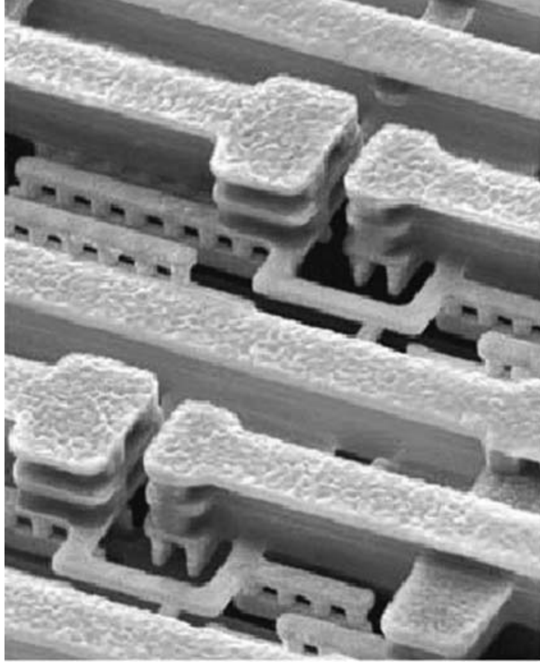


Fig. 6.15 Multilevel copper interconnects in an IC after removal of burying dielectric materials (Reprint Courtesy of International Business Machines Corporation, copyright 2008 © International Business Machines Corporation)

6.2.5 Pattern Transfer by Stencil Mask

Metallic patterns can also be transferred directly onto a substrate by a stencil mask. The process is illustrated in Fig. 6.16. A stencil mask is a framed membrane with etch-through patterns. In Section 3.6 of Chapter 3, stencil mask has been introduced as a mask for charged particle projection lithography (Fig. 3.34a). Pattern transfer by stencil mask is to use a stencil mask to partially block material deposition on a substrate. Deposited material forms a pattern on a substrate, which is the same as the etch-through pattern on the stencil mask; or in other words it leaves a shadow of the mask on substrate during material deposition. Hence, it can also be called shadow mask patterning. To ensure the shadowing effect, the deposition has to be directional. Therefore, thermal evaporation deposition is the only viable deposition technique to combine with the stencil mask pattern transfer process.

Pattern transfer through stencil mask is also called nanostencil lithography which was first proposed in 1999 [28]. Originally, it combined with a scanning probe microscope to deposit sub-100 nm metal nanostructures. Later it developed into a large area patterning technique for making nanostructures [29, 30]. Using stencil mask for making nanostructures has the following advantages:

- A thin film pattern can be deposited at any local position, because the stencil mask is not in contact with substrate. Pattern deposition can also be done on topographical surface.
- Thin film patterns are formed in one step, without resist pattering and lift-off processes like conventional lithography does. Because no resist patterning is

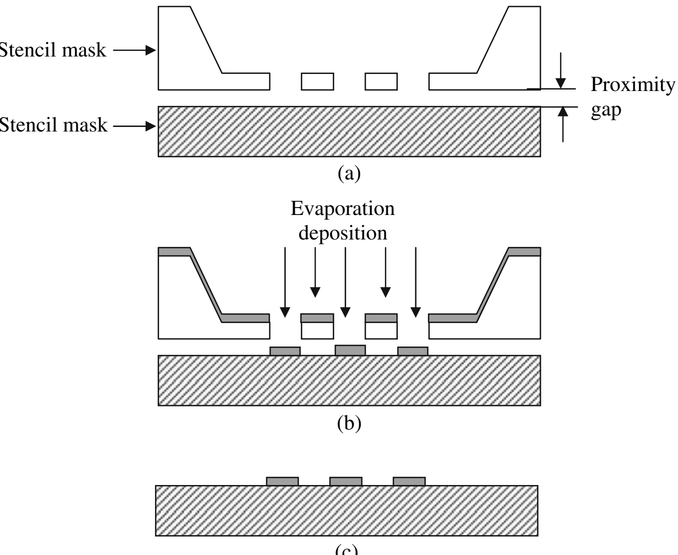


Fig. 6.16 (**a**) Stencil mask and substrate configuration, (**b**) thermal evaporation deposition through stencil mask, (**c**) pattern transfer to substrate

needed, material can be deposited directly on fresh substrate surface, avoiding potential surface contamination. As the deposition is carried out in vacuum, the possibility of substrate surface oxidation can also be excluded. Mechanically fragile or chemically functionalized surface can be structured due to the noncontact and process-free nature of the pattern transfer.
- Because no optical or e-beam lithography is involved, there is no proximity effect caused either by light diffraction or electron scattering. Large and small patterns can be formed simultaneously. For example, it has been reported that on the same stencil mask both sub-100 nm features and above 100 μm features were made in one-step deposition [30].

However, when transferring patterns by stencil masks, some caution has to be taken as there are limitations in this technique:

- The technique can only deposit single-layer pattern. For multilayer deposition, there is an alignment issue, unless the multilayer materials are deposited on the same location
- For thick-film deposition or repeated use of stencil mask, the deposition of material around the edge of stencil pattern can close in, resulting in smaller feature dimension on the mask, and eventually clog some small holes completely. Figure 6.17a shows schematically the ideal deposition without clogging and Fig. 6.17b is the deposition with clogging at the stencil mask aperture. The eventual close up of mask aperture is demonstrated by the SEM image of real stencil mask in Fig. 6.17c. Nevertheless, it is possible to remove clogging by wet chemical or RIE each time after use. It is the clogging during the deposition that imposes a limit on the amount of material to be deposited in a single process step
- Because there is a gap between stencil mask and substrate surface, there are always blurs like halos around the main pattern edge, due to molecules arriving in angular direction through the stencil mask opening. Further blurs are caused by kinetic movement of molecules, because the molecules evaporated thermally from the material source have kinetic energy and will move laterally before landing on the substrate surface. The blur, however, is much thinner than the main deposited pattern structure and can be corrected by an extra etching after the stencil mask pattern transfer [31]
- For large-size stencil mask, there is a thermal stability issue. Deposition of metal by thermal evaporation can generate a lot of heat from evaporation source. Radiation of heat can cause expansion of stencil mask, causing pattern shift and distortion. The deposition-induced stress can also deform the stencil membrane
- The fabrication of stencil mask involves other lithographic techniques and RIE, which limits the smallest feature achievable on the mask. For large-area stencil mask, the minimum mask features are around 100 nm. For small-size stencil mask, FIB sputtering can be used to make very small holes in membrane.

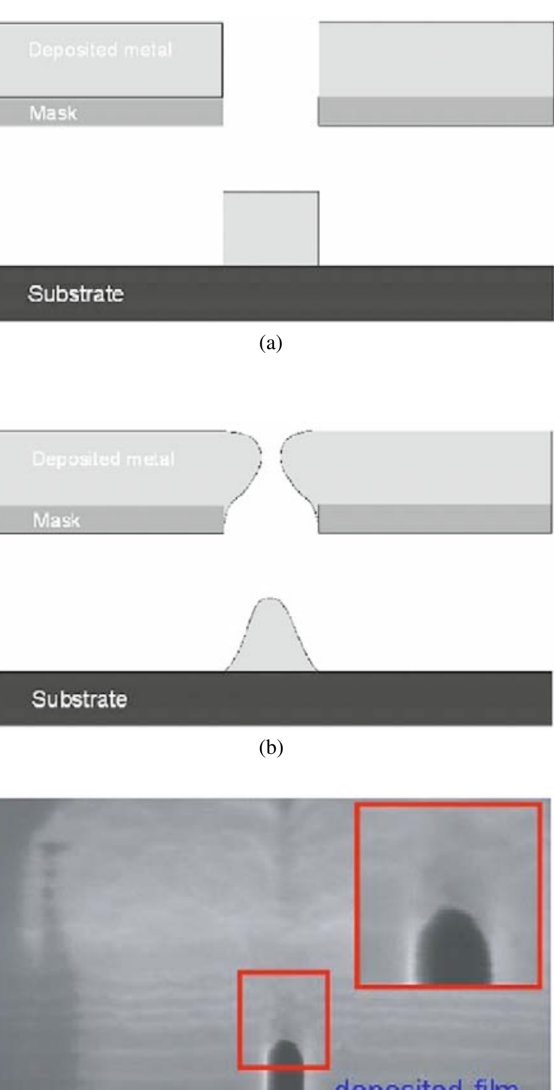


Fig. 6.17 (**a**) Ideal deposition through stencil mask aperture without clogging, (**b**) deposition with clogging, (**c**) SEM image of a stencil mask aperture which is closed up due to excess deposition (Reprint from [32] with permission)

6.3 Subtractive Pattern Transfer

Subtractive pattern transfer is mostly characterized by various etching techniques, whether wet or dry etching. In wet etching processes the etchants are liquid form of chemicals. In dry etching, the etchants are gas form of chemicals, or in some cases are energized ions. Lithography and etching are the two most important processes which go hand-in-hand in most micro- and nanofabrications. The main character of IC manufacturing is the repeated processes of lithography and etching, which lead to the buildup of multilevel circuit pattern layout on silicon. Making any other micro- and nanostructures, as long as processing by planar technology, cannot escape either one or another form of lithography and etching process. Etching has been a subject in many books and publications, including the author's previous book [33]. It is not the intention of this section to repeat what has already been described on various etching technologies. Here only the techniques that enable pattern transfer by etching at sub-100 nm dimension are introduced.

No matter wet or dry etching, there are two parameters which characterize all the etching techniques.

1. *Selectivity*. Etching is a process of removing substrate material through mask openings. The mask can be a polymer resist, or another thin film material such as silicon dioxide and silicon nitride. The thin film mask is also made by lithography and etching. Selectivity is the ratio of substrate etching rate over mask etching rate. High selectivity means there is little loss of masking material and the mask can endure long etching in an etchant.
2. *Anisotropy*. The purpose of etching is to transfer the mask pattern into a substrate. Ideally, etching should transfer the exact dimension of mask pattern. In practice, there are always distortions. The distortion can be caused by excessive etching of masking layer. This is to do with the aforementioned selectivity issue. The distortion can also be the result of lateral etching of the substrate, which is qualified by anisotropy. Figure 6.18 illustrates different degrees of anisotropy in etching. It ranges from isotropic etching, to partially anisotropic etching, and to completely anisotropic etching. In general, anisotropic etching transfers mask pattern in its exact dimension into a substrate. However, there are exceptions that anisotropic etching may not follow the mask pattern boundary, such as the alkaline-based wet chemical etching of crystalline silicon. On the other hand, isotropic etching may still be able to transfer the exact dimension of a mask pattern, if the etching depth is extremely shallow.

6.3.1 Isotropic Wet Etching

Wet chemical etching is generally isotropic; that is, etching progresses in depth as well as in lateral direction. Even though there is a mask, liquid etchant can penetrate underneath the mask, resulting in the etched pattern larger than the

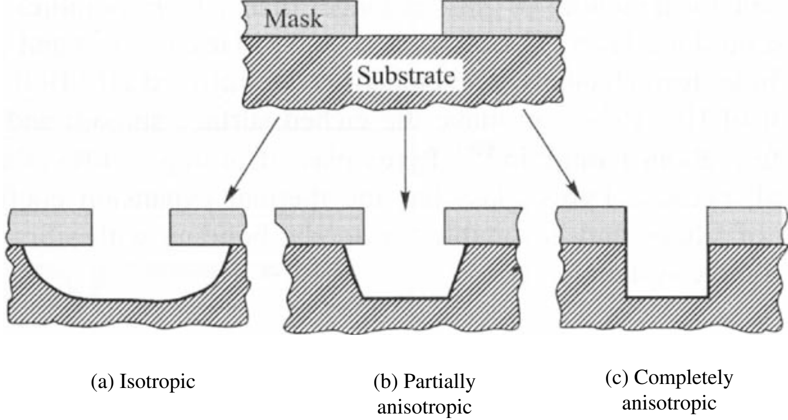


Fig. 6.18 Schematic of anisotropy of etching: (**a**) isotropic, (**b**) partially anisotropic, and (**c**) completely anisotropic

mask pattern, as illustrated in Fig. 6.18a. For this reason, wet chemical etching is normally not used for high-resolution pattern transfer. A typical example is the fabrication of photomasks. Photomasks for above 0.25-μm IC technology are made by wet acidic etching of chrome layer using a patterned polymer resist as masking layer. For below 0.25-μm IC technology, various resolution enhancement techniques are being used, which include phase-shifting masks (PSMs) and optical proximity correction (OPC) masks. Phase-shifting and OPC masks both contain sub-resolution pattern features which are in the order of 100 nm. Wet chemical etching is no longer capable of transferring these sub-resolution features faithfully into the chrome layer; RIE has to be used.

However, the etching in lateral direction is only as fast as in depth direction. If the etch depth is very small, the lateral etching can be negligible. Therefore, high-resolution pattern transfer is still possible. Wet chemical etching has indeed found its way into sub-100 nm pattern transfer applications when self-assembled monolayer (SAM) is used as a resist. Not all SAMs are qualified as a resist for nanoscale pattern transfer. The only SAM that has found applications in nanolithography is the family of alkanethiol SAM systems. There are two unique features that make the alkanethiol SAM systems suitable as a resist: one is its fast self-assembly process on the surface of coinage metals, notably the gold surface; the other is its very high etching selectivity in gold etchant. For example, gold surface protected by a monolayer of hexadecanethiol (HDT) is 10^6 times slower in cyanide (gold etchant) [34]. Patterning of SAM can be done by microcontact printing (μCP), as described in Chapter 5, or by a scanning probe (Dip-pen nanolithography) as described in Chapter 4. The layer thickness of a SAM pattern is only 1–2 nm. Transfer of SAM pattern has to be carried out in wet chemical etching for two reasons: one is that the SAM is so thin that it cannot withstand ion bombardment in a RIE process; the other is that there are no effective RIE processes suitable for

etching coinage metals. To realize sub-100-nm-scale pattern transfer in wet chemical etching, one needs to make the metal film extremely thin. Lines of 20-nm dimension have been produced in 10-nm-thick gold film using a STM patterning of HDT SAM and wet chemical etching of the gold film [34]. Although only coinage metals can be patterned using the SAM resist as mask, the pattern can be further transferred into other materials using the patterned metallic film as a mask. For example, various silicon structures have been generated by anisotropic etching of silicon with patterned gold or silver as the etching masks [35].

Like the aforementioned SAMs, polymers which are mainly hydrocarbons are generally resistant to acidic solutions. If a polymer resist is used to make the mask patterns for subsequent acidic etching, one should be aware that even a very thin layer of polymer resist residue (scum) can stop the etching of substrate. Therefore, a descum process usually follows the development of resist pattern before proceeding to acidic etching of substrate. Descum is normally done by a short oxygen plasma etching of developed resist pattern, assuming that the residue layer is very thin. For example, for wet etching pattern transfer in photomask-making process there is always a descum step before the etching of chrome layer. The descum etching by oxygen plasma can, however, affect the resist pattern profile. So it has to be brief, usually less than 30 s at low plasma density. Another caution is the effect of descum on substrate surface. In principle, the descum etching should not damage the underlying thin film or substrate surface.

6.3.2 Anisotropic Wet Etching

Although wet chemical etching in general is isotropic, an exception is silicon etching by alkaline-based etchant, which is dependent on crystalline orientation, therefore, is anisotropic. The anisotropy comes from the fact that alkaline echant dissolves silicon at different rates at different crystalline planes. The most commonly encountered crystalline planes of silicon wafers are <100>, <111>, and <110>. The etch rate decreases in the order of <110>, <100>, and <111>. The etch rate in <110> plane is two times higher than the <100> plane and they both are hundreds times higher than the <111> plane. Typically, the ratio of etch rates for the three crystalline planes <110>/<100>/<111> is 400/200/1 [36]. Because different crystalline planes intersect at different angles, the etch profile of silicon in alkaline etchant can be quite complicated. For example, etching of <100>-oriented silicon wafer proceeds fast in <100> direction but slow in <111> direction. Because the <111> plane intersects <100> plane at an angle of 54.74°, the etched profile is sloped, as illustrated in Fig. 6.19a. For <110>-oriented wafer, some of the <111> planes are perpendicular to the <110> plane. There is a direction in which the etched profile has vertical sidewalls, as illustrated in Fig. 6.19b. To obtain vertical sidewalls, at least two parallel boundaries of a mask pattern have to align to the <111> plane. Figure 6.20 shows

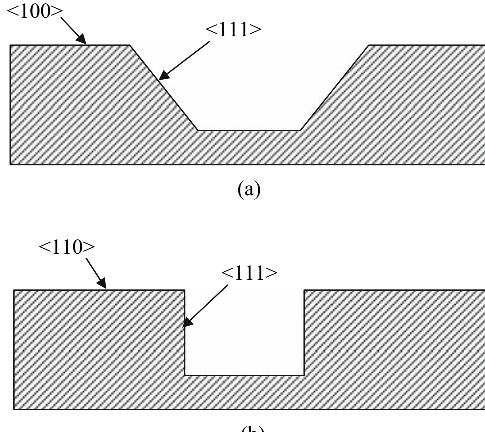


Fig. 6.19 Alkaline etching of silicon wafer in (**a**) <100> orientation and (**b**) <110> orientation

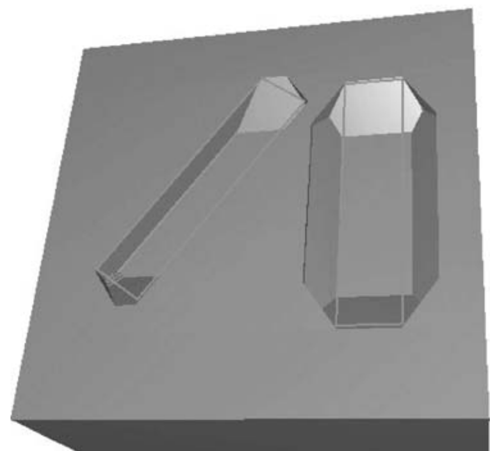


Fig. 6.20 Computer-simulated alkaline etching of <110> silicon wafer with mask patterns in different orientations

computer-simulated KOH etching of <110>-oriented silicon. The simulation was done using the AnisE software which is a part of IntelliSuite software package [37]. There are two mask patterns: one is aligned and another is not. The etched profiles of these two mask patterns are completely different.

Alkaline etching of silicon is a bulk silicon micromachining process which is generally used to create large structures of tens or hundreds of micrometers. With a <110> silicon wafer, the technique can be employed for nanoscale pattern transfer as long as the pattern edges are aligned to certain crystalline planes and the etch depth is limited. Line patterns less than 40 nm have been transferred into silicon by alkaline wet etching on <110> silicon wafer, which creates vertical sidewall profile with more than 5:1 aspect ratio and nearly atomic smooth sidewall surface [38]. In conventional bulk micromachining of

silicon, a large quantity of silicon has to be etched away. A good masking material with high selectivity is necessary. Silicon nitride is preferred because it has higher etch resistance in alkaline etchant than silicon dioxide. For nanoscale wet etching of silicon, the selectivity of masking material is less of an issue because the etch depth is so small. Oxide mask is generally used because an oxide mask pattern can be directly generated by local oxidation using the scanning probe in an atomic force microscope [39]. Otherwise, an oxide mask pattern is generated by e-beam lithography of PMMA resist and subsequent isotropic etching by buffered HF or conventional RIE process [40]. Only a very thin oxide masking layer is needed for such nanoscale etching.

There are three common alkaline etchants for wet chemical etching of silicon: the potassium hydroxide (KOH), ethylenediamine pyrocatechol (EDP), and trimethyl ammonium hydroxide (TMAH). For etching large structures, all the three etchants are fine. There are some differences in selectivity and etch rate. For example, KOH has higher silicon etch rate than the other two etchants, but also higher etch rate to oxide. More importantly, KOH etching of silicon creates rough surface, which is highly detrimental when the etched structure itself is at sub-100 nm dimension. Ethylenediamine pyrocatechol has lower oxide etch rate and is less anisotropic because it etches <111> plane faster than KOH does, but it is very toxic and special caution has to be taken when using it. Trimethyl ammonium hydroxide (TMAH) is the best suitable alkaline etchant for nanoscale etching of silicon because the etched surface is 10 times smoother than etching by KOH. Grooves 40-nm deep with 50-nm width and 30-nm high ridge with 40-nm width have been fabricated on Si<100> surface with TMAH etchant [41]. The other option is to use a mixture of KOH and isopropyl alcohol (IPA), which can also significantly reduce the surface roughness [40]. The additional benefit of using TMAH is that it does not contain alkaline metals like the KOH; therefore, it is safe to use in an IC manufacturing environment.

6.3.3 Reactive-Ion Etching (RIE)

Reactive-ion etching is the most widely used pattern transfer technique in micro- and nanofabrication. The most distinctive feature of RIE is that the etching takes place in plasma. There are no wet chemicals involved; therefore, the other name for RIE is dry etching, though dry etching includes much broad range of plasma-based etching techniques and RIE is only one of them [33]. Reactive-ion etching involves both physical and chemical processes. The physical process includes ionization of reactive gas molecules and ion sputtering of sample surface, both of which take place in a plasma. Plasma can be loosely defined as partially ionized gases. In a confined space such as a vacuum chamber, the ionized particles collectively set up a long-range electrostatic field which has important influence on the behaviors of these charged particles. The electric field is created because electrons in the plasma tend to move much faster than ions. The fast-moving electrons hit the chamber wall before ions do, causing the wall to be charged up

negatively. The negatively charged wall surface pushes other electrons away from the wall, creating a layer of positively charged region where an electrical field is created to accelerate the ions toward the wall surface [42].

In a practical RIE system with planar electrodes configuration, as schematically shown in Fig. 6.21, one of the electrodes is the cathode which is connected to RF power source by capacitance coupling. The other electrode is the anode which is grounded, including the reaction chamber walls. The RF power source provides a high-frequency electrical field between the anode and cathode, which accelerates and resonates a small number of free electrons. These electrons acquiring enough kinetic energy in the RF field will collide with gas molecules and ionize the gas molecules. More free electrons become available and more gas molecules are ionized, leading to an avalanche of free electrons and gaseous ions. The gas-filled space between cathode and anode changes from initially an insulator to a conductor. Electrical current flows through the space between the two electrodes. The electrical potential distribution between cathode and anode is shown in Fig. 6.21. A large part of the region is equipotential region, which is also called glow region, where electrons and ions recombine to form molecules. The gas molecules returning from excitation state to normal state through the recombination process release their energy in the form of photons. A glowing light is seen in the plasma chamber. Local electrical fields near the cathode and anode are formed because of the fast-moving electrons and charging of the electrode surface. The region near the

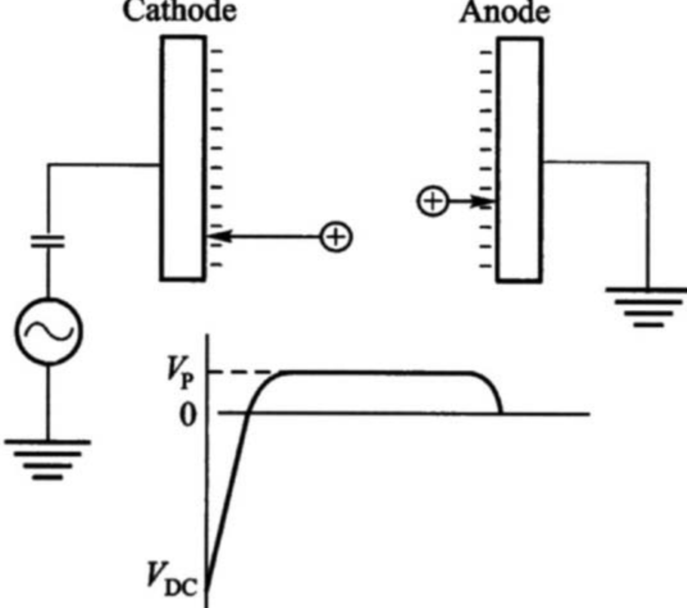


Fig. 6.21 Planar electrodes plasma system and potential distribution

cathode which is not grounded establishes higher field than that of the anode, which accelerates ions toward the cathode. A sample placed on the cathode electrode will experience bombardment of high-energy ions.

Although there is an electrical field which accelerates ions toward the sample surface, the physical sputtering is not a dominant phenomenon in RIE. The dominant etching process is caused by chemical reaction. Most RIE processes use halogen-containing gases, fluorine (F), chlorine (Cl), bromine (Br), and iodine (I), to generate the plasma, because they are the most chemically reactive and erosive gases to various inorganic materials. Fluorine is widely used for etching silicon and silicon compounds. Chlorine is mainly used for etching Ga-, Al-, and As-containing III–V compounds. For InP and other indium (In)-containing compounds, such as InGaAs and GaInAsP, the volatility of etched product in Cl plasma at room temperature is low. Either the substrate has to be heated to more than 150°C or Br or I-containing gases are used (but such gases tend to attack the vacuum system aggressively) [43]. There are, however, cases that non-halogen-containing gases (hydrogen or hydrogen/methane mixture) were used to etch compound semiconductors such as GaAs and InP [44]. The halogen compounds that can be used in RIE and the materials which can be etched are listed in Table 6.2 [45]. One can notice that for etching of some of the materials, there are several choices of etching gases. The gases listed in Table 6.2 only serve as references. There can be many more choices. Each gas has its own advantages and drawbacks. The choice of a particular etching gas is very much dependent on the requirements of a specific application. For example, F compounds and Cl gas or compounds can both be used for etching silicon. Flourine atoms react with Si spontaneously and they induce isotropic etching and cause undercutting below the mask. Chloride etching is anisotropic since it is an ion-assisted process rather than a spontaneous etching process. However, Cl_2 etching has much lower etch rate and selectivity than F-based etching [46]

Table 6.2 Materials and corresponding RIE gases

Materials to be etched	Chemical gases (multi choices)
Single-crystal silicon	CF_3Br, HBr/NF_3, SF_6/O_2
Polysilicon	$SiCl_4/Cl_2$, BCl_3/Cl_2, $HBr/Cl_2/O_2$, HBr/O_2, Br_2/SF_6
Al	$SiCl_4/Cl_2$, BCl_3/Cl_2, HBr/Cl_2
Al-Si-Cu, Al-Cu	BCl_3/Cl_2+N_2
W	SF_6, NF_3/Cl_2
TiW	SF_6
WSi_2, $TiSi_2$, $CoSi_2$	CCl_2F_2/NF_3, CF_4/Cl_2
SiO_2	CCl_2F_2, CHF_3/CF_4, CHF_3/O_2, CH_3CHF_2
Si_3N_4	CF_4/O_2, CF_4/H_2, CHF_3, CH_3CHF_2
GaAs	$SiCl_4/SF_6$, $SiCl_4/NF_3$, $SiCl_4/CF_4$
InP	CH_4/H_2
Photoresists	O_2

The compound forms of halogen gas, though chemically erosive, are rarely able to etch a material or etch at very low rate. They only become chemically reactive and erosive when these halogen atoms are dissociated from the compounds or lose their electrons to become free radicals. Here the energetic ions play an important role to produce the radicals. An interesting experiment demonstrated that when silicon surface was exposed to XeF_2 gas or Ar^+ beam alone, the silicon was removed at less than 1 nm min^{-1}. However, if the XeF_2 gas and Ar^+ beam were both present, the etch rate jumped to 6 nm min^{-1} [47]; therefore, the RIE process, as schematically shown in Fig. 6.22, is the action of four simultaneous processes [43]:

(a) The ions sputter the substrate surface, which helps initially to remove any hydrocarbon contaminants and native oxide layer to create a fresh material surface, promoting absorption of gas molecules onto the material surface. Ion sputtering can also remove material atoms as purely physical sputtering etching.
(b) The ions, if they are reactive species, can directly react with material atoms to form volatile compounds which escape from the substrate surface.
(c) The ions can dissociate the absorbed gas molecules to produce radicals on the substrate surface.
(d) The radicals can move around at the substrate surface and react with the material surface atoms to form volatile compounds leaving the surface. All the volatile compounds are pumped away, so that no redeposition occurs during the etching process.

It is clear that there are three factors that lead to the RIE: the energized ions, the reactive gas, and the formation of volatile compounds. First, if there are only energized ions without reactive gas, for example etching in pure Ar plasma, then the etching is pure physical sputtering. Ion milling is a typical sputter etching technique conducted in Ar plasma. Second, if there are only reactive gas molecules without energized ions, the radicals are only generated in plasma not on the

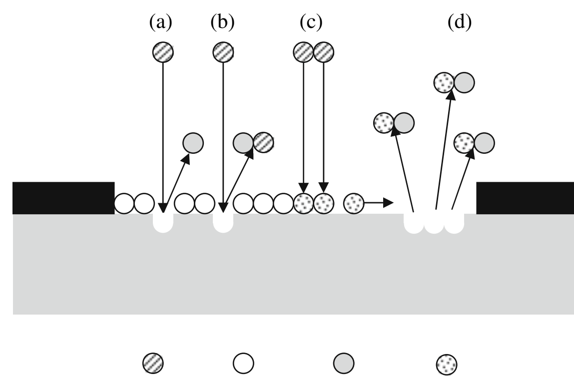


Fig. 6.22 Schematic of RIE process: (**a**) ion sputtering, (**b**) reactive etching, (**c**) radical formation, (**d**) radical etching

surface. The etching is a pure chemical reaction. For example, if the sample is placed on the anode shown in Fig. 6.21, there is little electrical field to accelerate the ions, or there are no electrodes at all, such as in a barrel etcher, the ion bombardment of surface is negligible. The etching is isotropic and similar to aforementioned wet chemical etching. The third factor is the volatility of etched products. If there are no volatile compounds formed during etching, the removed material by etching will redeposit somewhere else in the vicinity of etched pattern.

The volatility of halogen compounds can be characterized by their boiling temperature. The boiling temperatures of halogen compounds for some most commonly encountered elements in micro- and nanostructure fabrication are listed in Table 6.3 [48]. The data in the Table 6.3 can serve as a guideline for selecting appropriate halogen gas when etching a particular material. The general trend is that when going from F to Cl and Br chemistry, etch rates decrease and etching needs more ion assistance. An example has already been described previously about etching silicon with F and Cl gases. Fluorine plasma etches silicon faster but is less anisotropic. Chlorine plasma etches silicon slower but is more anisotropic. This behavior, which is in favor of an easier profile control, is in part due to the reduction of the etch product volatility. For compound materials such as GaAs, there is generally not one etch product for the compound material, but one etch product per constituent, each having its own volatility. This complicates the choice of the chemistry and has strong effects on the etching mechanisms and on the chemical composition of the etched surface. To increase the volatility of those etched products with high boiling temperatures, the sample has to be heated to high temperature. In some cases, adding hydrogen gas into plasma works better because some hydrides have much higher volatility than the halogen compounds, as shown in the Table 6.3.

6.3.4 Process Control in Nanoscale RIE

Reactive-ion etching is a complex process. The above description is only a qualitative and simplistic description of what is happening inside the plasma etching chamber. So far there are no theories that can satisfactorily describe the complete RIE process, though there have been various attempts to establish theoretical models of plasma etching processes [42]. Carrying out pattern transfer by RIE is highly experimental and empirical. There are RIE parameters which can be externally controlled, such as gas flow rate, RF power, chamber pressure, substrate temperature, substrate bias voltage, etc. There are also measurable indicators that indicate the quality of etching, such as the aforementioned etching selectivity and anisotropy, as well as etch rate and etch uniformity, etc. The control of an RIE process is to adjust these parameters experimentally to match them with a desirable outcome from the above indicators. Though theoretically predicting a plasma etching process is difficult, there are some broad guidelines with which a user should be able to anticipate what may happen when certain process parameters are changed. These guidelines are briefly described in the following:

Table 6.3 Volatility of etched compounds with different gas chemistries (Reproduced from [48] with permission)

Elements	Fluorides	Boiling temperature (°C)	Chlorides	Boiling temperature (°C)	Bromides	Boiling temperature (°C)	Hydrides, trimethyls	Boiling temperature (°C)
Al	AlF_3	1297 (subl.)	$AlCl_3$	178 (subl.)	$AlBr_3$	263		
As	AsF_3	−63	$AsCl_3$	130.2	$AsBr_3$	221	AsH_3	−55
	AsF_5	−53			$AsBr_5$			
C	CF_4	−128	CCl_4	77	CBr_4	189	CH_4	−164
Cr	CrF_2	>1300	CrO_2Cl_2	117	$CrBr_2$	842		
Cu	CuF	1100 (subl.)	CuCl	1490	CuBr	1345		
	CuF_2	950	$CuCl_2$	993			CuH	55−60
Ga	GaF_3	1000	$GaCl_3$	201.3	$GaBr_3$	278.8	$Ga(CH_3)_3$	134
Ge	GeF_4	−37 (subl.)	$GeCl_4$	84	$GeBr_4$	186.5	GeH_4	−88.5
In	InF_3	>1200	$InCl_3$	300 (subl.)			$In(CH_3)_3$	55.7
Mo	MoF_5	213.6	$MoCl_5$	268				
	MoF_6	35	$MoOCl_3$	100 (subl.)				
	MoO_2F_2	270 (subl.)						
	$MoOF_4$	180						
P	PF_3	−101.5	PCl_3	75	PBr_3	172.9	PH_3	−87.7
	PF_5	−75	PCl_5	162 (subl.)	PBr_5	106		
Si	SiF_4	−86	$SiCl_4$	57.6	$SiBr_4$	154	SiH_4	−111.8
Ta	TaF_5	229.5	$TaCl_5$	242	$TaBr_5$	348.8		
Ti	TiF_4	284 (subl.)	$TiCl_4$	136.4	$TiBr_4$	230		
W	WF_6	17.5	WCl_6	346.7				
	WOF_4	187.5	WCl_5	275.6	WBr_5	333		
			$WOCl_4$	227.5	$WOBr_4$	327		

1. *Flow rate of reactive gas*. The basic process in RIE is gaseous ions reacting with substrate surface and converting the surface material into volatile products. The reaction rate is directly linked to the supply rate of reactive gases. However, fast flow rate of reactive gases at a constant pressure can lead to short dwelling time of gas molecules in reaction chamber, because the constant pressure is maintained by high pumping speed. If plasma discharge power is not increased accordingly, the number of gaseous ions generated in reaction chamber will decrease, resulting in reduction of etching rate. On the other hand, if the gas flow rate is too low, the consumed reactive gases cannot be replenished sufficiently, which again can lead to drop of etching rate.

2. *RF power* Increasing RF power will increase electron energy, hence increase ionization probability; therefore, etching rate generally increases with the increase of RF power, particularly when gas flow rate has increased or chamber pressure has increased. As explained earlier, high-energy electrons not only generate more ions, but also build up a high negative field near the cathode surface, called self-bias field, which accelerates ions toward cathode. The potential difference at cathode region is 200–500 V. Accelerated ions bombarding cathode (platen) where the sample is situated can also increase the etching rate. In addition, the anisotropy of etching can be improved because ion sputtering is highly directional and almost has no lateral etching. However, ion sputtering is a physical process, irrespective of mask or substrate. Too much of ion sputtering will speed up the consumption of masking layer and reduce etching selectivity. Figure 6.23 shows the Si etch rate and selectivity as functions of platen power which represents the energy of bombarding ions in an induction coupling plasma etching system [49]. Although high platen power is favored for high etch rate, the etch selectivity suffers at the high platen power.

3. *Chamber pressure* Reactive-ion etching is normally taking place at low-pressure condition. The chamber vacuum is usually at $10^{-3} \sim 10^{-1}$ Torr. At low pressure, the density of gas molecule is low in the chamber. Electrons have longer mean free path and can acquire more acceleration energy between two collisions with gas molecules, which can increase ionization probability. Another benefit

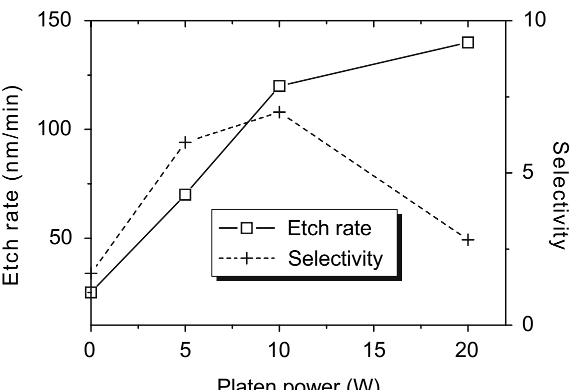


Fig. 6.23 Influence of platen power on Si etch rate and etch selectivity

of low chamber pressure is less lateral etching and improvement of anisotropy. This is due to fewer collisions between ions and gas molecules, resulting in more directional movement of ions toward substrate. Low chamber pressure also means high pumping speed, which helps to quickly move away the volatile products from etched surface. The reaction rate is improved, hence the etching rate. In general, at low chamber pressure RIE is dominated by ion sputtering etching, because chemically active gaseous molecules are less and ion energy is higher. At high chamber pressure, the etching is dominated by chemical reaction, because the mean free path of ions is reduced and the number of chemically active ions is large. At higher pressure, the etching is more isotropic. However, if the pressure is too low, the plasma cannot start or sustain. There exists an optimum pressure at which both etch rate and selectivity are maximum, as an example shown in Fig. 6.24, where both silicon etch rate and selectivity are decreased when chamber pressure is too low or too high [49].

4. *Substrate temperature* In general, high temperature can promote chemical reaction and also help to remove reacted products from substrate surface. As described earlier, etching of In-containing III–V semiconductor compounds with Cl plasma requires high substrate temperature, because of the low volatility of In–Cl compounds. In RIE, ion bombardment of substrate will naturally cause the rise of substrate temperature. The substrate surface temperature can reach 100°C–200°C, depending on the duration of etching. However, high substrate temperature can soften masking material if it is photoresist, leading to pattern distortion and higher photoresist etch rate. Using other inorganic materials instead of photoresist as mask will avoid the problem. Another effect of high substrate temperature is the increase of lateral etching, because the mobility of radicals is enhanced at high surface temperature and hence the chemical etching which is by nature an isotropic etching. For long duration of etching, the sample holder has to be cooled to reduce the substrate temperature.

5. *Gas additives* Although halogen compounds are the main gases for etching inorganic and metallic materials, additives are often used to enhance some

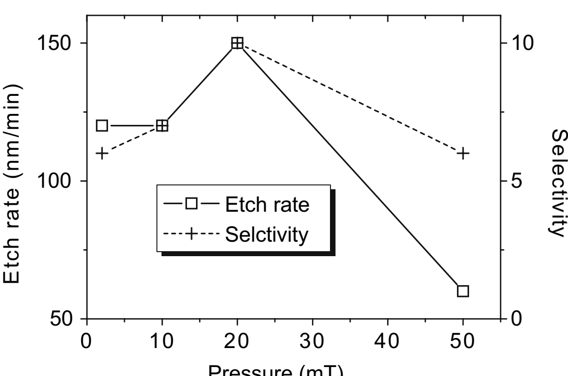


Fig. 6.24 Influence of chamber pressure on silicon etch rate and selctivity

aspect of the chemistry. For example, the addition of 10% O_2 to CF_4 increases the etch rate by 10 times. In the plasma, CF_4 is broken down into radicals CF_3 and CF_2. In the absence of oxygen these radicals can recombine (both in the gas phase and on surfaces) into stable molecules; for example,

$$CF_3 + CF_3 \rightarrow C_2F_6 \text{ and } CF_3 + F \rightarrow CF_4.$$

With oxygen present, it can react with the radicals to form CO, CO_2, and COF, thus releasing fluorine radicals

$$CF_2 + O_2 \rightarrow COF + F, \quad COF + O \rightarrow CO_2 + F.$$

Figure 6.25 shows Si etch rate and selectivity as functions of flow rate of O_2 additive [49]. Addition of oxygen has clearly enhanced the silicon etch rate. However, the selectivity is substantially reduced because oxygen is an effective etching gas to etch polymer resists which is the masking material in this case. Nevertheless, addition of oxygen also brings the benefit of eliminating micromasking effect and helps to produce smooth etched surface.

Apart from O_2 gas, Ar gas is often added into plasma. While Ar^+ can participate in physical sputter etching, they can also help to stabilize the plasma. For example, SF_6 is an electronegative gas. It absorbs electrons upon ionization, which may reduce the total number of electrons in the plasma. The addition of Ar can add electrons to the plasma when they are ionized, and so helps to produce a stable plasma [43]. However, adding too much Ar can dilute the concentration of reactive gas and also cause excessive physical sputtering, which leads to reduction of both etch rate and selectivity.

On the one hand, RIE is a complex process and is influenced by many parameters. On the other hand, there are more degrees of freedom to control the process conditions and to optimize. It is possible to fine-tune process parameters to reach a compromise between etch rate and selectivity and to achieve a particular etch profile. Figure 6.26a–c are three different profiles of

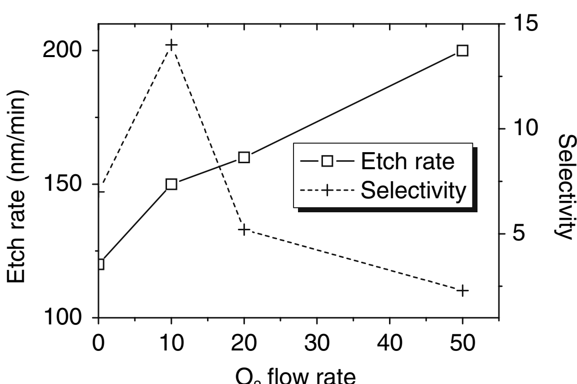
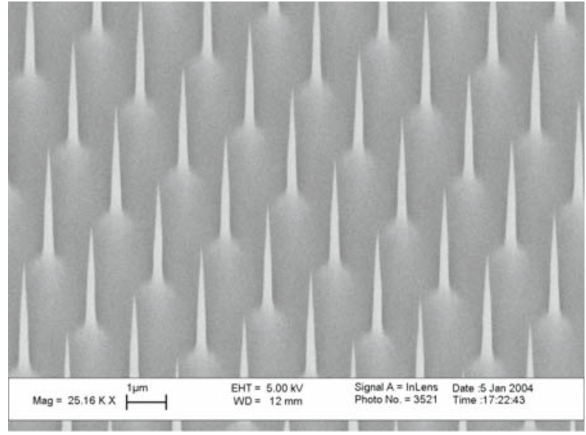
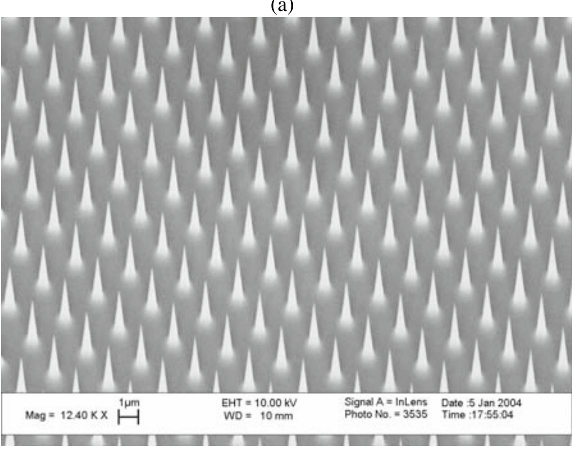


Fig. 6.25 Influence of oxygen flow rate on Si etch rate and selectivity

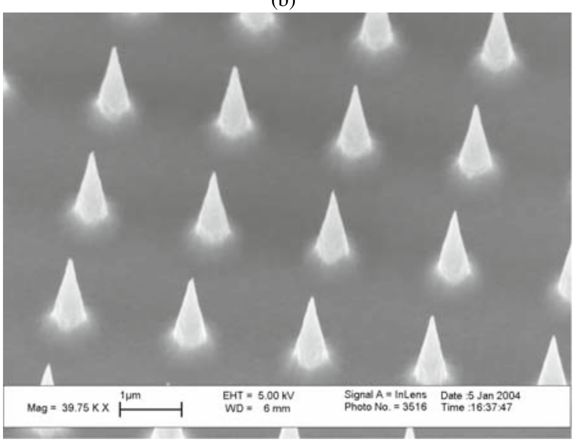
Fig. 6.26 Controlled etching to achieve different shapes of silicon tips: from very sharp tips (**a**), to medium sharp tips (**b**), and to blunt tips (**c**)



(a)



(b)



(c)

etched silicon tips. Although the basic etching chemistry is based on SF_6 plasma, different shapes of etch profile can be fine-controlled by adding CHF_3, O_2, and Ar gases [50]. The sharp end of these silicon tips can be controlled to less than 50 nm in diameter.

The control of anisotropy in RIE is tricky because etching by halogen radicals in plasma is nondirectional. It causes the etching to advance both in depth and laterally, resulting in the pyramid shape of etched profile as shown in Fig. 6.26. Only the ion bombardment of substrate is directional. However, halogen radical etching is very sensitive to polymers. A thin layer of polymer or carbon compounds can stop the etching. A rough guide to control the anisotropy of RIE is to find a way to create a passivation film on the etched sidewall to stop halogen radicals etching laterally, while any passivation on the bottom of etched structure can be removed by directional ion bombardment with the progress of etching process. It is not difficult to find carbon compounds in RIE because most of the etching processes employ polymer resists as the etching mask layer. Ion bombardment removes the polymer masking material, though at very low rate. The carbon compounds in the plasma can deposit onto the sidewall of etched structures, forming the passivation layer. It is also possible to form the passivation layer by introducing carbon compound gas, such as CHF_3, into the plasma. Sub-50 nm silicon nanowires have been produced with aspect ratio up to 50:1 using mixture of $SF_6 + CHF_3$ etching gases [51]. With the assistance of CHF_3, sub-100 nm silicon structures can be etched in RIE even without the use of a polymer masking layer. Figure 6.27 shows etched silicon nanostructure of 50 nm in width and 350–400 nm in height with near-vertical sidewall profile in which 10-nm chrome film was used as masking layer. The etching was done with CHF_3 and SF_6 gas mixture at RF power of 250 W and chamber pressure of 30 mTorr [52].

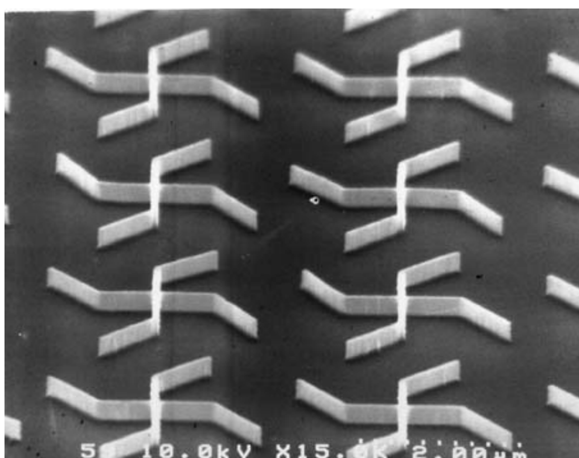


Fig. 6.27 Sub-100 nm etched silicon structure using RIE with CHF_3 and SF_6 gas mixture and chrome film as etching mask

Another sidewall passivation technique is low-temperature etching of silicon, also called cryogenic etching. During etching of Si, SF_6 gas reacting with Si forms volatile compound SiF_4. At room temperature, the compound is quickly removed from reaction chamber by vacuum pumping system. However, at low temperature (<–100°C) the compound can form a 10–20-nm protective film over Si surface. At low temperature, the reaction of F free radicals with Si is greatly weakened because of this protective film; also weakened is the reaction with photoresist or oxide. Both the sidewall protection and high selectivity of mask to Si are realized by low-temperature RIE, resulting in excellent anisotropic etching of silicon [53]. The optimum low temperature is –110°C at which no lateral etching of silicon was found. Below –140°C silicon etching also stops because the etching gas SF_6 starts to freeze on the cold substrate surface [54].

6.3.5 RIE by Inductively Coupled Plasma

It has been demonstrated that conventional RIE is perfectly capable of etching sub-100 nm structures. The only limitation is its relatively low etch rate, which is normally less than 200 nm min^{-1} for Si [55]. The etching rate in RIE directly depends on plasma density. For the conventional RIE system as shown in Fig. 6.21, plasma density increases with increase of RF power. However, increasing RF power will also increase the self-biasing voltage on the cathode where the etching sample is situated. The consequence is the increase of ion bombardment energy, hence deterioration of etching selectivity. This becomes a particularly serious problem in sub-100 nm RIE, because the masking layer is always thin to enable photon or e-beam patterning of polymer resist at sub-100 nm feature dimension.

Inductively coupled plasma (ICP) system has cleverly solved the problem. In an ICP source, the plasma generation is separated from etching chamber, as shown in Fig. 6.28. Radio frequency power is coupled into plasma chamber by an induction coil from outside. The sample stage is connected to a second RF power source as an auxiliary RF source to enhance the production of plasma. Electromagnetic field generated by inductive coupling coil can sustain electron cycling movement in plasma for a long period, which has greatly increased ionization probability while keeping the pressure low in the etching chamber. As the sample stage has an independent input of RF power, the self-biasing voltage can be independently controlled. Therefore, an ICP system can produce very high plasma density ($>5\times10^{11}$ cm^{-3}) compared to conventional RIE (10^8–10^{10} cm^{-3}), as well as maintain low ion bombardment energy. The conflict between high plasma density and high etch selectivity encountered in conventional RIE systems has been resolved. The current best ICP systems can produce over 20 μm min^{-1} etching rate and are poised to achieve 50 μm min^{-1} or even 100 μm min^{-1} in a few years time [56]. The selectivity between Si and

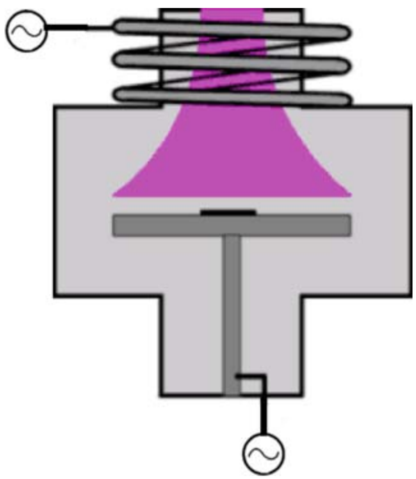


Fig. 6.28 Schematic of inductive coupling plasma source

photoresists can be as high as 100:1. Selectivity between silicon and oxide is over 200:1. With such high plasma density and high etching selectivity, high etching rate and deep RIE (DRIE) become possible.

To ensure high anisotropy of DRIE, a gas switching process was introduced to deliberately passivate the sidewall to stop lateral etching. The gas switching process, called “Bosch” process [57], is illustrated in Fig. 6.29. It is a repeated etching–passivation process by introducing alternating gases. The etching gas is SF_6 and passivation gas is C_4F_8. C_4F_8 in plasma can form a fluorine carbon compound which once deposited on silicon surface can prevent F^+ from reacting further with Si. Etching and passivation are switched over every 5–15 s a cycle. Although pure SF_6 gas reacting with Si results in isotropic etching, the etched surface is soon passivated by C_4F_8 after a few seconds of C_4F_8 reaction. In next round of SF_6 gas being introduced, ion bombardment in depth direction removes the passivation film to allow SF_6 reacting further with the fresh Si surface, whereas ions cannot sputter the sidewalls. The passivation film remains intact to protect sidewalls from being etched. Through the periodical etching–passivation–etching cycle, Si etching only proceeds in depth direction and eventually forms vertical sidewalls. The typical “Bosch” process conditions are listed in Table 6.4

Though it is capable of etching very deep structures with near vertical sidewall profiles, the “Bosch” process has a fundamental problem: the scalloping effect; for example, ripples are formed over the sidewall surface due to the switching between etching and passivation, as the example shown in Fig. 6.30. The surface roughness due to the scalloping is in the order of more than 100 nm [58]. The surface roughness caused by the scalloping effect may be tolerable if the feature dimension is tens or hundreds of micrometers, which is the case in MEMS devices. However, it cannot be intolerable if the feature dimension is in sub-100 nm range. An example is shown in Fig. 6.31 where the passivation time

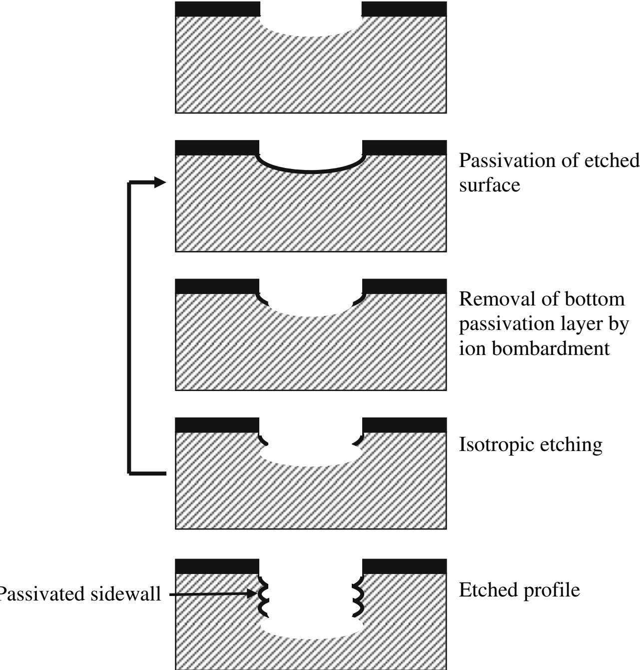


Fig. 6.29 Schematic of "Bosch" process

Table 6.4 Typical "Bosch" process conditions

Process parameters	Passivation	Etching
C_4F_8	85 sccm	0 sccm
SF_6	0 sccm	130 sccm
RF power at stage	0 W	12 W
RF power from coil	600 W	600 W
Cycle time	7.0 s	9.0 s
Delay time	0.5 s	0.5 s
Etch rate		1.5–3 μm min^{-1}

sccm: standard cubic centimeter per minute

is 5 s and etching time is 6 s. There are eight cycles of gas switching, which can almost be counted from the number of ripples on the sidewall [59].

The high roughness of sidewall is undesirable in many applications. For example, a nanoimprint mold made in Si by the "Bosch" process is almost impossible to use because the surface roughness can severely damage imprinted polymer nanostructure during demolding process. Many investigations have been carried out to reduce the scalloping effect. For large structure dimension, the scalloping may be smoothed by post-etch polishing in wet chemicals such as KOH + IPA [60, 61]. For sub-100 nm feature dimension, the wet chemical

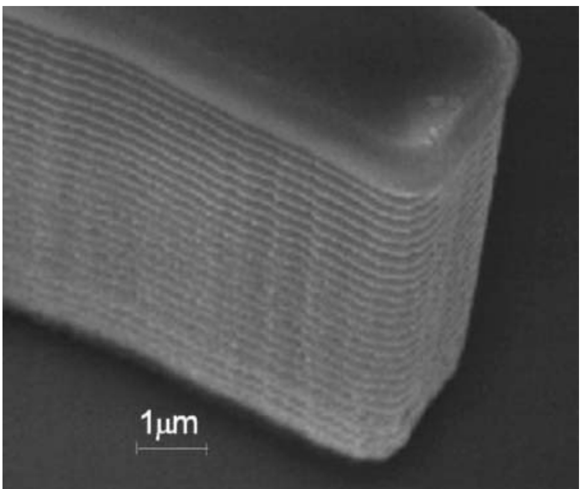


Fig. 6.30 Scalloping effect created in "Bosch" process

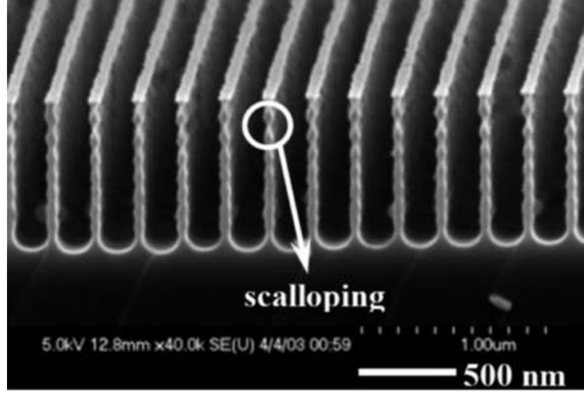


Fig. 6.31 Scalloping effect in etched sub-100 nm silicon structures (Reprint from [59] with permission)

polishing may distort the structure. Other techniques have been developed, including prolonging the passivation time [62], using a Teflon-coated sample bed [63], or simply mixing the SF_6 etch gas with the C_4F_8 passivation gas without switching [64]. Cryogenic etching is also an option, in which no passivation gas is needed and the etching chemistry is the same as that of conventional RIE, for example, $SF_6 + O_2$. Normally the roughness caused by scalloping is in the range of a few hundreds of nanometers. With the aforementioned techniques the scalloping-induced roughness can be reduced to less than 20 nm. Table 6.5 compares the performances of three nonswitching DRIE processes for etching Si nanostructures: HBr process, room temperature $SF_6 + C_4F_8$ process (RT F-base), and cryogenic process [65]. Each process has different etch rate and selectivity, but all are capable of producing very smooth sidewalls (<10-nm roughness). Figure 6.32 shows 16-nm etched Si line structures using the non-switching $SF_6 + C_4F_8$ DRIE process.

Table 6.5 Comparison of three nonswitching DRIE processes

Process	HBr	RT F-base	Cryogenic
Depth	0.05–1 μm	0.05–10 μm	0.2– > 100 μm
Feature size	≥5 nm	5 nm	≥5 nm
Aspect ratio	≥5:1	≥5:1	≥10:1
Etch rate (nm/min)	≥100	≥200	≥300
Uniformity	< ± 5%	< ± 5%	< ± 5%
Selectivity over oxide	>100:1	> 10:1	>30:1
Sclectivity over resist	>3:1	>5:1	>15:1
Profile	80–92°	80–92°	80–92°
Sidewall roughness	<10 nm	<10 nm	<10 nm

Reproduced from [65] with permission.

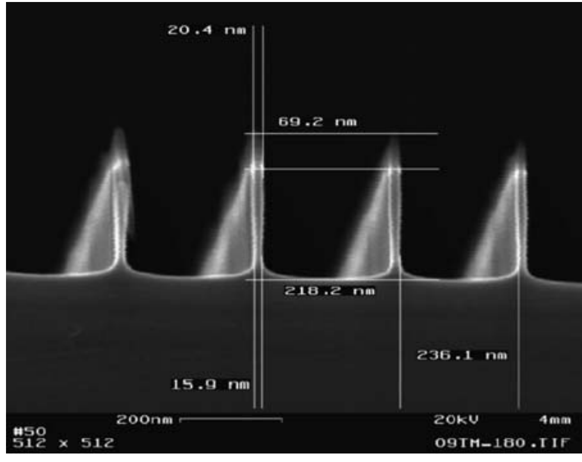


Fig. 6.32 High-aspect-ratio Si nanostructures etched by nonswitching DRIE process (16-nm linewidth) (Reprint from [65] with permission)

Although the above results are for Si etching, ICP also have been widely used for etching other materials. The ICP with high plasma density offers much higher etch rate and selectivity than it is possible by conventional RIE. For example, SiO_2 is known to be difficult to etch. With a conventional RIE, the etch rate is less than 50 nm min^{-1}. Using an ICP the etch rate is nearly 10 times enhanced. For example, 500 nm min^{-1} etch rate for SiO_2 has been demonstrated, with 21:1 selectivity using polysilicon mask or 10:1 selectivity using photoresist mask [66]. Inductively coupled plasma systems are also widely used to etch III–V semiconductor materials which are of great importance in high-frequency and optoelectronic applications [67].

6.3.6 Critical Issues in RIE

Whether using conventional RIE or ICP-based DRIE for pattern transfer, there are a number of critical issues one should be aware of. These issues are

related to the etch uniformity, control of etch profile, or damage of substrate materials.

1. *Loading effect* Reactive-ion etching rate depends on the supply rate of chemically active species. When the consumption of radicals surpasses their supply while the gas flow rate and other conditions remain the same, the etching rate will decrease, causing the so-called loading effect. There are three types of loading: macroloading, microloading, and aspect-ratio-dependent etching (ARDE) [68, 69]. They all mean the same thing, that is, nonuniform etching due to the difference in supply and consumption of reactive radicals at the etching surface. Macroloading effect is the decrease of overall etch rate due to the increase of total area of materials to be etched away, as shown in Fig. 6.33a, which poses the problem that a well-characterized etching process for one set of the pattern design may not be applicable to another set of design if the overall etching areas are different. Microloading effect is related to localized depletion of reactive species or accumulation of etching production as a result of local variation of pattern density, as schematically illustrated in Fig. 6.33b. High-density pattern areas demand high rate of supply of reactive species as well as high rate of removal of etched volatile compounds than low pattern density areas. Such differences in

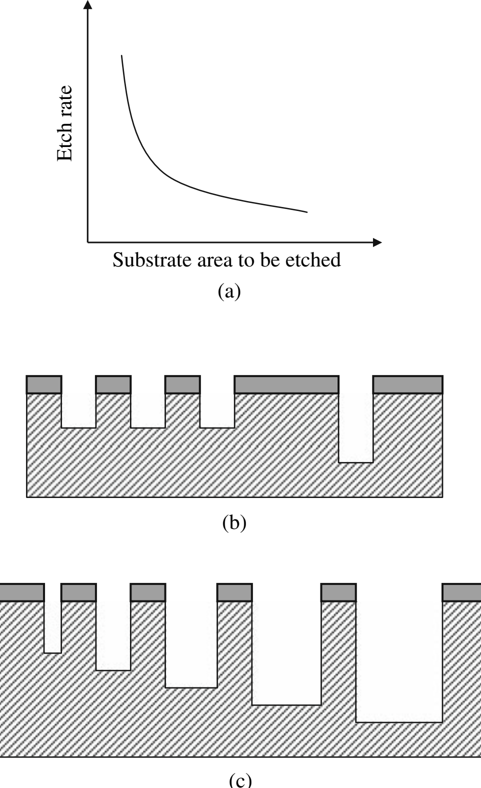


Fig. 6.33 Loading effect in RIE. (**a**) Macroloading, (**b**) microloading, (**c**) aspect-ratio-dependent etching

the rates of radical supply and etched product removal can cause difference in etch depth between high-density pattern area and low pattern density area. Generally, isolated patterns are etched deeper than dense patterns. Aspect-ratio-dependent etching is strongly dependent on feature dimensions. Other names for the ARDE are RIE lag or aperture effect. As illustrated in Fig. 6.33c, the smaller the feature dimension (or pattern opening to etching) the less is the etch depth under the same etching condition. The cause of ADRE is very complicated. There are as many as eight different mechanisms that may explain the ADRE phenomenon [68].

Loading effect is most significant in DRIE because of prolonged etching and large etching depth. The longer the etching time the more significant is the loading effect. There can be 100% variation in etching depth [69]. Conventional RIE also has the problem of aperture effect, when both large and small features have to be etched on the same substrate. For example, the etch rate of SiO_2 has shown to drop from 32 nm min^{-1} to 12 nm min^{-1} when the etched feature size changes from 20 μm to 52 nm [70]. There are no very effective measures that can eliminate the loading effect. Generally, high pumping speed (low chamber pressure), diluting gases, or using pulsed plasma can reduce to some extent the loading effect [71]. These measures help to remove the etched products fast or to reduce local nonuniformity of reactive species. Alternatively, the pattern layout can be modified to reduce or eliminate the nonuniformity in feature density or feature dimension, for example, introducing dummy features purely to balance the feature density or subdividing large features in multistep etching [61]. If the pattern design cannot be easily altered, one has to decide what the critical feature dimension in a design layout is if different feature sizes coexist in the design. Then the RIE condition should be optimized for the critical feature dimension, allowing variation in etch depth for other noncritical features.

2. *Microtrenching effect* Microtrenching is a phenomenon that the etch rate is larger near the corners of a trench compared to the center of the trench. Typical microtrenching in etched features is shown in Fig. 6.34 where deep grooves appear adjacent to the sidewalls of etched structure. Microtrenching is believed to be produced by the impact of high-energy particles (mostly ions) at grazing angles ($>80^\circ$) on the sidewalls then reflected toward the bottom of the etched surface, as shown in Fig. 6.35. These ions retain a large fraction of their energy and directionality. It is like a ‘‘focusing’’ of the high-energy particles at the base of the sidewalls of the feature, resulting in higher etch rates at those locations [72]. Such a mechanism has been proved by Monte Carlo simulation of ions sliding at sloped sidewalls [73]. Both sidewall slope angle and incident angle of ions can significantly influence the microtrenching formation. Generally, the deeper the etch depth, the more significant is the microtrenching because a larger sidewall surface becomes available for reflecting ions. Negative charging of sidewalls can also contribute to microtrenching, because the incident ions can be deflected by the negative field toward the sidewall [74]. By deliberately introducing flux of electrons into the etched trenches, the local electrical field can be neutralized and the microtrenching effect can be reduced [75].

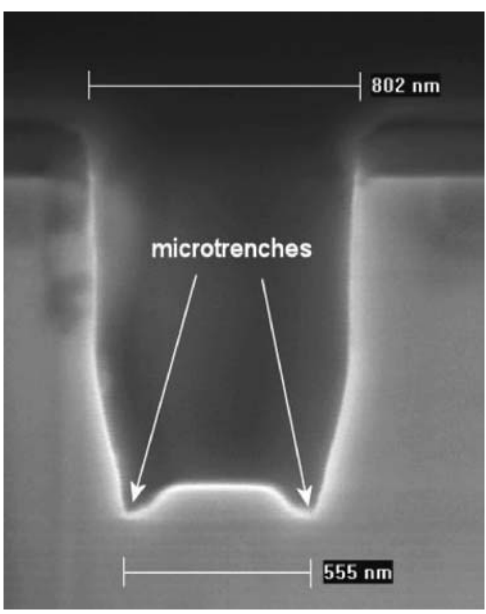


Fig. 6.34 Microtrenching effect in RIE

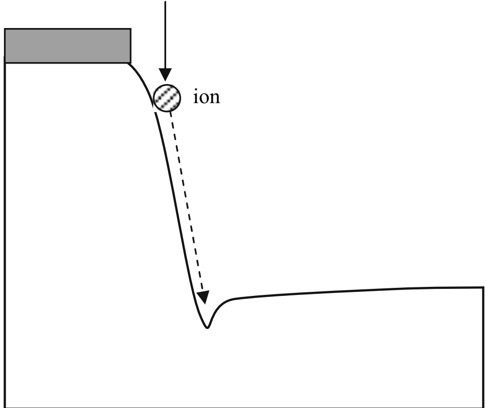


Fig. 6.35 Mechanism of microtrenching

3. *Notching effect* Notching is the phenomenon of lateral etching at the foot of etched structure, which occurs when etching nonuniform size of features in silicon-on-insulator (SOI) substrates. Typical notching effect is shown in Fig. 6.36 [76]. Silicon-on-insulator wafers are special Si substrates in which a thin SiO_2 layer is sandwiched between a thick silicon wafer (handle layer) and a thin silicon wafer (device layer) [77]. Theoretically, the SiO_2 insulator underneath the device layer provides an excellent etch stop layer which ensures all features, no matter big or small, are all etched to the same depth (the thickness of device layer). There should be no microloading effect or ARDE. However,

it is found that big features are continuously etched laterally at the foot of trenches even though the etching has stopped in depth direction. Small features can also have notches if they are overetched. The notching, also called "footing", is believed to be caused by positive charging at the bottom insulating layer, as illustrated in Fig. 6.37. Positive charges accumulated at the SiO_2 layer creates a local positive field which deflects the incoming ions toward sides of the surface, causing the lateral etching [78]. There is nothing one can do to prevent the notching, unless the positive charge can be conducted away [79], or applying a low-frequency pulsed bias power to the substrate stage to give the insulating layer a chance to dissipate the accumulated charge during the bias-off period [80].

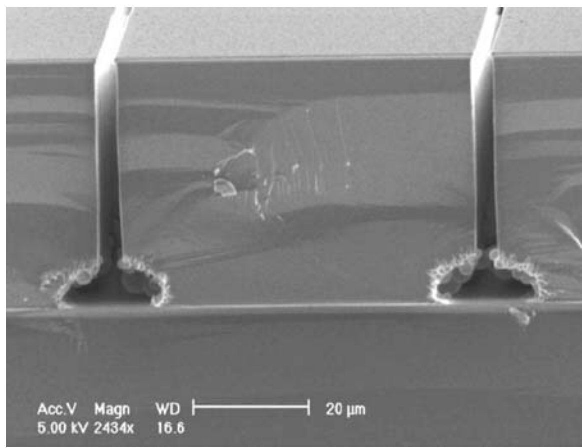


Fig. 6.36 Notching in DRIE of SOI substrate (Reprint from [76] with permission)

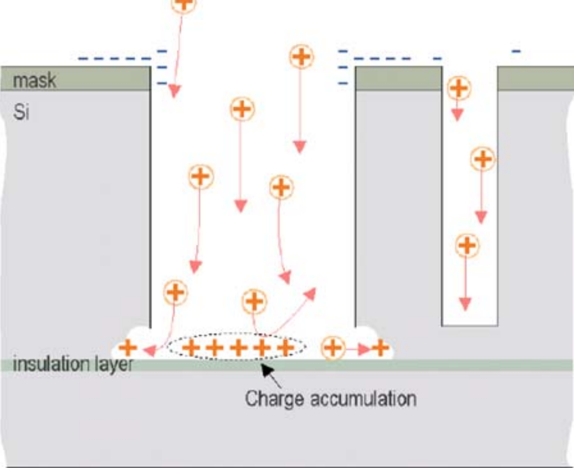


Fig. 6.37 Mechanism of notching during etching of SOI substrate (Reprint from [79] with permission)

3. *Ion-induced damage effect*. Ion sputtering during RIE can introduce damage to the substrate. Though the energy of ions is much lower than FIB sputtering, being in the range of 50–500 eV, and the penetration depth is only a few nanometers, it can change the surface conductivity, the mobility of charge carriers, with depth up to 100 nm [81]. The much deeper damage range is due to the ion channeling effect; for example, ions can travel much larger distance in a crystalline material if the incident ions happened to be in the path of low index lattice. It is generally believed that this damage is due to the introduction of traps, vacancies, and interstitial atoms into the crystalline lattice. For single crystalline materials such as Si, the ion-induced damage can always be reduced or eliminated by post-etching annealing. However, high annealing temperature can cause decomposition of compound semiconductors such as GaAs, etc. Reducing ion energy is the most effective way of reducing the damage. For example, it is found that when RF power increased from 20 W to 40 W the carrier mobility and concentration at the etched surface decreased by 13.8% and 29.2%, respectively [82]. Use of ICP etchers can certainly help to control the ion bombardment energy, because the substrate stage bias voltage can be independently controlled. However, detailed study found that even with ICP one should avoid generating atomic ions in the plasma because etching with atomic ions showed more damage than with molecular ions [81]. On the other hand, reducing RF power may have to involve changing other conditions in order to maintain necessary etch rate, selectivity, and anisotropy. As a general rule, every etching process has to be adjusted and optimized to balance the different requirements.

6.3.7 Ion Milling

Subtractive pattern transfer at sub-100 nm dimension is also possible by ion milling. Ion milling is a pure physical ion sputtering process. Ion milling is needed because not all materials can be etched by RIE. Only those materials which chemically react with reactive ions and form volatile compounds can be etched by RIE. For other materials, ion milling is the alternative for subtractive pattern transfer.

Argon gas is commonly used to generate gaseous ions in ion milling. As Ar is an inert gas, Ar ions do not react chemically with any sample materials. There are different physical reactions between Ar^+ and sample surface, depending on the Ar^+ energy. If the energy of Ar^+ is less than 10 eV, surface absorption or surface damage is the dominant process. Ion damage to sample surface confines to only a few atomic layers. If Ar^+ energy is over 10 keV, ions can penetrate deep inside the sample and become ion implantation. Ion milling only occurs at Ar^+ energy between 10 eV and 5000 eV. It has been shown in previous sections that there is always ion sputtering taking place in RIE. However, ion milling is different from the ion sputtering in plasma or in RIE. In RIE, ions are only accelerated by the field near the sample surface. The sputtering efficiency is low and material removal by pure ion sputtering is not the dominant etching mechanism.

The key difference between ion milling and ion sputtering in RIE is that an ion milling system separates the generation of ions from sample etching. Figure 6.38 is a schematic of typical ion milling system, where ion source is separated from sputtering chamber. In the ion source, electrons emitted from a thermionic cathode are accelerated by the field between cathode and anode. Gas molecules are ionized by energized electrons and plasma is formed in the ion source chamber. The gaseous ions are then extracted from ion source chamber in the form of a broad beam and accelerated toward the sample in the sputtering chamber. Therefore, ion milling is also called ion beam milling. An ion beam milling system is normally fitted with an electron gun, as shown in Fig. 6.38. The electron gun is a hot filament emitting electrons to neutralize positive charge accumulated at the sample surface if it is an insulating material. Typical ion milling rate is 100–3000 A min^{-1} at 1000-eV Ar ion energy and 1 mA cm^{-2} ion current density.

Because Ar^{+} is chemically inert to sample materials, the etch rate by ion milling only depends on the sputtering yield of a particular sample material. In principle ion milling can etch any material. Although ion milling rate is much lower than RIE rate, its nondependence on material types makes the etching technique a good supplement to RIE and other etching techniques. It opens up the possibility to fabricate nanostructures in a much broader range of materials other than those commonly used in microelectronics industry. It is particularly

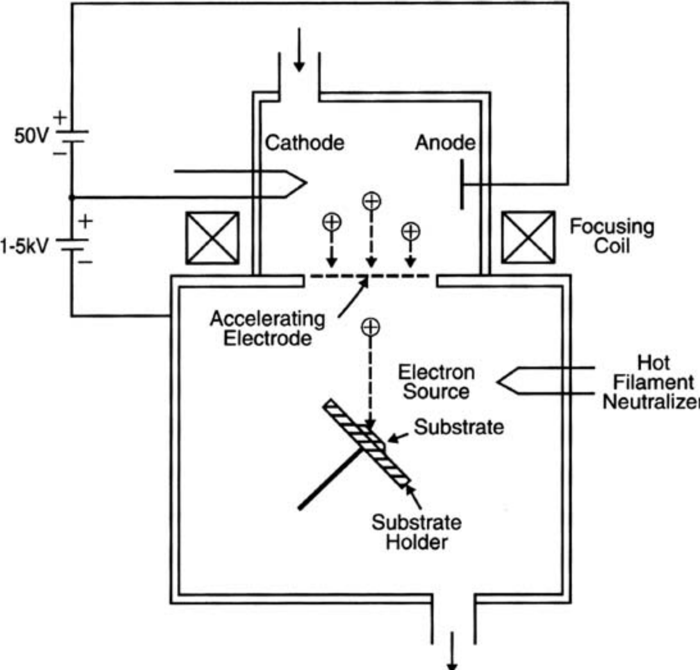


Fig. 6.38 Schematic of ion beam milling system

useful in fabrication of nanomagnetic devices in which many magnetic materials cannot be etched by RIE [83, 84]. However, one must be aware of the following issues when using ion milling for nanoscale pattern transfer:

1. Physical sputtering by Ar ions is basically nonselective to materials. The etch selectivity only depends on the difference of sputtering yield for different materials, which are relatively small. Table 6.6 depicts the ion milling rates for some commonly used materials in electronics industry [85]. It is apparent that photoresist, which is the basic masking material in most RIE processes, is only marginally suitable as mask for ion milling of gold film, but not for nichrome or alumina.

For some materials, the selectivity of photoresist to the etched material can be improved by introducing certain chemical gases, so that the ion milling becomes partially reactive. Figure 6.39 has compared ion beam milling of hafnium oxide (HfO_2) as the function of ion energy with and without adding CHF_3 gas [86]. Without the participation of CHF_3 in the milling process (Ar only), the etching rate of photoresist is higher than that of HfO_2. By mixing two parts of CHF_3 with one part of Ar, the etch rate of HfO_2 is much enhanced whereas the etch rate of photoresist is reduced. In this way, the photoresist serves as an ion milling mask better. It should be noted that the reactive-ion milling is by no means a true sense of RIE because the etched product is still nonvolatile.

Table 6.6 Etching rates of some materials by ion beam sputtering

Materials	Etch rate (A°/min)
Gold	1000
Copper	700
Photoresist (AZ-1350)	200
Nichrome (NiCr)	170
Alumina (Al_2O_3)	90

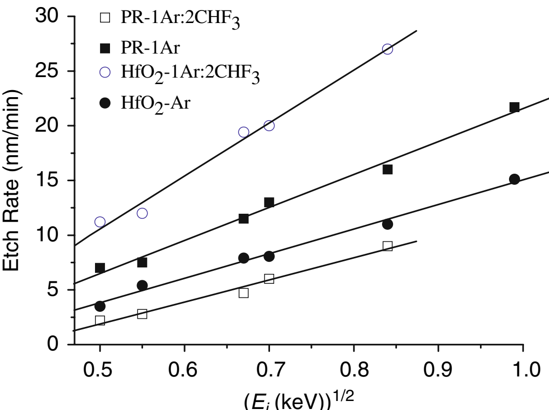
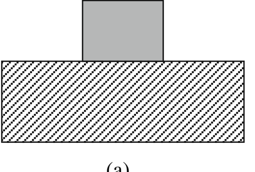


Fig. 6.39 Etch rates of hafnium oxide (HfO_2) and photoresist (PR) as the function of square-root ion energy $(E_i)^{1/2}$

For sub-100 nm pattern transfer by ion milling, e-beam resists have to be used as the masking layer. Though PMMA is the popular high-resolution e-beam resist, it is not a good masking material for ion milling [84]. The negative-tone inorganic resist HSQ works better. Pattern features as small as 30 nm have been made with HSQ by e-beam lithography and serve as ion milling mask to etch gold nanostructures [87]. The gold nanostructures produced by combining negative-tone HSQ mask and ion milling are equivalent to those fabricated by lift-off in positive resist mask but have much smoother edges than those done by lift-off. The thickness of HSQ masking layer has to be similar to the thickness of metal layer to be etched, because the selectivity of HSQ in ion milling of gold is still low [87].

In general, the etch depth by ion milling is limited because of the fast consumption of mask material. As mask thickness is reduced toward the end of etching, the final etched feature may deviate from its original design. The most pronounced profile distortion encountered in ion milling pattern transfer is faceting. Faceting is a phenomenon that top corners of a masking pattern are eroded faster than its planar surface, as illustrated in Fig. 6.40. Figure 6.40a is the initial mask pattern structure before ion milling. Because of the fast sputtering at the pattern corner, the mask pattern becomes sloped in the course of ion milling, so does the underlying etched pattern structure as shown in Fig. 6.40b. The faceting is most severe for polymer resist mask because much thick masking layer has to be used. To reduce the faceting problem, the masking layer should be made very thin. Hard metallic masking materials can be used, such as tungsten, which have been proved to alleviate the faceting problem [88]. However, patterning the hard mask needs an additional fabrication step, which inevitably complicates the pattern transfer process.

(2) Ion milling is a physical etching process and does not form volatile compounds. The sputtered material can redeposit everywhere, particularly onto adjacent patterns. Figure 6.41 shows a group of HfO_2 patterns etched by ion milling [86]. The redeposition of nonvolatile materials is so severe that unwanted overhanging structures were created on top of each etched structure. By rotating while tilting the sample during sputtering, it helps to clean up effectively the redeposition of sputtered debris on the structure sidewalls [84]. Though a gentle wet acidic etch after ion milling can also help to clean up the redeposition, the acid that removes the redeposited material can harm the etched structures as well. To overcome the dilemma, some reactive gases can



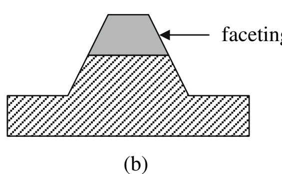


Fig. 6.40 Faceting in ion milling. (**a**) Initial mask pattern profile, (**b**) faceting of etched profile caused by ion sputtering

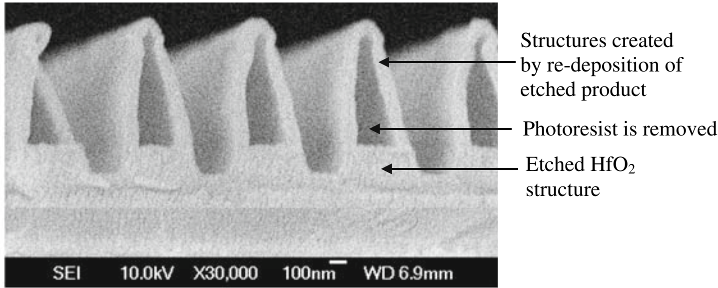


Fig. 6.41 Redeposition in ion milling of hafnium oxide

be introduced into the ion milling process, which may produce some nonvolatile compounds dissolvable in acid but without attacking the pattern structures. For example, by ion milling of HfO_2 with Ar/CHF_3 gas mixture, the reaction product is hafnium fluoride which can be dissolved in HCl acid, but the HCl does not react with HfO_2.

References

1. Elshabini-Riad, A.A.R. and F.D. Barlow III, *Thin Film Technology Handbook*. 1997, McGraw-Hill.
2. Chambers, A., R.K. Fitch, and B.S. Halliday, *Basic Vacuum Technology* 2nd ed. 1998, CRC Press.
3. Chapman, B., *Glow Discharge Processes*. 1980, John Wiley & Sons, Inc.
4. Rossnagel, S.M., *Sputter deposition for semiconductor manufacturing*. IBM J. Res. Develop., 1999. **43**(1/2): p. 163.
5. Ilataakis, M., Canavcuo, and I.M. Shaw, *Single-step optical lift-off process*. IBM J. Res. Develop., 1980. **24**(4): pp. 452–460.
6. Beaumont, S.P., et al., *Sub-20-nm-wide metal lines by electron-beam exposure of thin poly(methyl methacrylate) films and liftoff*. Appl. Phys. Lett., 1981. **38**(6): pp. 436–439.
7. Chou, S.Y., et al., *Sub-10 nm imprint lithography and applications*. J. Vac. Sci. Technol., 1997. **B15**(6): p. 2897.
8. Cui, Z., *Focused ion beam technology*, in *Micro-Nanofabrication Technologies and Applications*. 2006, Springer.
9. LOR product information available from Microchem Inc. (web link: http://www.microchem.com/) .
10. Chen, Y., K. Peng, and Z. Cui, *A lift-off process for high resolution patterns using PMMA/LOR resist stack*. Microelectronic Engineering, 2004. **73/74**: p. 278.
11. Xia, X., et al., *Fabrication of near-infrared and optical meta-materials on insulating substrates by lift-off using PMMA/Al stack*. Microelectronic Engineering, 2007. **84**(5–8): pp. 1144–1147.
12. Radulescu, F., et al., *Introduction of Complete Sputtering Metallization in Conjunction with CO_2 Snow Lift-Off for High Volume GaAs Manufacturing,* in *Proc. GaAs MANTECH Conference*, 2002, San Diego.
13. Voigt, A., et al., *A single layer negative tone lift-off photo resist for patterning a magnetron sputtered Ti/Pt/Au contact system and for solder bumps*. Microelectronic Engineering, 2005. **78–79:** pp. 503–508.

14. Yang, H., et al., *Electron beam lithography of HSQ/PMMA bilayer resists for negative tone lift-off process*. Microelectronic Engineering, 2008. **85**(5–6): p. 814.
15. Romankiw, L.T., L.M. Croll, and M. Hatzakis, *Batch fabricated thin film magnetic recording head*. IEEE Trans. Mag., 1970. **6**(4): p. 729.
16. Romankiw, L.T. and E.J.M. O'Sullivan, *Plating techniques*, in *Handbook of Micolithography, Micromachining and Microfabrication*, P. Rai-choudhury, Editor. 1997, SPIE Optical Engineering Press & IEE.
17. Jiang, M., *Processing considerations for CMP on thin-film head wafers*. Solid State Technology, 2004 (September).
18. Jackson, R.L., et al., *Processing and integration of copper interconnects*. Solid State Technology 1998 (March).
19. Becker, E.W., et al., *Fabrication of microstructures with high aspect ratios and great structural heights by synchrotron radiation lithography, galvanoforming and plastic moulding (LIGA process)*. Microelectronic Engineering, 1986. **4**: p. 35.
20. Cui, Z. and R.A. Lawes, *Low cost fabrication of micromechanical systems*. Microelectronic Engineering, 1997. **35**: p. 389.
21. Cui, Z., et al., *High sensitive magnetically actuated micromirrors for magnetic field measurement*. Sensors and Actuators A: Physical, 2007. **138**(1): pp. 145–150.
22. Attwood, D.T., *Soft X-Rays and Extreme Ultraviolet Radiation*. 2000, Cambridge University Press.
23. Chao, W., et al., *Soft X-ray microscopy at a spatial resolution better than 15 nm*. Nature, 2005. **435**: p. 1210.
24. Simon, G., et al., *Electroplating: an alternative transfer technology in the 20 nm range*. Microelectronic Engineering, 1997. **35**: pp. 51–54.
25. Anderson, E.H., et al., *Nanofabrication and diffractive optics for high-resolution X-ray applications*. J. Vac. Sci. Technol., 2000. **B18**: p. 2970.
26. Haatainen, T., et al., *Nickel stamp fabrication using step & stamp imprint lithography*. Microelectronic Engineering, 2006. **83**: pp. 948–954.
27. Wolf, S., *Introduction to dual-damascene interconnect processes*. Silicon Processing for the VLSI Era, 2004. **4**: pp. 674–679.
28. Luethi, R., et al., *Parallel nanodevice fabrication using a combination of shadow mask and scanning probe methods*. Appl. Phys. Lett., 1999. **75**(9): p. 1314.
29. Brugger, J., et al., *Resistless patterning of sub-micron structures by evaporation through nanostencils*. Microelectronic Engineering, 2000. **53**: p. 403.
30. Kim, G.M., M.A.F. Van den Boogaart, and J. Brugger, *Fabrication and application of a full wafer size micro/nanostencil for multiple length scale surface patterning*. Microelectronic Engineering, 2003. **67–68**: p. 609.
31. Arcamone, J., et al., *Dry etching for the correction of gap-induced blurring and improved pattern resolution in nanostencil lithography*. J. Micro/Nanolith. MEMS MOEMS, 2007. **6**(1): p. 013005.
32. Lishchynska, M., et al., *Predicting mask distortion, clogging and pattern transfer for stencil lithography*. Microelectronic Engineering, 2007. **84**: pp. 42–53.
33. Cui, Z., *Etching Technology*, in *Micro-Nanofabrication Technologies and Applications*. 2006, Springer.
34. Biebuyck, H.A., et al., *Lithography beyond light: Microcontact printing with monolayer resists*. IBM J. Res. Develop., 1997. **41**(1/2): p. 159.
35. Xia, Y. and G.M. Whitesides, *Soft lithography*. Angew. Chem. Int. Ed., 1998. **37**: pp. 550–575.
36. Kendall, D.L. and R.A. Shoultz, *Wet chemical etching of silicon and SiO_2 and ten challenges for micromachiners*, in *Handbook of Microlithography, Micromachining and Microfabrication*, P. Rai-Coudhury, Editor. 1997, SPIE Optical Engineering Press & IEE.
37. IntelliSuite™ from IntelliSense Software Corp. (web link: http://www.inteliisensesoftware.com/)

38. Li, B., et al., *Fabrication and characterization of patterned single-crystal silicon nanolines*. Nano Lett., 2008. **8**(1): pp. 92–98.
39. Wieser, U., et al., *Nanoscale patterning of Si/SiGe heterostructures by electron-beam lithography and selective wet-chemical etching*. Semicond. Sci. Technol., 2000. **15**: pp. 862–867.
40. Zhang, Y.Y., et al., *Fabrication of silicon-based multilevel nanostructures via scanning probe oxidation and anisotropic wet etch*. Nanotechnology, 2005. **16**: pp. 422–428.
41. Sugimura, H. and N. Nakagiri, *Fabrication of silicon nanostructures through scanning probe anodization followed by chemical etching*. Nanotechnology, 1995. **6**: pp. 29–33.
42. Hitchon, W.N.G., *Plsama Processes for Semiconductor Fabrication*. Cambridge Studies in Semiconductor Physics and Microelectronic Engineering. 1999, Cambridge University Press.
43. Wilkinson, C.D.W. and M. Rahman, *Dry etching and sputtering*. Phil. Trans. R. Soc. Lond., 2003. **A362**: pp. 125–138.
44. Ibbotson, D.E., *Plasma and gaseous etching of compounds of Groups III–V*. Pure & Appl. Chem., 1988. **60**(5): pp. 703–708.
45. Peters, L., *Plasma Etch Chemistry: The Untold Story*. Semiconductor International, 1992. **15**(6): p. 66.
46. Tian, W.-C., J.W. Weigold, and S.W. Pang, *Comparison of Cl_2 and F-based dry etching for high aspect ratio Si microstructures etched with an inductively coupled plasma source*. J. Vac. Sci. Technol., 2000. **B18**(4): p. 1890.
47. Coburn, J.W. and H.F. Winters, *Ion- and electron-assisted gas-surface chemistry – An important effect in plasma etching*. J. Appl. Phys., 1979. **50**(5): p. 3189.
48. Cardinaud, C., M.-C. Peignon, and P.-Y. Tessier, *Plasma etching: principles, mechanisms, application to micro- and nano-technologies*. Appl. Surf. Sci., 2000. **164**: pp. 72–83.
49. Wang, X., et al., *Fabrication of nanoimprint template in Si with high etch rate by non-switch DRIE process*. Microelectronic Engineering, 2008. **85**(5–6): p. 1015.
50. Tao, J., et al., *A systematic study of dry etch process for profile control of silicon tips*. Microelectronic Engineering, 2005. **78–79**: p. 147–151.
51. Grigoropoulos, S., et al., *Highly anisotropic silicon reactive ion etching for nanofabrication using mixtures of SF6/CHF3 gases*. J. Vac. Sci. Technol., 1997. **B15**(3): p. 640.
52. Chen, Y., et al., *Nanoimprint lithography for planar chiral photonic meta-materials*. Microelectronic Engineering, 2005. **78–79**: pp. 612–617.
53. Walker, M., *Comparison of Bosch and cryogenic processes for patterning high aspect ratio features in silicon*. SPIE, 2001. **4407**: p. 89.
54. Tachi, S., K. Tsujimoto, and S. Okudaira, *Low-temperature reactive ion etching and microwave plasma etching of silicon*. Appl. Phys. Lett., 1988. **52**(8): pp. 616–618.
55. Schutz, R.J., *Reactive plasma etching*, in *VLSI Technology*, S.M. Sze, Editor. 1988, McGraw-Hill Book Co.
56. Doe, P., *Bosch: Deep etch tools on target for 100 μm/min throughput in 2–3 years*. Solid State Technology, 2007 (September).
57. Laermer, E. and A. Schilp, *Method of anisotropically etching silicon (US Patent 5,501,893)*, 1996.
58. Ayon, A.A., R.L. Bayt, and K.S. Breuer, *Deep reactive ion etching: a promising technology for micro- and nanosatellites*. Smart Mater. Struct., 2001. **10**: pp. 1135–1144.
59. Choi, C.-H. and C.-J. Kim. *Fabrication of silicon nanostructures with various sidewall profiles and sharp tips,* in *Proceedings of Transducers'05: The 13th International Conference on Solid-State Sensors, Actuators and Microsystems*. 2005, Seoul, IEEE.
60. Nilsson, D., S. Jensen, and A. Menon, *Fabrication of silicon molds for polymer optics* J. Micromech. Microeng., 2003. **13**: p. S57–S61.
61. Agarwal, R., S. Samson, and S. Bhansali, *Fabrication of vertical mirrors using plasma etch and KOH:IPA polishing*. J. Micromech. Microeng., 2007. **17**: pp. 26–35.

62. Teo, S.H.G., et al., *High resolution and aspect ratio two-dimensional photonic band-gap crystal*. J. Vac. Sci. Technol., 2004. **B 22**(6): p. 2540.
63. Kawata, H., M. Yasuda, and Y. Hirai, *Fabrication of Si mold with smooth side wall by new plasma etching process*. Microelectronic Engineering, 2007. **84**: pp. 1140–1143.
64. Wang, X., et al., *Fabrication of nanoimprint template in Si with high etch rate by non-switch DRIE process*, in *33rd International Conference on Micro- and Nano-Engineering*. 2007, Copenhagen, Denmark.
65. Welch, C.C., et al., *Silicon etch process options for micro- and nanotechnology using inductively coupled plasmas*. Microelectronic Engineering, 2006. **83**: pp. 1170–1173.
66. *Thick SiO_2 etching*. [cited; Available from: Adixen http://www.adixen.com/].
67. Deng, L., et al., *ICP etching of InP and related materials using photoresist as mask*. Proc. SPIE, 2004. **5280**: pp. 838–843.
68. Gottscho, R.A. and C.W. Jurgensen, *Microscopic uniformity in plasma etching*. J. Vac. Sci. Technol., 1992. **B10**(5): p. 2133.
69. Karttunen, J., J. Kiihamäki, and S. Franssila, *Loading effects in deep silicon etching*. Proc. SPIE, 2000. **4174**: pp. 90–97.
70. Chen, A., et al., *Fabrication of sub-100 nm patterns in SiO_2 templates by electron-beam lithography for the growth of periodic III–V semiconductor nanostructures*. Nanotechnology, 2006. **17**: pp. 3903–3908.
71. Feurprier, Y., et al., *Microloading effect in ultrafine SiO_2 hole/trench etching*. J. Vac. Sci. Technol., 1999. **A17**(4): p. 1556.
72. Dalton, T.J., J.C. Arnold, and H.H. Sawin, *Microtrench formation in polysilicon plasma etching over thin gate oxide*. J. Electrochem. Soc., 1993. **140**(8): p. 2395.
73. Hoekstra, R.J. and M.J. Kushner, *Microtrenching resulting from specular reflection during chlorine etching of silicon*. J. Vac. Sci. Technol., 1998. **B16**(4): p. 2102.
74. Schaepkens, M. and G.S. Oehrlein, *Asymmetric microtrenching during inductively coupled plasma oxide etching in the presence of a weak magnetic field*. Appl. Phys. Lett., 1998. **72**(11): p. 1293.
75. Watanabe, M., D.M. Shaw, and G.J. Collins, *Reduction of microtrenching and island formation in oxide plasma etching by employing electron beam charge neutralization*. Appl. Phys. Lett., 2001. **79**(22): p. 2698.
76. Kim, J., et al., *Robust SOI process without footing and its application to ultra high-performance microgyroscopes*. Sensors and Actuators A: Physical, 2004. **114**: pp. 236–243.
77. Lasky, J.B., *Wafer bonding for silicon-on-insulator technology*. Appl. Phys. Lett., 1986. **48**: pp. 78–80.
78. Hwang, G.S. and K.P. Giapis, *On the origin of the notching effect during etching in uniform high density plasmas*. J. Vac. Sci. Technol., 1997. **B15**(1): p. 70.
79. Kim, C.-H. and Y.-K. Kim, *Prevention method of a notching caused by surface charging in silicon reactive ion etching*. J. Micromech. Microeng., 2005. **15**: pp. 358–361.
80. Laermer, F. and A. Urban, *Challenges, developments and applications of silicon deep reactive ion etching*. Microelectronic Engineering, 2003. **67–68**: pp. 349–355.
81. Wilkinson, C.D.W. and M. Rahman, *Dry etching and sputtering*. Phil. Trans. R. Soc. Lond., 2004. **A362**: pp. 125–138.
82. Li, X., et al., *A low damage RIE process for the fabrication of compound semiconductor based transistors with sub-100 nm tungsten gates*. Microelectronic Engineering, 2006. **83**: pp. 1159–1162.
83. Maat, S., et al., *Ultrathin CoPt-pinned current perpendicular to the plane spin valves*. J. Appl. Phys., 2005. **98**(11).
84. Moneck, M.T., et al., *Fabrication of flyable perpendicular discrete track media*. IEEE Trans. Magn., 2007. **43**(6): pp. 2127–2129.
85. *Ion beam milling information*. [cited; Available from: http://www.ionbeammilling.com/].

86. Wang, X., et al., *Reactive ion beam etching of HfO_2 film and removal of sidewall redeposition*. J. Vac. Sci. Technol., 2006. **A24**(4): p. 1067.
87. Stade, F., et al., *Fabrication of metallic nanostructures for investigating plasmon-induced field enhancement*. Microelectronic Engineering, 2007. **84**: pp. 1589–1592.
88. Walsh, M.E., et al., *Optimization of a lithographic and ion beam etching process for nanostructuring magnetoresistive thin film stacks*. J. Vac. Sci. Technol., 2000. **B 18**(6): p. 3539.

Chapter 7
Indirect Nanofabrication

7.1 Introduction

The nanofabrication technologies introduced in the previous chapters so far have a common feature, that is, the dimension of a structure or pattern is defined directly by lithographic means, whether through a mask (optical lithography or charged particle projection lithography), or through direct patterning (electron beam (e-beam), focused ion beam (FIB), scanning probes), or with a mold (nanoimprint, microcontact printing). The structural dimension is more or less determined by the resolution capability of a particular lithographic technique, though pattern transfer can cause, to some extent, a deviation in the final dimension. Each lithographic technique has its own capability and limitation. Table 7.1 summarizes the capabilities of each of these direct nanofabrication techniques introduced in the previous chapters. These figures only serve as references. It is difficult to define the ultimate capability of a particular nanofabrication technique, because the resolution limit is very much dependent on what equipment is used and at what process conditions. For example, pattern features less than 10 nm have been fabricated by e-beam lithography [1], though in Table 7.1 the resolution capability of e-beam lithography is quoted as 10 nm. However, these individual cases reported are exceptions rather than representative of the current level of e-beam lithography. The capability figures listed in Table 7.1 are based on majority of reported results and they are generally reproducible with certain type of equipment and process conditions no matter who is doing it or where the fabrication is done.

Table 7.1 Capabilities of direction nanofabrication techniques

Technology	Capability
Optical lithography	~30 nm
Electron beam lithography	~10 nm
Nanoimprint lithography	~10 nm
Scanning probe lithography	~20 nm
Reactive-ion etching	~50 nm
Metal lift-off	~10 nm
Focused ion beam	~30 nm

Z. Cui, *Nanofabrication*, DOI: 10.1007/978-0-387-75577-9_7,

© Springer Science+Business Media, LLC 2008

One should be aware that to achieve the resolution capability quoted in Table 7.1, one is assumed to have access to the state-of-the-art equipments, which is a luxury many may not have. Some nanofabrication practitioners may have to go with whatever equipments they can access and try their best to explore the potential of existing fabrication equipments, or to find ways to get around the resolution limit. The indirect nanofabrication technologies to be introduced in this chapter are those clever ways of getting around the limits of existing fabrication equipment. They are indirect because the pattern dimension is not predefined by a direct lithographic technique. Here is a simple example: if one cannot make a very narrow line pattern, say sub-10 nm, by optical or e-beam or scanning probe, one can start with a wide line pattern and take away bit by bit from both sides until it becomes very narrow. This is the sort of indirect nanofabrication technique to be discussed and the principle can be applied to many other ways of narrowing down a pattern structural dimension.

It is difficult to categorize each indirect nanofabrication technique because there can be many diverse methods to overcome the existing fabrication limits. The above example can be considered as an extension of pattern transfer techniques introduced in Chapter 6. It is a lateral subtraction instead of vertical, without direct involvement of lithographic patterning. Other methods utilize material deposition to create pattern structures whose lateral dimension is not defined by lithography but by the thickness of the deposited film. Generally, indirect nanofabrication is not a straightforward "lithography plus pattern transfer" process, but rather more complex and several process steps may be involved. In most cases they do not have the flexibility and versatility enjoyed by direct nanofabrication technologies; for example, making arbitrary shapes of patterns can be difficult with indirect techniques. Nevertheless, indirect techniques can be very useful to break the limit of current nanofabrication technologies. In fact, indirect techniques have been vigorously pursued currently by semiconductor industry in an attempt to extend the life of deep-UV projection optical lithography to future generation of IC manufacturing.

7.2 Sidewall Lithography

Sidewall lithography was the name given to a technique which achieves narrow width of pattern structure by thin-film coating on the sidewalls of supporting structures and selectively etching away the supporting structures. The technique is schematically explained in Fig. 7.1. The supporting pattern structures are first made with low-resolution direct lithographic process, for example, optical lithography (Fig. 7.1a). The second step is to conformably coat a thin film over the supporting structure (Fig. 7.1b). The thin film has to be a material different from the supporting structure. A selective reactive-ion etching (RIE) process removes the top cover of the supporting structure (Fig. 7.1c). The final step is to selectively etch away the supporting structure, leaving the sidewall coating

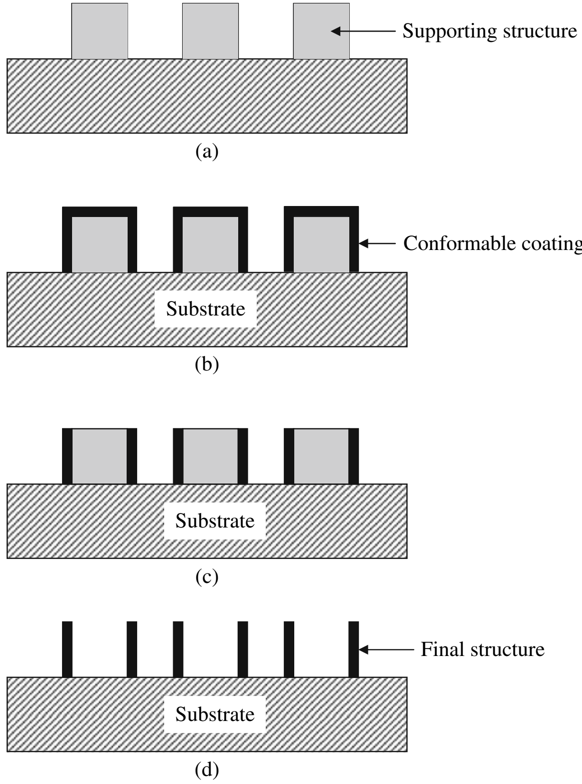


Fig. 7.1 Schematic of sidewall lithography process. (**a**) Patterning supporting structure, (**b**) conformable coating, (**c**) removal of top cover thin film by RIE, (**d**) removal of supporting structure

standing alone as the final pattern (Fig. 7.1d). It is apparent that the lateral width of final pattern structure depends only on the thickness of thin film coated on the sidewalls of supporting structures. So is the name of sidewall lithography. There are two critical requirements that enable the sidewall lithography: the supporting structures must have a vertical sidewall, so that when the top cover is removed by RIE there is a minimum removal of the sidewall coating material; the two materials for the thin film and the supporting structure have to be incompatible in RIE or chemical wet etching, so that the supporting structure can be removed without damaging the sidewall thin-film structure.

Sidewall lithography was probably the earliest indirect patterning technique proposed to break the resolution limit of conventional lithography. The technique was first reported in 1984 by IBM with the name of sidewall image transfer (SIT) process [2]. The original process used photoresist to pattern the supporting structure. Conformable coating of SiO_2 was performed using plasma-enhanced chemical vapor deposition (PECVD). Reactive-ion etching using CF_4 gas then removed the top cover SiO_2, followed by O_2 RIE to remove

the photoresist supporting structure. Later a much simple process was proposed, in which the supporting photoresist was silylated to replace the conformable thin-film coating step [3]. Silylation is a chemical reaction between the novolak resin within a photoresist and silylation agent, which builds silicon atoms into the resin so that the photoresist becomes etch resistant to oxygen plasma. The diffusion depth of silicon atoms into photoresist depends on the silylation conditions and can be well controlled. The silylated photoresist sidewall pattern, after removing the nonsilylated photoresist, can be used as a mask to further transfer into substrate. Minimum 50-nm wide and 200-nm-high silicon nitride patterns have been fabricated by RIE transfer of sidewall patterns, which was inconceivable at the time by any other lithographic techniques.

Since it was reported, sidewall lithography has been continuously employed till present day to make fine pattern structures without using high-resolution lithography tools [4]. The basic concept is the same. The only variations are different combination of thin film and supporting structure materials. There have been reports of using platinum coating on amorphous silicon to make 50-nm Pt patterns [5], using SiO_2 as the supporter and Si_3N_4 as the thin film to make 100-nm line and space grating for nanoimprint [6], using polysilicon and SiO_2 combination to make sub-20 nm SiO_2pattern structures [7], or using patterned silicon as supporting structures and thermal oxidation process to modify the sidewall surface of the silicon supporters while protected by Si_3N_4 at the top [8], which is very similar to the aforementioned photoresist silylation process. The oxide thickness can be well controlled by oxidation time. The final sidewall oxide pattern is revealed after RIE removal of the silicon pattern structure. Figure 7.2 shows a sidewall oxide pattern which is approximately 40 nm in width. Minimum 20-nm pattern width has also been achieved by this process.

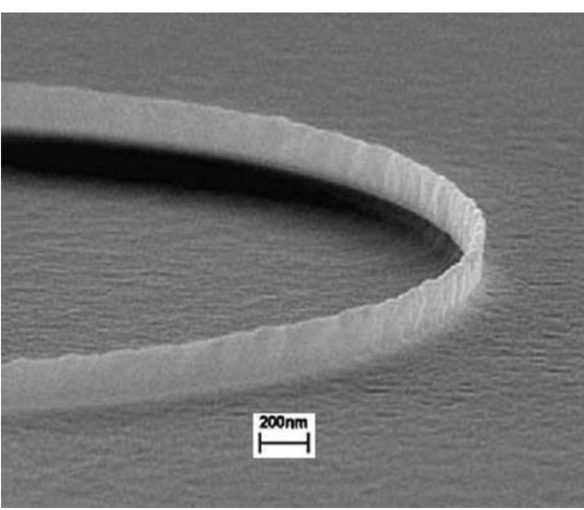


Fig. 7.2 Silicon dioxide pattern of 40-nm width obtained by sidewall lithography (Reprint from [9] with permission)

Whatever the combination of materials is, as long as the two aforementioned critical requirements are observed, sub-100 nm patterns can be made with low cost, albeit with a slight increase in process complexity. However, the flexibility of making any shapes of 2D pattern layout is restricted with the technique, because the final pattern layout relies on the pattern of supporting structures. Line and space pattern is probably the easiest and more effective pattern structure to make by the sidewall lithography, though other pattern layouts are possible by ingenious design of the supporting patterns.

Even though the sidewalls of supporting structure are not vertical, it is still possible to mold the sidewall shape by the sidewall lithography. Figure 7.3 explains schematically how a conical thin-walled microstructure can be made by the sidewall molding technique. A metal thin film is conformably coated

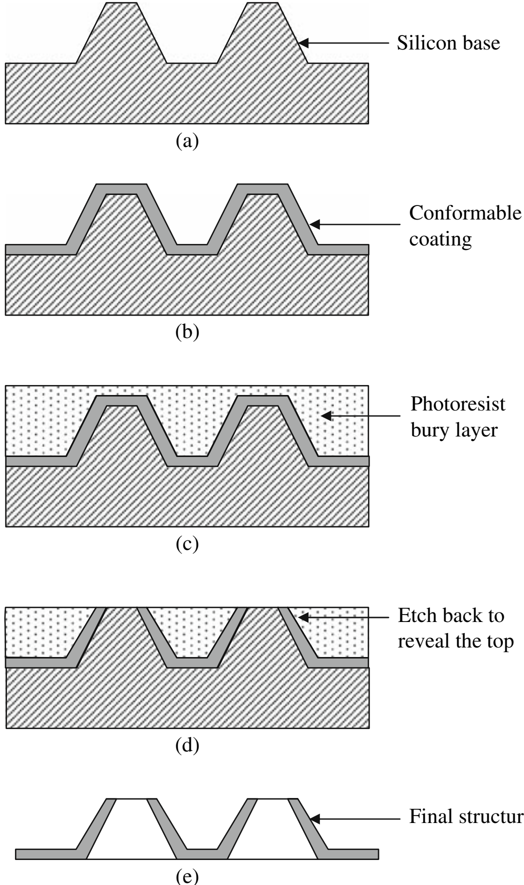


Fig. 7.3 Indirect patterning from sloped sidewalls. (**a**) Initial silicon base, (**b**) conformable coating, (**c**) photoresist planarization, (**d**) etch back, (**e**) released final structure

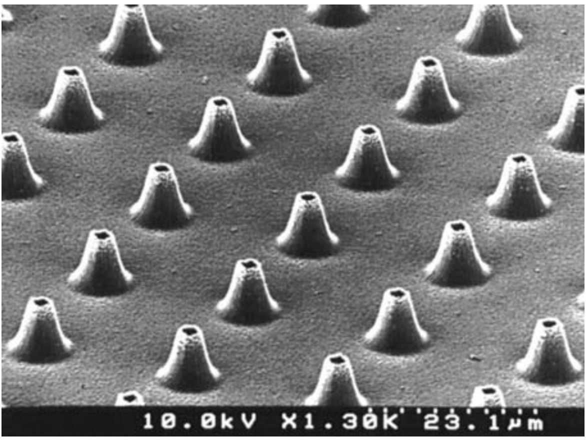


Fig. 7.4 Scanning electron microscope photo of micronozzle array made by sidewall patterning

onto a silicon base which has an array of cone-shaped structures. To protect the sidewalls while etching the top cover film, a thick layer of photoresist is coated to bury the pattern structures. An etching-back in oxygen plasma removes the top layer of photoresist, just to reveal the top surface of coated silicon cones. A wet chemical etch removes the top cover metal and a subsequent silicon etch removes the silicon base, as well as the photoresist bury layer. The final structure is an array of hollow metallic micronozzles as the scanning electron microscope (SEM) photo shown in Fig. 7.4 [10].

7.3 Dimensional Subtraction and Addition

7.3.1 Lateral Subtraction

Taking away bit by bit from sideway of a wide line to make it narrow is the idea of indirect patterning by lateral subtraction. The isotropic nature of wet chemical etching or RIE offers the possibility of subtracting materials laterally from a patterned structure. One of such schemes is shown schematically in Fig. 7.5. It shows how a silicon nanowire can be made by KOH etching of silicon laterally [11]. Because alkaline wet chemical etching of silicon is highly dependent on crystalline orientation, it is possible to control lateral etching of silicon fast in one direction but slow in another direction. As shown in Fig. 7.5, by aligning the two oblique sides of pattern mask to the <100> direction, the KOH removes silicon fast in the <100> orientation but very slow in the <111> direction. The fast-forward etching leaves a very narrow silicon line between the two dislocated silicon pads. Figure 7.6 shows four consecutive SEM photos of silicon etching progress, showing how the silicon nanowire is formed. By careful design of etching mask and fine control of etching time, a 40-nm wide and 700-nm-long silicon nanowire has been fabricated [11].

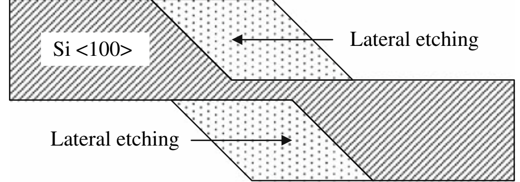


Fig. 7.5 Top view of etching mask pattern, the two horizontal arrows indicate the Si edge movement underneath the etch mask

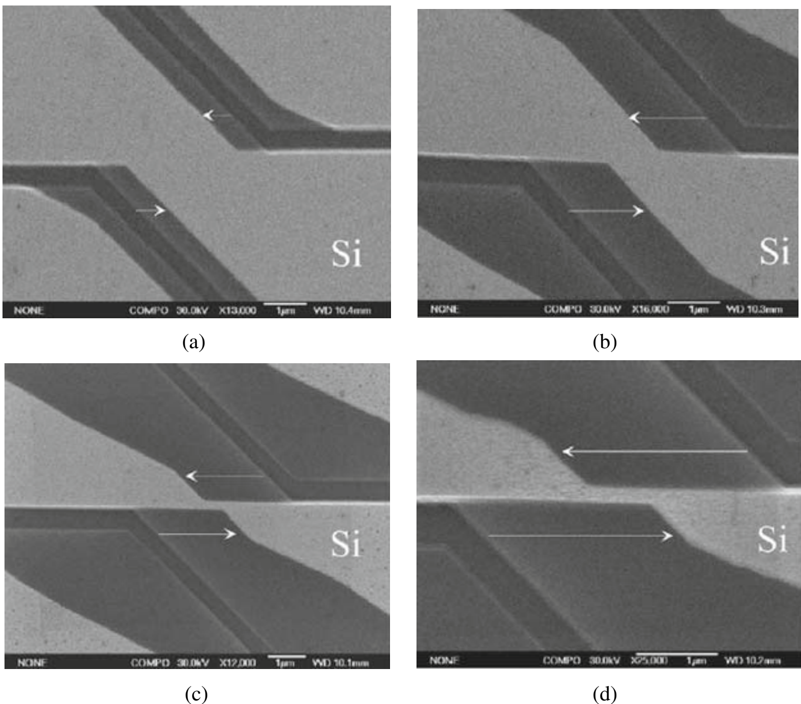


Fig. 7.6 Scanning electron microscope photos showing the formation process of silicon nanowire at different etching time: (**a**) 5 min, (**b**) 10 min, (**c**) 15 min, and (**d**) 20 min. The *white arrows* indicate the etching advance direction (Reprint from [11] with permission)

Lateral subtraction in RIE is much easier to implement. As introduced in Chapter 6, RIE always has certain degree of isotropic etching. While the etching is going deeper, it is also going wider laterally. The anisotropy can be controlled by tuning the plasma condition. Figure 7.7 shows a series of SEM photos of etched silicon stripes [12]. It demonstrates how the stripes are narrowed with the progression of etching time. Eventually, the structure dimension at the top can be made infinitely small. This is how very sharp silicon tips are made, as shown in Fig. 7.8.

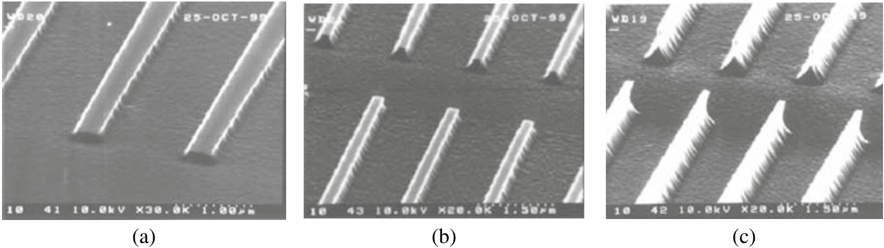


Fig. 7.7 Lateral subtraction of silicon line pattern in RIE: (**a**) etching 1 min, (**b**) etching 2 min, (**c**) etching 3 min

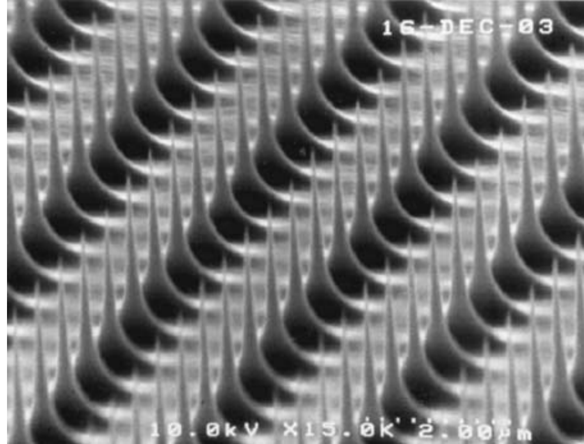


Fig. 7.8 Silicon tips made by RIE with nanometer-scale apex

Another lateral subtraction technique, particularly suitable for scaling down silicon structures, is oxidation sharpening. It is known that silicon can be turned into silicon dioxide in thermal oxidation process. Once becoming oxide, it can be etched away in hydrogen fluoride acid. The process is schematically illustrated in Fig. 7.9. An initial large silicon cone structure is thermally oxidized. The thermal oxidation "consumes" away the silicon from surface down. The longer the thermal oxidation process the more the silicon is converted into silicon dioxide. After the thermal oxidation, a HF wet etching removes the oxide "shell", leaving a much sharpened silicon tip. Generally, a "fat" structure can be "slimmed" down by the oxidation and HF etching. Oxidation sharpening technique was first reported in 1991 [13]. In that report, the thermal oxidation was carried out at 950°C. Higher temperature may cause the oxide to flow because the viscosity is reduced at oxidation temperature >1050°C, which may induce stress in a high-curvature structure such as a sharp silicon tip. Since the oxidation slows down with the oxidation layer growing thicker, multiple oxidation-etching may be performed if the removal of silicon is significant.

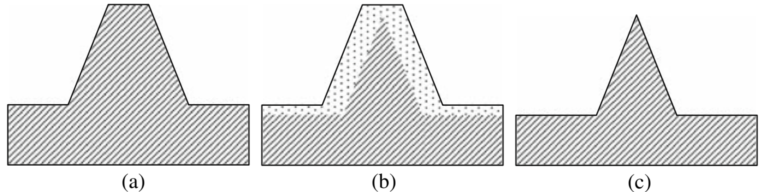


Fig. 7.9 Schematic of oxidation sharpening process. (**a**) Initial tip geometry, (**b**) oxidation, (**c**) final tip geometry after removal of oxide layer

Atomically sharp silicon tips, with a tip radius of <1 nm, have been made with two consecutive oxidation and HF etching processes [13]. The oxidation sharpening technique, since proposed, has become a key technology to make sharp silicon tips which have many applications in field electron emission devices and scanning probe microscopes.

Alternatively, the oxide layer can remain while the silicon is etched away, which results in holes or trenches whose dimension is determined by the amount of remaining silicon. Such a process has been utilized to fabricate sub-100 nm fluidic channels on a silicon substrate [14]. The fabrication process is illustrated in Fig. 7.10. Silicon nitride lines of much wider dimension (2.5–3 μm) are patterned first on a silicon substrate (Fig. 7.10a). This can be done by standard low-resolution contact photolithography. Once the silicon nitride lines are defined, wet chemical etching of silicon by KOH is performed at the exposed silicon area (Fig. 7.10b). The etching goes in depth as well as laterally. The next step is to oxidize the silicon (Fig. 7.10c). As seen from the previous oxidation sharpening of silicon tips, the silicon is oxidized in every direction. The longer the oxidation, the less the silicon remains under the nitride mask caps. The final step is to remove the nitride masks and etch away the silicon sandwiched between the oxide walls (Fig. 7.10d). Thin channels are then formed. With this technique nanochannels as narrow as 40 nm have been fabricated without using any expensive lithography tools [15]. A further sputter deposition of oxide can seal off the top so that they become buried nanofluidic channels [14]. This is a typical lateral subtraction process in that the channel dimension is not dependent on lithography but on the lateral oxidation. It takes advantage of the fact that silicon, nitride, and oxide can each be selectively etched without harming the others.

In addition to oxidation, there are other ways of trimming down a “fat” structure. For example, RIE has been used to trim e-beam-exposed hydrogen silsesquioxane (HSQ) resist structure to achieve sub-10 nm pattern resolution [16]. Hydrogen silsesquioxane is a high-resolution negative-tone resist for e-beam lithography, which has been introduced in Chapter 3. Although e-beam lithography is of very high resolution, the proximity effect due to electron scattering has limited the achievable resolution. To further reduce the dimension of exposed feature, RIE has been used to laterally reduce the feature width. It has

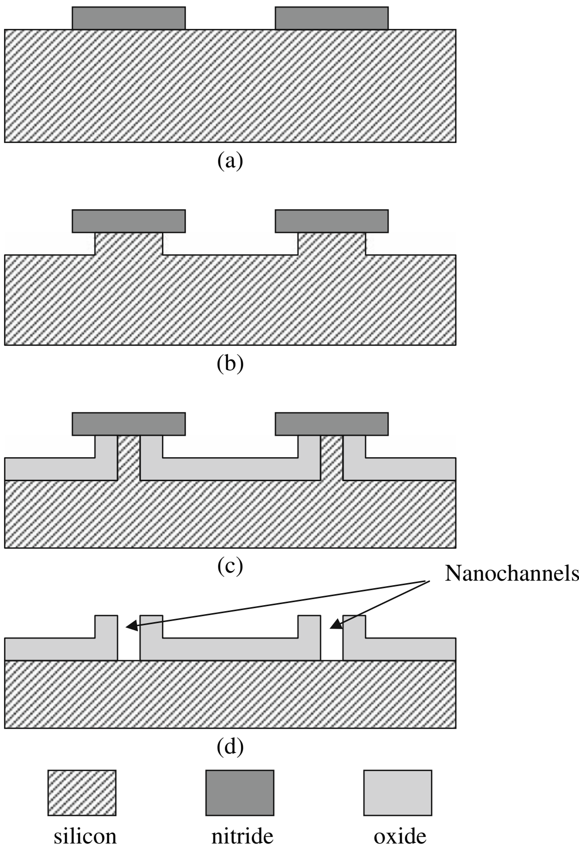


Fig. 7.10 Schematic of nanochannel fabrication by lateral oxidation

been demonstrated that with CF_4 plasma trimming an initial 21-nm HSQ line feature can be reduced to 5.3 nm [16]. The RIE trimming rate was estimated to be about 4 nm min^{-1}. By optimizing the RIE condition to let the plasma etching more isotropic while ensuring the HSQ feature sufficiently tall, the lateral trimming of HSQ feature will not jeopardize the integrity of HSQ pattern. In fact, post-fabrication trimming is a widely used technique for feature size reduction. In some applications, FIB is the best suited tool to perform the trimming task. Due to its scanning and sputter etching nature, FIB is flexible to sputter away materials from a very precise location. The best example of FIB trimming is the manufacturing of magnetic read/write heads, whereby the FIB trimming has become a standard volume manufacturing process. A magnetic read/write head initially patterned by optical lithography can be further trimmed down to sub-100 nm dimension by FIB sputter etching. Figure 7.11 is an example of FIB-trimmed magnetic head [17]. An initial square-shaped pattern has been trimmed into a thin line.

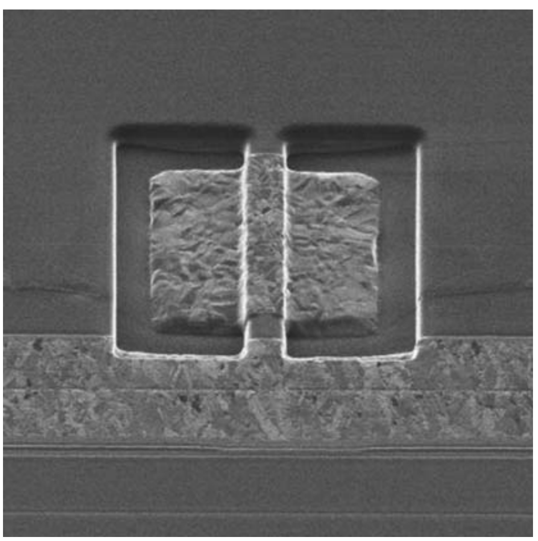
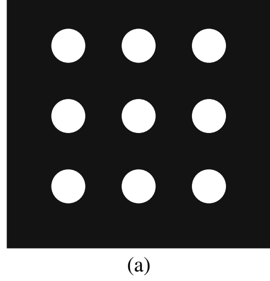
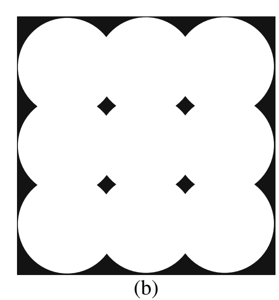


Fig. 7.11 Magnetic read/write head trimmed by FIB (Reprint from [17] with permission)

There can be many clever ways to realize feature size reduction. It is known from Chapters 2 and 3 that photoresists or e-beam resists have positive tone and negative tone. Positive-tone resist dissolves and negative tone remains after exposure and development. Figure 7.12 illustrates how overexposure of positive-tone resist can produce much reduced size of negative-tone resist image. The standard mask pattern looks like what is shown in Fig. 7.12a. By overexposure of the mask pattern, each dot expands laterally. After development, the remaining resist becomes the pattern of reversed tone with much reduced dimension, as shown in Fig. 7.12b. The size and period of reversed-tone dots array can be varied by varying the overexposure dose and the initial mask pattern design. It has been shown that for initial 125-nm positive-tone dot design the exposure at 230 mJ cm^{-2} can produce 107-nm negative-tone dots, against the normal exposure dose of 110 mJ cm^{-2}. At 350 mJ cm^{-2}, the negative-tone dots are reduced to 53 nm [18].



(a)



(b)

Fig. 7.12 Feature size reduction by overexposure of positive-tone resist. (**a**) Initial positive-tone pattern. (**b**) Final negative-tone pattern after overexposure of the pattern in (**a**)

7.3.2 Lateral Addition

Lateral subtraction can produce narrow lines, as mentioned above. The opposite is to produce narrow gaps by lateral addition. Conventional technique to create a narrow gap is to make a narrow line pattern by e-beam exposure of a positive-tone resist, then transfer the resist pattern into substrate by RIE, or a narrow gap is created by direct FIB cutting of a continuous line. Narrow gaps produced by both of the techniques are restricted by the diameter of beam, resolution of the resist, and proximity effect due to electron scattering. The lateral addition technique is to start with a wider gap and then narrow it by adding material bit by bit to close in the gap. The initial gap can be produced by conventional lithography and etching, as shown in Fig. 7.13a. Subsequent deposition of material closes the gap by hanging over the gap, as shown in Fig. 7.13b. By carefully controlling the initial gap distance and material deposition, 2–3-nm gap between two gold electrodes has been created [19].

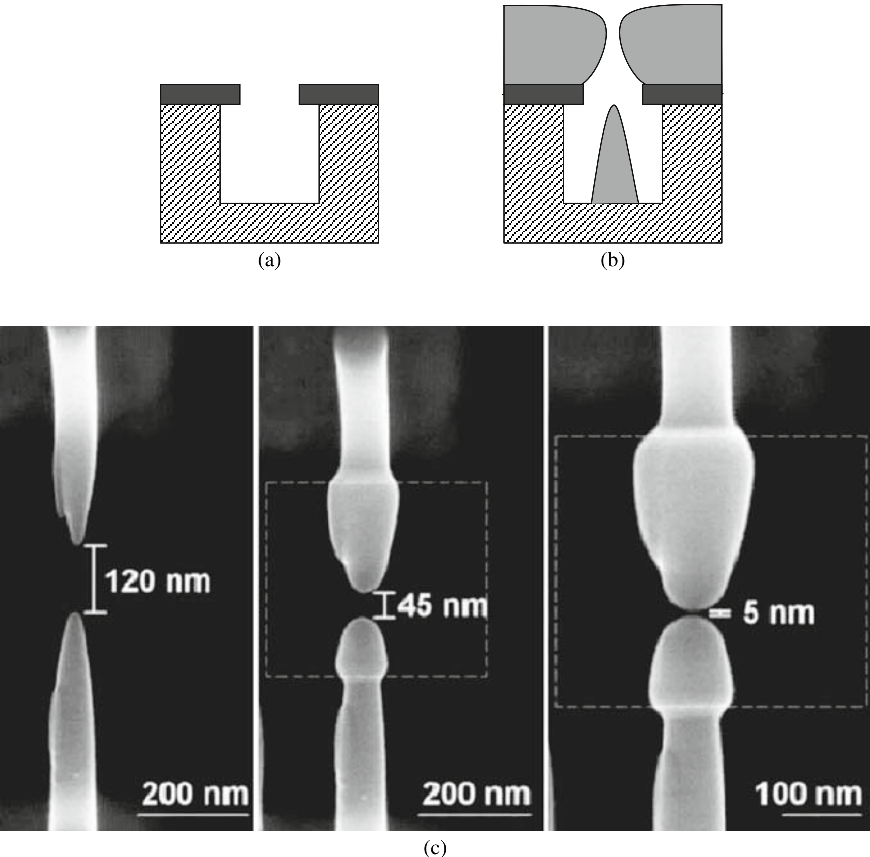


Fig. 7.13 Creation of narrow gap by lateral addition of material. (**a**) Initial gap produced by conventional lithography and etching, (**b**) close the wide gap by material deposition, (**c**) narrowing a wide gap by e-beam induced deposition (Reprint from [20] with permission)

The most flexible lateral addition process is to utilize e-beam- or FIB-induced deposition, by which the amount of material deposited and the location of deposition can be precisely controlled. Figure 7.13c shows a sequence of processes to produce a narrow gap by e-beam induced deposition. It demonstrates that the initial 120-nm gap has been narrowed to 5 nm by gradual deposition of material.

The same principle can be applied to shrink holes. As illustrated in Fig. 7.14, the initial holes made by optical lithography may be big (Fig. 7.14a). By depositing a thin film the diameter of the holes can be reduced, depending on the thickness of thin film (Fig. 7.14b). Plasma etching-back removes the thin film from the top and bottom of the holes (Fig. 7.14c), but leaves the sidewall thin film intact. A subsequent pattern transfer by RIE results in much smaller holes in the substrate than original photoresist pattern (Fig. 7.14d). It has been demonstrated that the technique can shrink original 100-nm holes to 40 nm. This nonlithographic technique can potentially produce holes as small as 10 nm [21]. The only catch is that the pitch of the hole array is also changed, accompanying the reduction of hole diameter. Nevertheless, its potential in further reduction of feature size beyond the resolution of current optical lithography has certainly provided an alternative for manufacturing of future generation of integrated circuits.

It has been found that ion irradiation of a free-standing thin film can induce free movement of atoms in the thin film. If there is a hole in a free-standing thin film, ion beam irradiation of the hole and surrounding area can cause the hole to shrink, even close up completely. This so-called "ion beam sculpting" technique has been used to produce nanopores of less than 10 nm in thin membranes [22]. Conventionally, nanopores are made by RIE or direct FIB etching of a membrane. The direct technique is always limited by the resolution of lithography tool and process. The smallest hole etched by FIB in a 30-nm thick silicon nitride membrane is around 30 nm [23]. Further reduction of nanopore size beyond the resolution limit of FIB can be realized by the ion beam sculpting process. For example, an initial 100-nm hole in a 100-nm thick silicon nitride membrane has been reduced to 15 nm after 18 scans of Ga ion beam in a 1–2-μm^2 area surrounding the hole. The complete process takes about 1 min.

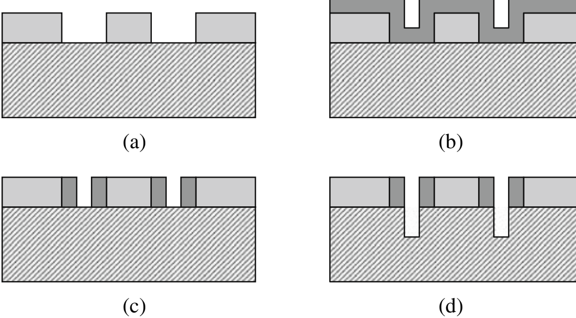


Fig. 7.14 Creation of nanoholes by lateral addition of material. (**a**) Original photoresist pattern, (**b**) thin film deposition, (**c**) plasma etching back, (**d**) transfer to substrate

Figure 7.15 shows a series of SEM photos of a nanopore, representing gradual reduction of pore size from 100 nm to 5 nm through scanning of FIB. The nanopore can be made even smaller if electrons instead of ions are used. By employing 200-keV e-beam in a transmission electron microscope (TEM), the 5-nm nanopore shown in Fig. 7.15e has been reduced to 1 nm. The exact mechanism of beam-induced nanopore shrink is still not clear. It is believed that high-energy ions or electron irradiation of the thin membrane has increased the mobility and diffusivity of atoms surrounding a hole. The "melting" of material causes the close up of nanopore.

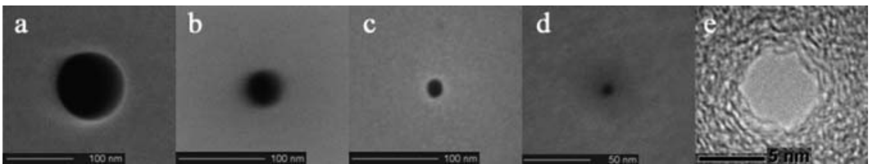


Fig. 7.15 Reduction of nanopore size by ion beam irradiation: (**a**) 100 nm, (**b**) 40 nm, (**c**) 20 nm, (**d**) 10 nm, (**e**) 5 nm (TEM image) (Reprint from [23] with permission)

7.3.3 Vertical Subtraction

Vertical subtraction is the opposite of sidewall lithography. In sidewall lithography technique, the thin film deposited on the sidewalls of supporting structure serves as the final pattern structure once the supporting structure is removed. In vertical subtraction, the sidewall thin film serves as a sacrificial layer. Its thickness determines the dimension of vertical thin channels. The fabrication process is illustrated in Fig. 7.16. First, supporting structures are made on a silicon-on-insulator (SOI) substrate (Fig. 7.16a). They are etched into the top silicon layer and stopped at the SiO_2 insulating layer. Thermal oxidation converts the sidewalls of etched structure into SiO_2 (Fig. 7.16b). Next, deposition of silicon nitride by PECVD fills up the etched cavities (Fig. 7.16c). A chemical mechanical polishing (CMP) removes the top layer nitride to reveal the SiO_2 grown on the sidewalls (Fig. 7.16d). Deep silicon etching removes the bottom-layer silicon and a thin membrane is created (Fig. 7.16e). A final etching of SiO_2 between the silicon supporting structure and the silicon nitride filling creates the fine channels through the membrane (Fig. 7.16f), which can serve as a filter to perform nanoscale filtration. It is apparent that the width of the channels depends only on the thickness of thermal oxide, which can be controlled below 100 nm. The technique has been used to create lateral nanochannel for nanofluidic application [24, 25]. Same as in the sidewall lithography, the only condition for this technique to work is to make sure the materials for the supporting structure and the sacrificial structure

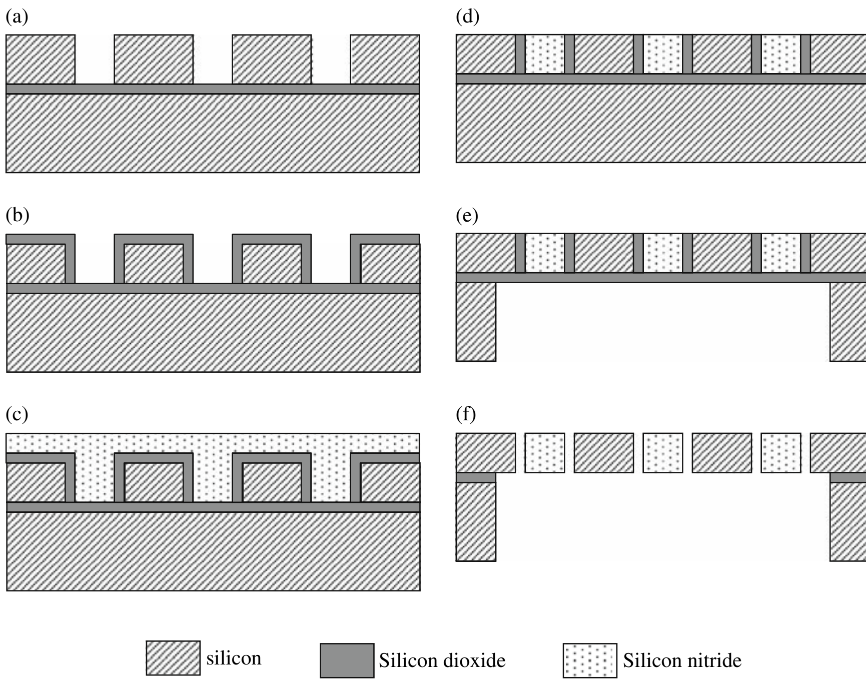


Fig. 7.16 Nanoscale filter made by vertical subtraction technique. (**a**) Supporting structure, (**b**) thermal oxidation, (**c**) silicon nitride filling, (**d**) surface polishing back, (**e**) creating membrane by bottom side etching, (**f**) etching silicon dioxide to create the nanochannels

have different etching chemistry; that is, etching one material does not damage the other.

See-through holes or channels can also be created by anisotropic wet chemical etching of silicon. As described in Chapter 6, wet chemical etching of silicon in alkaline solution is dependent on crystalline orientation; i.e. the etching is fast in one direction but slow in another direction. A typical example is the etching of <100> silicon in KOH. The slow etching in <111> direction creates an etched profile with sloped sidewall. By controlling the mask opening size and etch depth, a reversed cone-like cavity can be created as shown in Fig. 7.17a. In this case, the reversed cone is made in a SOI substrate and the etching is controlled to stop just when the tip of the reversed cone reaches the SiO_2 insulating layer. A second deep silicon etching from the bottom side removes the silicon underneath the reversed cone. A further etching of the SiO_2 layer reveals the cone tip which becomes a see-through hole in the silicon membrane, as shown in Fig. 7.17b. Depending on how precisely one can control the etch depth, the size of see-through hole can be made as small as 50–100 nm. The technique has been used to fabricate nanoscale stencil mask, through which metal patterns are deposited with the smallest feature dimension of 100 nm [26].

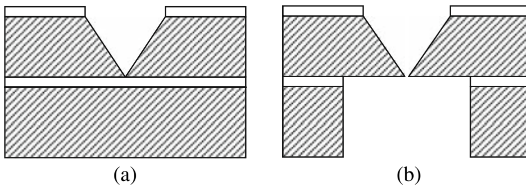


Fig. 7.17 Wet chemical etching of silicon to create nanoscale see-through hole. (**a**) Reversed cone is made in the top layer silicon of a SOI wafer. (**b**) Removal of bottom layer silicon to reveal the hole

The technique is very simple and low cost, because only conventional photolithography and wet chemical etching processes are used.

Another simple subtraction technique is to alternatively grow GaAs and AlGaAs layers with precisely controlled thickness, then selectively etch away one of the materials, leaving the other as the final pattern structure. In order to get etchant solution into the buried layers, the substrate after growing the layers has to be cleaved, so that the alternating layer structures are exposed to wet chemical etchant. The process is illustrated in Fig. 7.18. It is known that molecular beam epitaxy (MBE) can grow crystalline materials in atomic layer accuracy. Therefore, very thin layers of crystalline structure can be formed. On the other hand, GaAs and AlGaAs have the property that they can be etched in different acidic solutions. For example, GaAs can be etched away in the mixture of citric acid and H_2O_2 (5:1 mixture) without affecting the $Al_{0.32}Ga_{0.68}As$ [27], or $Al_{0.7}Ga_{0.3}As$ can be etched away in a diluted solution of hydrofluoric acid without affecting the GaAs [28]. After MBE growth, the sample is cleaved to reveal the alternating crystalline layers (Fig. 7.18a). Acidic etching removes one of the crystalline material and leaves the other alternating layers as the densely packed line pattern structure (Fig. 7.18b). It is apparent that these line patterns will have the width equal to the thickness of MBE-grown layers. The technique has been employed to create a spacer to separate two electrodes with the gap distance as small as 23 nm [27]. Another application is to use the densely packed crystalline grating pattern as nanoimprint stamp to replicate the pattern structure into

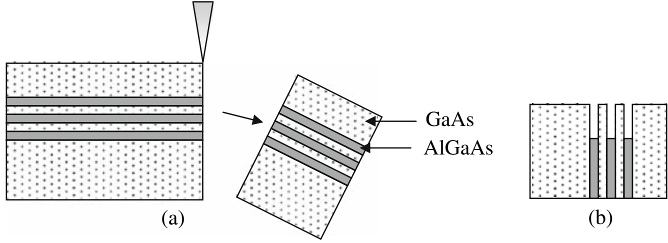


Fig. 7.18 Selective etching of MBE-grown alternating layers. (**a**) MBE growth and sample cleaving, (**b**) selective etching

UV-cured polymer. Dense lines in polymer with half-pitch as small as 6 nm have been created in this way [29].

7.3.4 Nanosphere Lithography

There are many techniques for making dot array pattern without using photons or charged particle beam lithography. Some of the techniques involve self-assembly of molecules, which will be introduced in the next Chapter. One of the simple techniques is nanosphere lithography, which utilizes tightly packed polystyrene spheres on a substrate surface as a masking layer. If a single layer of tightly packed spheres with uniform diameter is lying on a substrate surface, as shown in Fig. 7.19a, the gaps between the spheres form a regular array of dots. Material deposited by thermal evaporation over the spheres-covered surface can only reach the substrate through these gaps. After removing the polystyrene spheres in CH_2Cl_2 acid, similar to the lift-off process introduced in Chapter 6, the substrate surface is left with array of triangular dots, as shown in Fig. 7.19b. This is the principle of nanosphere lithography [30]. Alternatively, assembled silica spheres can be used as an etching mask. Reactive-ion etching reaches to the substrate through the gaps between the spheres. An array of holes are formed in the substrate material, instead of aforementioned dots [31]. From the diameter of these spheres it is not difficult to calculate the size of dots formed through the gaps and distance between the dots. Assuming the sphere diameter is D, the dot formed through a single-layer nanosphere mask will be 0.233 D and the distance between dots is 0.577 D. A double layer of spheres can further reduce the gap, hence the size of dots becomes 0.155 D. The minimum dots that can be made by this technique is about 40 nm. However, both the dot size and the period of dot

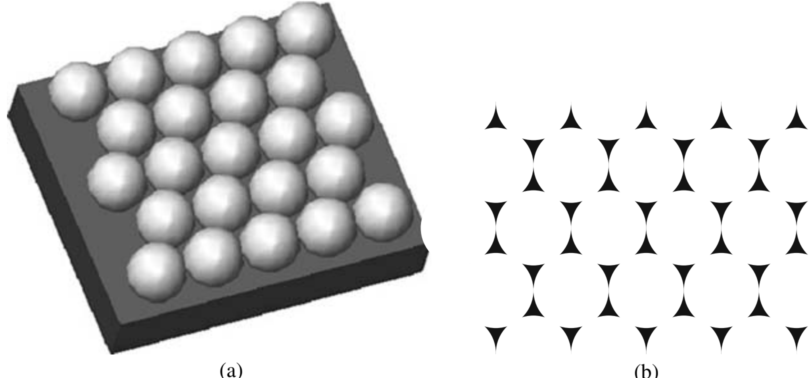


Fig. 7.19 Nanosphere lithography. (**a**) *Polystyrene sphere* covering a flat substrate surface. (**b**) *Dot array* is formed after material deposition and removal of the spheres

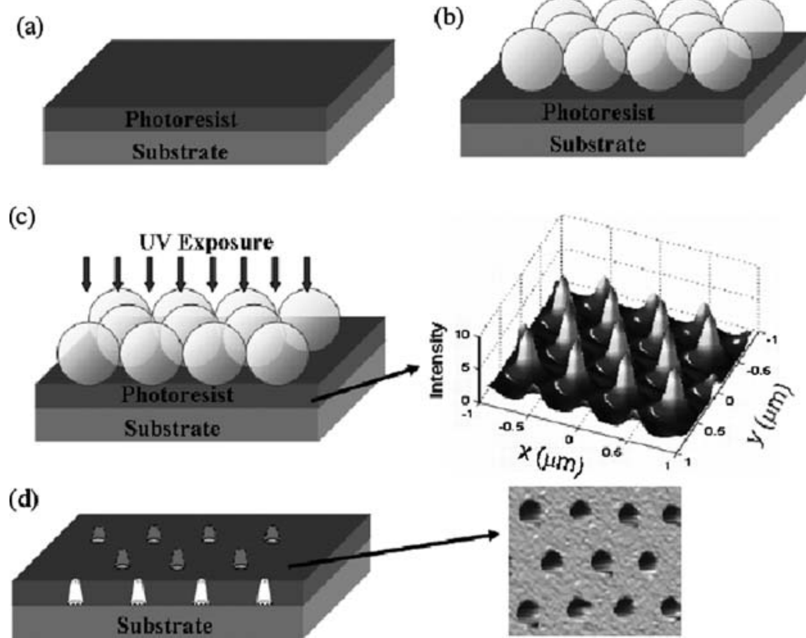


Fig. 7.20 Schematic of nanosphere optical lithography (Reproduced from [32] with permission)

array are restricted by the available size of polystyrene spheres. The lack of flexibility is the major drawback of this technique.

Another kind of nanosphere lithography utilizes self-assembled polystyrene or silica nanospheres on a substrate surface as a lens array. By shining UV light onto the nanosphere lens array, light spots are formed under each nanosphere and a regular array of holes can be created if the underlayer is photoresist [32]. Figure 7.20 illustrates the process of nanosphere optical lithography. With illumination wavelength of 365 nm and conventional photoresist AZ5214, holes with minimum diameter of about 250 nm can be formed in the photoresist layer using silica sphere of ~0.97-μm diameter. The diameter of holes can be controlled by varying optical exposure dose and the pitch can be controlled by using different diameters of spheres. The minimum diameter of light spot formed under the silica sphere is found dependent on the wavelength, not on the diameter of spheres. The ratio of minimum hole diameter to wavelength is about 0.625. Therefore, to pattern sub-100 nm holes, deep UV wavelength will have to be used. The diameter of exposed holes is also dependent on the exposure dose. Higher exposure dose generally creates larger holes. In addition to vertical illumination as illustrated in Fig. 7.20, the nanosphere array can also be illuminated at an angle. In this way, the exposed photoresist pattern no longer remains as holes but takes other shapes. A multiscan of UV light from

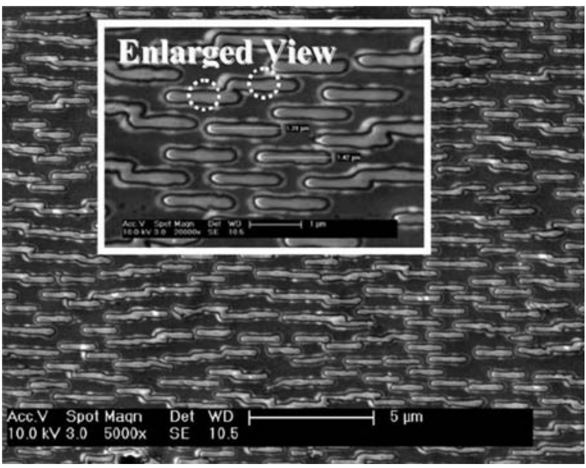


Fig. 7.21 Arbitrary pattern formed by angular scanning exposure through silica sphere array (Reproduced from [33] with permission)

different angles onto the nanosphere-covered surface can create arbitrarily shaped patterns [33]. Figure 7.21 shows an example of short bar patterns created by angular scan of silica sphere array. The angular exposure can also make the pattern dimension smaller. For example, 1-μm silica spheres illuminated by KrF laser (248-nm wavelength) at vertical position ($\alpha = 0$) creates feature size of ~300 nm. If illuminated at an angle ($\alpha = 30$), the feature size is reduced to ~80 nm [33].

The nanosphere array can also be used as a mold to generate hemispherical nanowells on a surface. A laser-assisted embossing process has been developed. Instead of shinning UV light through the nanosphere array, as shown in Fig. 7.20, a laser beam illuminates a mold which consists of array of silica spheres and is in contact with a silicon surface. The laser energy is high enough to melt the surface of silicon substrate. A mechanical pressure is then applied to the nanosphere mold to create hemispherical dents on the soften silicon surface. The size of the hemispherical indentation is obviously dependent on the size of silica spheres, the energy density of the incident laser pulse, and the applied mechanical pressure [34].

7.4 Multistep Processing

Multistep processing is a strategy being widely used to overcome limitations of single-step processing in microscale or nanoscale fabrication. In the previous chapters of this book, several multistep processing techniques have been mentioned. A typical example is the optical and e-beam mix-and-match lithography whereby large-size features are patterned by low-resolution optical lithography and fine features are patterned by high-resolution e-beam lithography. If all the features, large and small, are patterned by e-beam lithography

alone, the exposure time would be extremely long. The two-step lithography can overcome this problem. Top surface imaging (TSI) is another example. Optical or e-beam imaging is done at the top thin layer resist in a multilayer resist stack, which makes high-resolution lithography easier. Once the TSI resist layer is developed, the resist pattern is then transferred into the bottom thick layer resist by RIE. It solves the problem of trying to achieve both high-resolution and high-aspect-ratio patterning in conventional single-layer resist process.

Multistep processing can also be a very effective way of achieving high pattern density. As described in Chapters 2 and 3, high-density pattern structures are always the most difficult pattern structures to fabricate whether by optical lithography or by e-beam lithography. In optical lithography it is the light wave diffraction and in e-beam lithography it is the proximity effect, both of which prevent exposure patterns from being too close to each other. Such a limitation can be overcome by multistep processing, for example, patterning less dense features in separate steps and overlap them to form the final dense features. Figure 7.22 illustrates the basic steps of a double patterning process. The dense pattern structure in substrate is obtained by two steps of resist exposure and pattern transfer. In each step of resist exposure, the pattern density is low and less difficult to achieve.

It is apparent that the high-density pattern fabricated by the double processing is at the expense of more complicated fabrication process. The fabrication cost is high. The multistep processing should only be the last resort when there are no other ways to achieve the required resolution and pattern density. This has been the case in IC manufacturing at below 45-nm half-pitch resolution. As mentioned in Chapter 2, there are no lithography solutions available for 32-nm generation of IC manufacturing. Extreme UV lithography is a possible technology but still in development stage and not ready for volume production. High-numerical aperture (NA) immersion optical lithography at 193-nm wavelength needs to find high-refractive-index liquid and lens materials. The only available solution is the double patterning technique which uses all the existing 193-nm

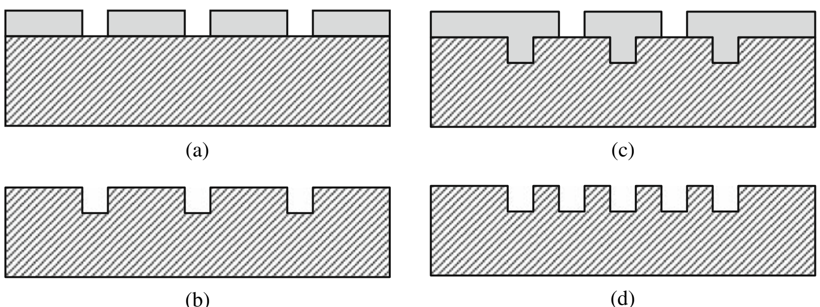


Fig. 7.22 Double patterning process. (**a**) First exposure of resist, (**b**) pattern transfer, (**c**) second exposure of resist, (**d**) pattern transfer

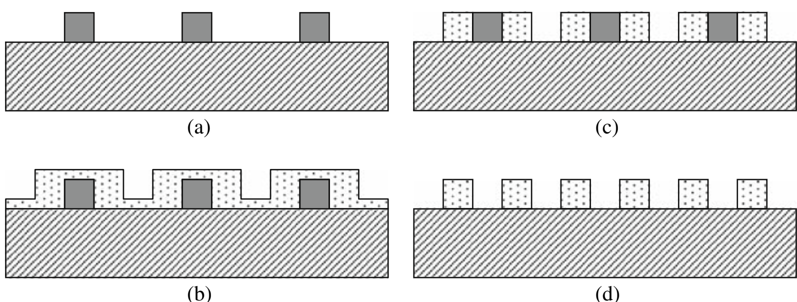


Fig. 7.23 Two-step process to double the pattern density. (**a**) Low density SiO_2 supporting pattern, (**b**) LPCVD Si_3N_4, (**c**) RIE of Si_3N_4, (**d**) removal of SiO_2 supporting pattern by wet etching

lithography processes. However, the double patterning process is by no means low cost and it also sacrifices the throughput because of the double process steps. It is possible to reduce the process cost by double exposure of resist but single-step pattern transfer. This requires that the photoresist is of very high contrast, which ensures the two exposures in the same resist layer do not interfere each other. A compromise is to render the resist photoinsensitive, such as cross-linking, then overcoat with another layer of resist for the second exposure [35]. Even the double-exposure process can still potentially lose 40% of throughput.

Pattern density can also be doubled by a technique similar to the sidewall lithography introduced in Section 7.2. As schematically illustrated in Fig. 7.23, low-density SiO_2 supporting pattern is first made on a silicon substrate (Fig. 7.23a). The substrate is then covered with Si_3N_4 by low-pressure CVD (LPCVD) process (Fig. 7.23b). Reactive-ion etching with CHF_3 plasma removes the Si_3N_4 vertically but keeps the sidewall Si_3N_4 intact because of its anisotropic etching nature (Fig. 7.23c). Finally, the supporting SiO_2 is removed in buffered HF wet etching (Fig. 7.23d). The pattern density is now doubled and the Si_3N_4 pattern becomes the final masking pattern for further transfer into the silicon substrate by RIE. A 100-nm line and space pattern has been achieved using this technique [6].

7.5 Super Resolution Patterning

It is known from Chapter 2 that the ultimate resolution of optical lithography is limited by diffraction. However, in proximity optical lithography, that is, optical imaging by a maskaligner, there is an interesting phenomenon that the light intensity distribution changes its shape with the changing of gap distance between mask and substrate. In principle, the mask should be as close as possible to the substrate so that the mask features can be truthfully printed onto the photoresist layer in 1:1 ratio. Figure 7.24 shows the simulated light intensity distributions at two different gap distances between mask and

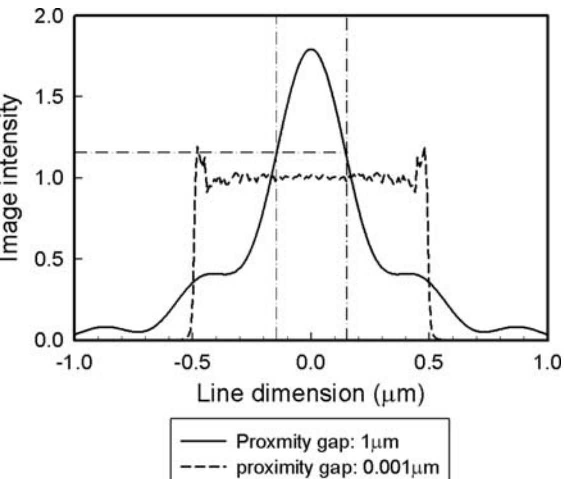
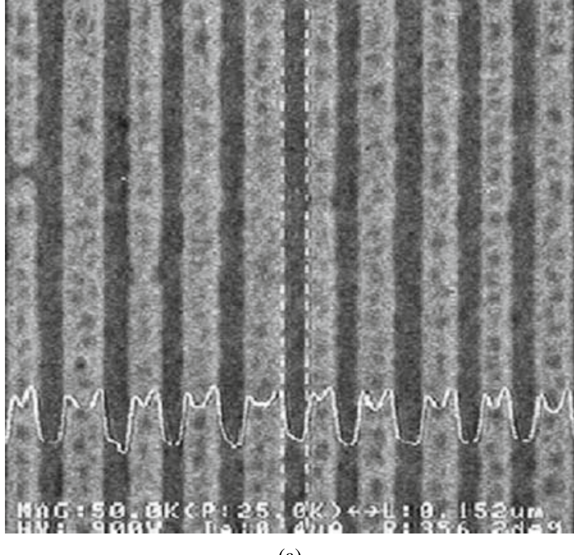


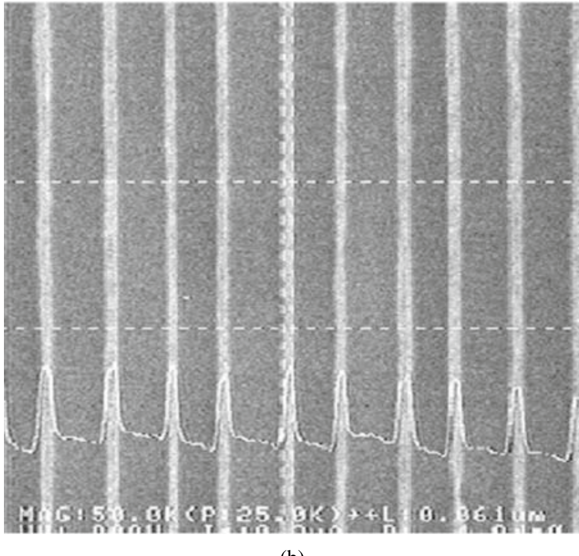
Fig. 7.24 Light intensity distribution on photoresist surface at different mask–substrate distances

substrate (photoresist surface). If the gap distance is infinitely small (assumed 0.001 μm in simulation), the light distribution has almost the same width as the mask feature opening (1 μm). The image in the photoresist is expected to be 1 μm as well. When the mask moves away from the photoresist surface, the light distribution will change due to diffraction. However, there is a gap distance (1 μm in this case) at which the top section of the distribution is much narrower than the opening width of mask feature. Further increasing the gap distance, the distribution becomes flat and the high-resolution peak disappears. If the mask–substrate gap is controlled at this optimum distance and the exposure dose is controlled to enable only the top section of light intensity to participate in the photoresist exposure, then one can expect the photoresist image to be much smaller than the mask feature width. In the simulation example shown in Fig. 7.24, the line feature, obtained by separating the mask from photoresist surface with 1-μm gap and utilizing only the top part of light intensity, will be only 300 nm instead of 1 μm of the mask design.

This "super resolution" property has long been utilized to pattern high-resolution features with a low-resolution mask in proximity optical lithography. The same concept has also been applied to X-ray lithography, since X-ray lithography is a proximity imaging too [36]. Figure 7.25 compares the SEM image of X-ray mask features which is measured to be 152-nm line opening (Fig. 7.25a) and exposed line image in SAL605 negative-tone resist which is measured to be 61 nm (Fig. 7.25b). The demagnification is around 2.5. In order to use a low-resolution mask to print high-resolution patterns, not only should the mask–substrate gap distance be carefully controlled and optimized, but also the control of exposure dose and development time, because only the irradiation near the peak intensity should be allowed to expose the resist. This can be achieved by using low-exposure dose or low-sensitivity resist, or longer development time, all of which have the same purpose, that is, filtering out the contribution of low-intensity



(a)



(b)

Fig. 7.25 Demagnification of X-ray proximity printing. (**a**) Mask feature at 152-nm opening width, (**b**) exposure resist feature at 61-nm line width (Reprint from [36] with permission)

irradiation from the exposure. For example, the 61-nm lines in SAL605 negative-tone resist was obtained by 1 min 45 s development time. The same resist feature became 43–46 nm if developed for 2 min [36]. The longer development time has tailored the exposure level toward higher intensity at which the distribution is narrower. In Chapter 2 it has been shown that the period of grating generated by interferometric lithography can be varied by varying the exposure dose. This is

based on the same principle as mentioned here, which is to move the threshold level up or down the intensity distribution peak, shown in Fig. 2.49, to achieve different line widths in photoresist.

Controlling exposure dose and development time to acquire different exposed feature dimensions in resists is also a common practice in e-beam lithography. It is a well known fact that the highest-resolution e-beam lithography is only achievable in resists having the lowest sensitivity. The reason is that an e-beam has a Gaussian distribution both in beam current and in electron energy, and that only the high-energy electrons and peak intensity of beam current, which is very narrow in distribution, can expose a low-sensitivity resist. The resultant feature will be narrow because the broad portion of the beam, which mainly consists of low-energy scattering electrons and low number of electrons, will have negligible effect on the polymer molecular chains in the resist. Therefore, even with a fixed diameter of beam the line features obtained by single-pass exposure can have different dimensions, depending the exposure dose and development time.

A clever use of beam profile and scattering electrons can create a narrow line feature not by direct exposure of the beam but by scattering electrons from the beam. The technique is illustrated in Fig. 7.26. Normal line exposure by single-pass scanning of a beam is shown in Fig. 7.26a. The scanning is continuous and the overlap of beam spot (shown as the circles) depends on the scanning step size. The actual line width (shown as the solid outlines) then depends on the exposure dose and development condition, which is normally bigger than the beam spot size. If the scan is deliberately broken (e.g., some pixels are blanked during the single pass scan as shown in Fig. 7.26b), the exposed line may not be broken if the blanked gap is sufficiently small, because the resist can still be exposed by scattering electrons and low-energy tail of a Gaussian beam across over the separation. This section of line is expected to be much narrower than the normal line because the exposure dose level is much lower. Figure 7.26c shows the SEM image of a short 3-nm line between the normally exposed lines, created by the aforementioned technique [37]. It can be very tricky to control the marginal exposure of the narrow line. The 3-nm line shown in Fig. 7.26c

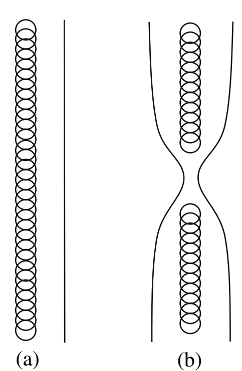


Fig. 7.26 Narrow line feature created by broken line exposure. (**a**) Normal single-pass scan, (**b**) broken single-pass scan, (**c**) SEM image of 3-nm line created between normally exposed line (Reprint from [37] with permission)

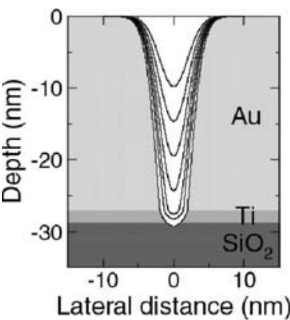


Fig. 7.27 Computer-simulated contours of FIB sputtering of Au–Ti film, which represent the successive sputtering profiles as the sputtering dose increases (Reprint from [38] with permission)

was 30-nm long and at exposure dose of 1.8 nC cm^{-1} (line dose). Reducing the separation to 20 nm, the line was no longer able to delineate, because the exposure contribution from backscattering electron became significant and the broken line was simply merged.

The same practice can be applied to FIB as well. Although the beam current distribution in FIB is generally Gaussian like, it has a larger tail than an e-beam because of the high energy spread in a gallium liquid metal ion source. This is one of the reasons that the resolution of patterning by FIB is generally not as high as that by an e-beam. However, utilizing the beam profile it is possible to fabricate structures much smaller than the beam size. Figure 7.27 shows how a narrow gap can be made by FIB sputter etching [38]. Those are computer-simulated contours of FIB sputtering in thin Au film (27-nm thickness) on top of SiO_2 with a 1.6-nm Ti seed layer. These contours represent the successive sputtering profiles in the Au film as the FIB sputtering dose increases. There is a sputtering dose at which the Au–Ti film is just broken through by the FIB etching and a narrow gap is created in the film. The sputtering profile is a true reflection of the beam profile which is much narrow at the tip. The narrow gap is created by the tip, not by the whole beam, which is the reason that the gap can be so small. Figure 7.28 is the SEM image of an actual 3-nm nanogap created by FIB sputtering with a beam diameter of $\sim$12 nm [38]. To produce this nanogap, it is very critical to be able to control the FIB sputtering to stop just at the moment when the beam tip breaks through the Au–Ti film. To do this, an in-situ measurement of current flowing through the electrode is performed, which is connected to the FIB scan controller. The beam sputtering scan is switched off immediately when the current flow through the thin film electrode drops to zero, which indicates that the electrode has been cut through by FIB sputtering.

The Gaussian distribution of a FIB determines that a hole sputtered by FIB will also be in the shape of a Gaussian profile, as shown in Fig. 7.29a. Then taking away materials layer by layer from the other side would eventually reveal the tip of this Gaussian-shaped hole and create a very small pore whose diameter is determined by the Gaussian distribution width, as illustrated in Fig. 7.29b. The key step in making the nanopore is the precise control of ion

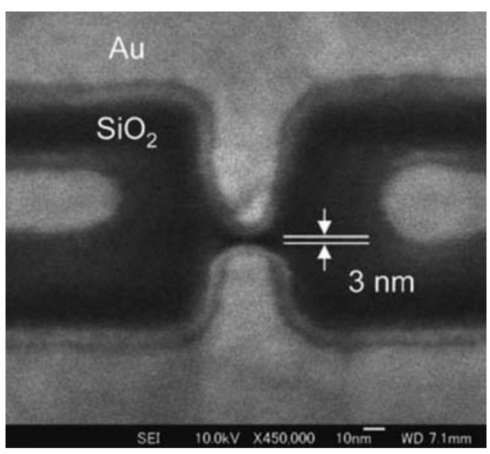


Fig. 7.28 Scanning electron microscope image of 3-nm nanogap created by FIB sputtering with beam diameter of ~12 nm (Reprint from [38] with permission)

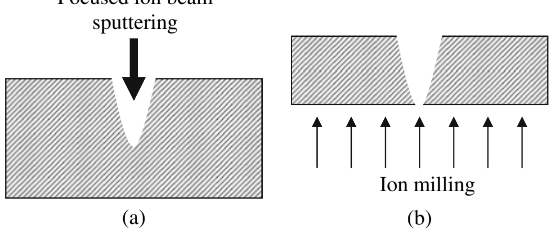


Fig. 7.29 Nanopore produced by FIB sputtering and ion milling. (**a**) FIB sputtering from *top side* of a membrane, (**b**) ion milling from *bottom side* of a membrane

milling depth in the backside. The ion milling has to stop just as the membrane has been broken through. To achieve this precise control, an ion counter can be mounted on the other side of the membrane. When ions start to flow through and are detected by the ion counter, it signals the breakthrough of the membrane and the time to stop the ion milling. The size of nanopore can be further reduced by low-dose exposure of ions after a hole is created, which promotes the mobility of atoms surrounding the hole and shrinks the hole, as mentioned previously. Nanopore with diameter as small as 1.8 nm have been created in Si_3N_4 membrane by the technique [22].

References

1. Hu, W., et al., *Sub-10 nm electron beam lithography using cold development of poly(methylmethacrylate)*. J. Vac. Sci. Technol., 2004. **B22**(4): p. 1711.
2. Johnson, C., et al., IBM Tech. Dic. Bull., 1984. **26**: p. 4587.
3. Vettiger, P., et al., *Nanometer sidewall lithography by resist silylation*. J. Vac. Sci. Technol., 1989. **B7**(6): p. 1756.

4. Park, C.W., O.V. Mena, and J. Brugger, *Patterning of parallel nanobridge structures by reverse nanostencil lithography using an edge-patterned stencil*. Nanotechnology, 2007. **18**: p. 044002.
5. Hsu, D.S.Y., *50 nm linewidth ~platinum sidewall lithography by effusive-source metal precursor chemical deposition and ion-assisted etching*. Appl. Phys. Lett., 1991. **59**(17): p. 2192.
6. Yu, Z., et al., *Fabrication of large area 100 nm pitch grating by spatial frequency doubling and nanoimprint lithography for subwavelength optical applications*. J. Vac. Sci. Technol., 2001. **B19**(6): p. 2816.
7. Choi, Y.-K., et al., *Sublithographic nanofabrication technology for nanocatalysts and DNA chips*. J. Vac. Sci. Technol., 2004. **B21**(6): p. 2951.
8. Zaborowski, M., et al., *Nano-line width control and standards using lateral pattern definition technique*. Microelectro. Eng., 2006. **83**: pp. 1555–1558.
9. Grabiec, P.B., et al., *Nano-width lines using lateral pattern definition technique for nanoimprint template fabrication*. Microelectron. Eng., 2004. **73–74**: pp. 599–603.
10. Cui, Z., *Fabrication of nozzle array by surface moulding technique*, in *International Microprocesses and Nanotechnology Conference, Digest of Papers*. 2002, Tokyo. p. 94.
11. Ciucci, S., et al., *Silicon nanowires fabricated by means of an underetching technique*. Microelectron. Eng., 2005. **78–79**: pp. 338–342.
12. Cui, Z., et al., *Fabrication of magnetic rings for high density memory devices*. Microelectron. Eng., 2002. **61–62**: pp. 577–583.
13. Ravi, T.S., R.B. Marcus, and D. Liu, *Oxidation sharpening of silicon tips*. J. Vac. Sci. Technol., 1991. **B9**(6): pp. 2733–2737.
14. Cho, Y.H., et al., *Fabrication of silicon dioxide submicron channels without nanolithography for single biomolecule detection*. Nanotechnology, 2007. **18**: p. 465303.
15. *Low-cost chip stretches single molecules*. Jan. 2008 [cited; Available from: http://nanotechweb.org/].
16. Baek, I.-B., J.-H. Yang, and W.-J. Cho, *Electron beam lithography patterning of sub-10 nm line using hydrogen silsesquioxane for nanoscale device applications*. J. Vac. Sci. Technol., 2005. **B23**(6): p. 3120.
17. Reyntjens, S. and R. Puers, *A review of focused ion beam applications in microsystem technology*. J. Micromech. Microeng., 2001. **11**: p. 287.
18. Luo, Y. and V. Misra, *Large-area long-range ordered anisotropic magnetic nanostructure fabrication by photolithography*. Nanotechnology 2006. **17**: pp. 4909–4911.
19. Steinmann, P. and J.M.R. Weaver, *Fabrication of sub-5 nm gaps between metallic electrodes using conventional lithographic techniques*. J. Vac. Sci. Technol., 2004. **B22**(6): p. 3178.
20. Gazzadi, G.C., et al., *Fabrication of 5 nm gap pillar electrodes by electron-beam Pt deposition*. J. Vac. Sci. Technol., 2005. **B23**(2): p. L1.
21. Hand, A., *Post-Litho Shrink Breaks Advanced Barrier*. Semiconductor International, 2007. **August**.
22. Li, J., et al., *Ion-beam sculpting at nanometre length scales*. Nature, 2001. **412**: p. 166.
23. Lo, C.J., T. Aref, and A. Bezryadin, *Fabrication of symmetric sub-5 nm nanopores using focused ion and electron beams*. Nanotechnology, 2006. **17**: pp. 3264–3267.
24. Stern, M.B., M.W. Geis, and J.E. Curtin, *Nanochannel fabrication for chemical sensors*. J. Vac. Sci. Technol., 1997. **B15**(6): p. 2887.
25. Zeng, H., Zhiliang, W., and A.D. Feinerman, *Fabrication of micro/nano fluidic channels with sacrificial galvanic coupled metals*. Nanotechnology, 2006. **17**: pp. 3183–3188.
26. Blech, V., T. Nobuyuki, and B. Kim, *Nanostenciling through a cm^2-wide silicon membrane*. J. Vac. Sci. Technol., 2006. **B24**(1): p. 55.
27. Luber, S.M., et al., *Nanometre spaced electrodes on a cleaved AlGaAs surface*. Nanotechnology, 2005. **16**: pp. 1182–1185.
28. Austin, M.D., et al., *Fabrication of 5 nm linewidth and 14 nm pitch features by nanoimprint lithography*. Appl. Phys. Lett., 2004. **84**(26): p. 5299.
29. Austin, M.D., et al., *6 nm half-pitch lines and 0.04 μm static random access memory patterns by nanoimprint lithography*. Nanotechnology, 2005. **16**(8): pp. 1058–1061.

30. Hulteen, J. and V.D. RP, *Nanosphere lithography: A material general fabrication process for periodic particle array surfaces*. J. Vac. Sci. Technol., 1995. A13(3): p. 1553.
31. Ng, W.N., et al., *Nanostructuring GaN using microsphere lithography*. J. Vac. Sci. Technol., 2008. **B26**(1): p. 76.
32. Wu, W., et al., *A deep sub-wavelength process for the formation of highly uniform arrays of nanoholes and nanopillars*. Nanotechnology, 2007. **18**: p. 485302.
33. Guo, W., et al., *Near-field laser parallel nanofabrication of arbitrary-shaped patterns*. Appl. Phys. Lett., 2007. **90**: p. 243101.
34. Barton, J.E. and T.W. Odom, *Mass-limited growth in zeptoliter beakers: A general approach for the synthesis of nanocrystal*. Nano Lett., 2004. **4**: pp. 1525–1528.
35. Lin, B.J., *The ending of optical lithography and the prospects of its successors*. Microelectron. Eng., 2006. **83**: pp. 604–613.
36. Vladimirsky, Y., et al., *Demagnification in proximity X-ray lithography and extensibility to 25 nm by optimizing Fresnel diffraction*. J. Phys. D: Appl. Phys., 1999. **32**: pp. L114–L118.
37. Cumming, D.R.S., et al., *Fabrication of 3 nm wires using 100 keV electron beam lithography and poly(methyl methacrylate) resist*. Appl. Phys. Lett., 1996. **68**(3): p. 322.
38. Nagase, T., et al., *Maskless fabrication of nanoelectrode structures with nanogaps by using Ga focused ion beams*. Microelectron. Eng. 2005. **78–79**: pp. 253–259.

Chapter 8
Nanofabrication by Self-Assembly

8.1 Introduction

The nanofabrication technologies described in the previous chapters so far can be characterized as conventional technologies, in the sense that they always involve in one way or another lithographic patterning and pattern transfer. They are still the same as, or not too much different from, the technologies for manufacturing integrated circuits (ICs), which originated nearly 50 years ago in 1961 when the first patent on planar IC was granted, though today's technologies can make pattern structures a thousand times smaller. It is a "top-down" approach that complex structures are built up by patterning layers upon layers from the surface of a planar substrate. The capabilities of top-down nanofabrication technologies have been amply demonstrated in previous chapters. The smallest features which can be made by top-down approach are, however, always limited by the available fabrication tools, being either lithography or pattern transfer. On the other hand, nature has demonstrated its ability to produce extremely complex living organisms by self-organization and self-construction for billions of years, all of which are taking place at molecular level which is an order of magnitude smaller than a nanometer. One surely could ask the question why a desired nanostructure cannot be made by molecular self-assembly, which is a "bottom-up" approach where molecules autonomously "grow" into nanostructures.

Such an attempt had started more than a decade ago, pioneered by, not surprisingly, chemists and biologists [1]. To them, manipulation of molecules is a familiar subject. In some areas, nanosystems constructed by molecular self-assembly have been quite successful. For example, self-assembled monolayers (SAMs) have been widely used in nanoscale patterning. Creation and application of carbon nanotubes (CNTs) is another example of how nanostructures and objects can be grown in a controllable manner. DNA molecular strands have been used as basic scaffolds to construct nanostructures. Controlled aggregation of colloidal particles is being widely utilized in formation of functional patterns and structures. There are other examples which demonstrate that certain two-dimensional (2D) or 3D pattern structures can

Z. Cui, *Nanofabrication*, DOI: 10.1007/978-0-387-75577-9_8,
© Springer Science+Business Media, LLC 2008

be formed by guided self-assembly or self-organization by the interaction of molecular forces or other interparticle forces.

The goals of nanofabrication by self-assembly are twofold: one is to be able to construct structures in true nanoscale, that is, at molecule level, and another is to be able to make nanostructures cheaply. Nanostructuring by molecular self-assembly should be, in principle, low cost. However, there is a long way to go before reaching these goals. The grand challenge is how to control the synthesis and assembly of molecules into molecular assemblies and into desired nanostructures, microstructures, and into macro objects. As Professor George M. Whitesides of Harvard University put it, "Although there are countless examples of self-assembly all around us – from molecular crystals to mammals – the basic rules that govern these assemblies are not understood in useful detail, and self-assembling processes cannot, in general, be designed and carried out "to order". Many of the ideas that are crucial to the development of this area – "molecular shape", the interplay between enthalpy and entropy, the nature of non-covalent forces that connect the particles in self-assembled molecular aggregates – are simply not yet under the control of investigators." [2].

One has to be aware that the current level of nanofabrication by self-assembly is still very primitive. Most of the nanostructures made from molecular self-assembly are array like, which originated from the formation of crystal lattices by molecular attractions. In majority of cases, the self-assembly has to be guided or has to take place on a template, which requires assistance from conventional top-down nanofabrication processes. The combination of top-down and bottom-up approaches, or the so-called "top-down meets bottom-up", has been, and will be in the future, the main feature in the development of self-assembly nanofabrication. To become a mainstream nanofabrication technique, molecular self-assembly has to come out of the realm of chemists and biologists, to be accepted by physicists and engineers. This chapter attempts to examine these self-assembly techniques, not from a chemist's or biologist's point of view but from a nanofabricator's point of view. Their practicality and applicability are analyzed.

8.2 Self-Assembly Processes

Self-assembly has a very broad meaning. Any spontaneous formation of an organization from individual parts without human intervention can be called self-assembly. What the nanofabrication is concerned about is the organization into ordered 2D patterns or 3D structures, ideally with designed shapes and geometries. Self-assembly can take place either at molecular level or meso/macroscopic level. Self-assembly at molecular level involves noncovalent or weak covalent interaction, typically van der Waals, electrostatic, and hydrophobic interactions or interfacial hydrogen bonding. In large-scale, gravitational force, capillary force, or external electromagnetic force may come into

play [3]. Those individual parts have to be mobile in order for the self-assembly process to occur. Therefore, self-assembly process normally takes place in liquid phase or on a smooth surface so that the basic building blocks can move around and interact with each other. The actual formation of an organization or ordered structure is influenced by environment, either the input of external energy or confined geometrical boundaries. The formation of an ordered structure is usually the result of equilibrium state, at which individual parts keep equal distance to each other, not a random aggregation. The self-assembled structures are long-range ordered beyond a few molecules once the interaction reaches equilibration.

Self-assembly has been classified into three categories: chemical, physical, and colloidal self-assemblies [4]. Chemical self-assembly involves molecular-level ordering. It is a chemical synthesis process whereby molecules are combined into new crystal lattice or chain structures by noncovalent bonds. Formation of SAM is a typical example of chemical self-assembly process. Chemically self-assembled structures are easy to prepare and form quickly from solutions of the assembling molecules. They are molecularly ordered and robust under many conditions of use. They represent thermodynamic stability, form spontaneously, and tend to reject defects, and they permit control of film thickness to within approximately 1 nm [5]. Physical self-assembly is the ordering of inorganic atoms during a physical deposition process such as molecular beam epitaxy (MBE) or chemical vapor deposition (CVD). Both of the processes involve reorganization of atoms at a substrate surface to form fine-grained and high-purity thin films. Colloidal self-assembly refers to self-assembly of particles or spheres with diameter from micrometers to nanometers in a liquid suspension. Particles in such a dimensional range have properties neither of individual atoms nor of bulk solid-state materials. As already mentioned earlier, several different forces can cause self-assembly of the particles into ordered 2D or 3D organizations.

Self-assembly processes can also be characterized as either static or dynamic processes, depending on if there are exchanges of energy between the assembly structures and their surrounding environment. Static self-assembly involves systems that are at global or local equilibrium and do not dissipate energy. Formation of the ordered structure may require energy, but once it is formed, it is stable. Most of the self-assembly systems investigated so far are static. In dynamic self-assembly, the interactions responsible for the formation of structures or patterns between components only occur if the system is dissipating energy, such as the self-assembly of living organisms [3]. Understanding and control of dynamic self-assembly system is the greatest challenge the humankind is facing. In general, self-assembly is a complex process. There are different forces in action, often not a single but several in combination. In the example of nanosphere in a suspension liquid, it is not only the capillary forces that can drive the aggregation of the particles, there are electrostatic forces and van der Waals forces which can interrupt the Brownian movement of the particles as well [6].

The above description may seem abstractive. There are many practical examples in nature showing self-assembly at work. Crystallization is the simplest example of molecular self-assembly into long-range ordered lattice structures. In crystals, atoms organize to form ordered solids when attractive and repulsive interatomic forces are properly balanced. A specific lattice structure is highly dependent on the condition of crystallization. Another familiar example is the formation of snowflakes. Water molecules crystallize from liquid phase to solid phase when temperature is below 0°C. The ice crystals are of many different shapes, depending on temperature, pressure, air current, and humidity, as the examples shown in Fig. 8.1. Generally, colder temperatures produce snowflakes with sharper tips on the sides of the crystals and may lead to more branching of the snowflake arms (dendrites). Snowflakes that grow under warmer conditions grow more slowly, resulting in smoother, less intricate shapes.

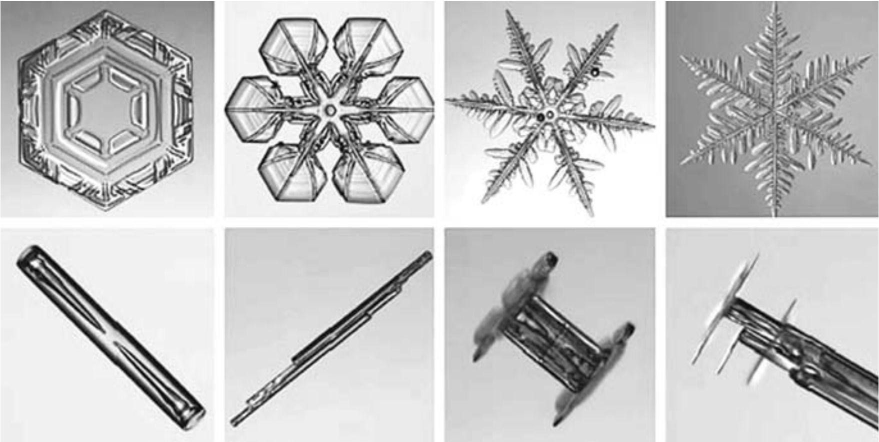


Fig. 8.1 Various shapes of snowflakes from crystallization of water molecules (Reprint from [7] with permission)

Crystallization normally involves formation of covalent bonds between atoms, which requires high-temperature or high-pressure environment. Majority of self-assembly processes that are of interest in nanofabrication are 2D phenomena involving noncovalent binding forces on a surface. Surface tension force is one of such forces. It has long been observed that microsized polystyrene spheres (PSs) in liquid suspension can form 2D ordered aggregation if the thickness of suspension liquid layer is approximately equal to the particle diameter. The process starts with a drop of suspension liquid containing the PS on a glass substrate surface. The drop spontaneously spreads and forms a slightly concave layer. Microscopic observations show that the PSs in this layer are in intensive Brownian motion. Initially they randomly scatter in the liquid layer. Several minutes later one can see that the particle concentration in close

vicinity to the glass surface becomes higher than in the bulk of the aqueous layer due to mutual attraction of the particles. As the layer thickness gradually decreases due to the water evaporation, the particles come closer and closer, and eventually form closely packed hexagonal array when the water level drops below the height of the spheres. Figure 8.2 shows the optical image of orderly packed PSs assembled from its initially random distribution. The ordering is due to surface tension or capillary force of suspension liquid between the particles, which pulls the particles together.

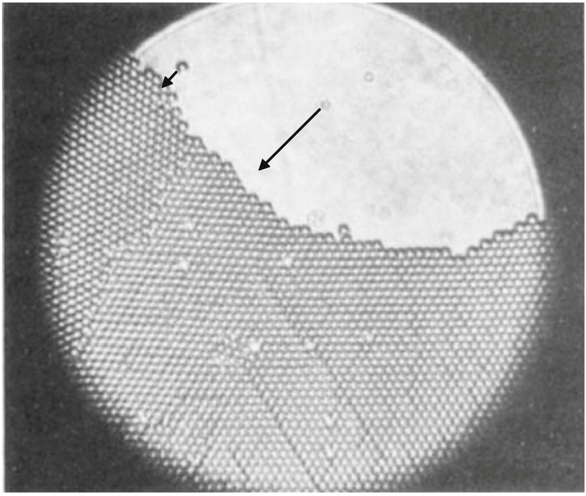


Fig. 8.2 Orderly packed polystyrene spheres formed by capillary force-driven self-assembly (Reprint from [8] with permission)

Surface tension is a much stronger force which can drive the assembly of much bigger objects than the polystyrene particles. It has become an effective technique to assemble micromechanical parts into a 3D microsystem [9]. Figure 8.3 has illustrated how the surface tension of photoresist can pull up a flat polysilicon structure into vertical position. The polysilicon movable structure is initially fabricated by silicon surface micromachining processes, in which silicon dioxide serves as the sacrificial layer. A layer of photoresist is coated on top of the flat polysilicon structure and patterned by optical lithography to fill up only one side of the gap (Fig. 8.3a). Upon heating the photoresist into melting state, the photoresist contracts into a spherical shape because the surface tension always drives a free liquid surface into minimum surface area, so as to minimize the surface energy. With the contraction, the movable flat polysilicon structure is pulled up into a vertical position (Fig. 8.3b). Figure 8.3c is an example of upright microstructure assembled by the surface tension force. Not only the photoresist, many other meltable materials can serve as the self-assembly drivers, such as solders and glass.

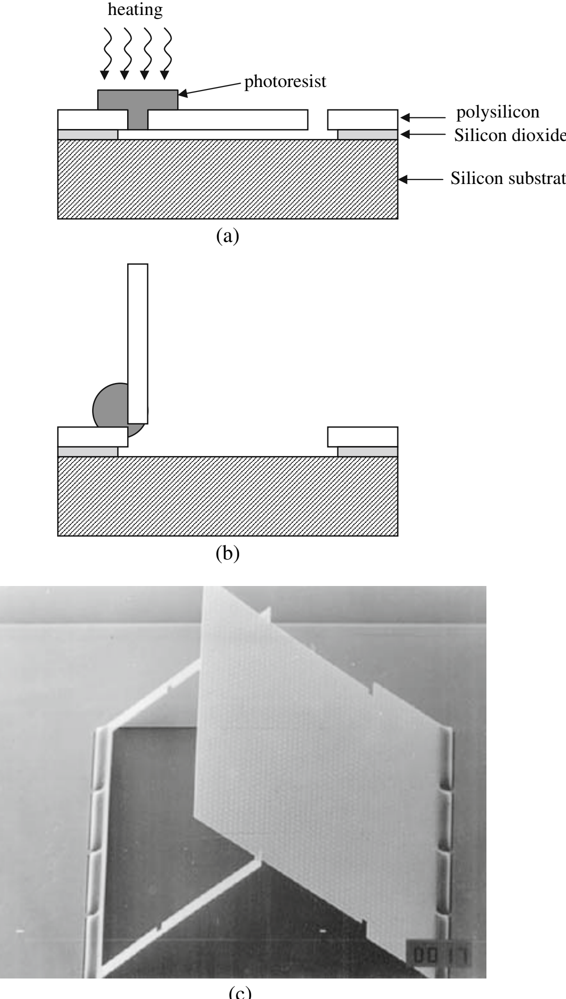


Fig. 8.3 Self-assembly of microstructures by surface tension force. (**a**) Initial flat structure. (**b**) The movable polysilicon structure is pulled up by surface tension of melting photoresist. (**c**) An example of upright microstructure assembled by surface tension force (Reprint from [10] with permission)

At nanoscale, heating can drive randomly distributed molecules into forming nanocrystals, which can also change the surface morphology. The principle has been used to create nanometer-size holes in an initially amorphous silicon film. The fabrication process is illustrated in Fig. 8.4. A high-quality three-layer film stack (20-nm SiO_2/15-nm a-Si/20-nm SiO_2) is deposited on the front surface of a silicon substrate by sputter-coating (Fig. 8.4a). The substrate is then briefly exposed to high temperature (715°C–770°C for 30 s) in a rapid thermal processing chamber. The a-Si is crystallized into a nanocrystalline film, creating nanoscale pores in the film because of the clustering of silicon atoms (Fig. 8.4b). Wet etching of silicon-supporting substrate from back side and removal of the protective SiO_2 produce a porous nanocrystalline silicon membrane with pore size ranging from 5 nm to 25 nm

(Fig. 8.4c). In the nanopore formation process, the thermal annealing temperature is very critical. At 700°C or below, no nanopores are formed. The higher the temperature the bigger the nanopore size becomes. The annealing temperature is therefore a means to tune the pore size. Annealing at 715°C, 729°C, and 753°C have produced average pore sizes of 7.3 nm, 13.9 nm, and 21.3 nm, respectively [11]. Although the nanopores formed in this way are not regular array and not uniform in size, the thin porous membrane can still serve as an effective filter which has been used to separate differently sized or differently charged proteins.

The same principle has been applied to create platinum nanocrystals. It has been found that platinum nanostructures on alumina become highly faceted at temperatures greater than 700°C. This is done by rapidly heating a 2-nm thin platinum film on an alumina substrate. The original thin film becomes clustered with multifaceted sidewalls, which greatly increases the surface area and may have applications in fuel cells [12]. Another similar technique utilizes the mismatch of stress in a thin film system to create cracks. For example, a 5-μm thick

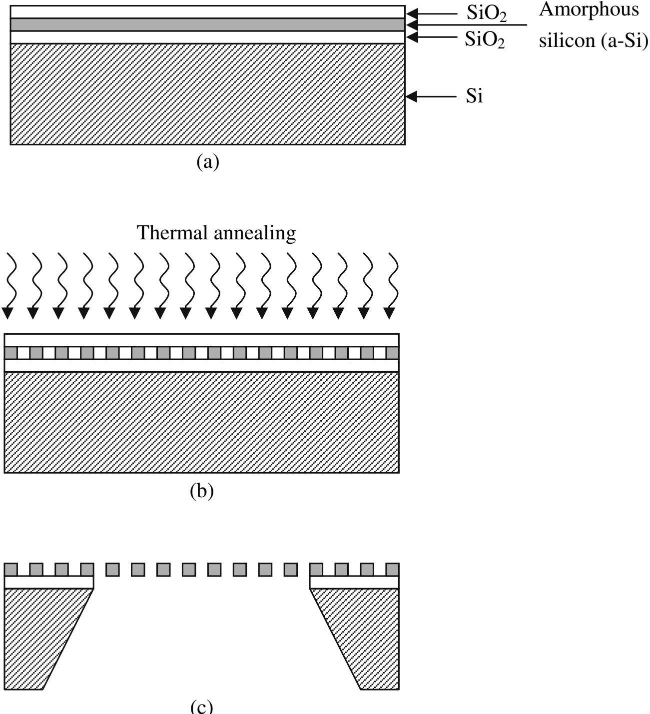


Fig. 8.4 Fabrication of nanosized thin porous membrane. (**a**) Sputter-coating triple layers with thin amorphous silicon (a-Si) sandwiched between SiO_2 layers. (**b**) Nanocrystallization of a-Si upon thermal annealing. (**c**) Removal of backside of silicon-supporting substrate and protective SiO_2 layers to create the porous silicon membrane

SiO_2 film deposited on a silicon wafer is found to have up to 200-MPa compressive stress. A further heating of the film converts part of the compressive stress into tensile stress, causing the film to crack [13]. Though cracks are highly undesirable in a normal thin film application, by careful design of open window patterns in the thin film the cracks can be directed in ordered fashion. Filling up the cracks with electroplated nickel and removing the rest of the thin film, metallic wires with dimension of around 100 nm can be produced. Fractures can also be formed in polymers with surprising regularity. It is found that by prising apart two relatively rigid flat plates which have been glued together with a thin layer of polymer, the polymer layer remaining on each side of the plates are fractured into highly regular grating structures. The grating period is always four times the thickness of polymer film, regardless of its molecular weight and chemical composition [14]. Sub-100 nm grating patterns can be created by simply reducing the film thickness.

Self-assembly at molecular level requires mutual attraction between molecules. This is only possible if there is a chemical affinity existing between the molecules; that is, the dissimilar molecules can pair with each other, or "recognize" each other. The most successful self-assembly by molecular recognition is the formation of SAM. SAMs are ordered molecular assemblies formed by the adsorption of an active surfactant on a solid surface. The order in these 2D systems is produced by a spontaneous chemical synthesis at the interface, as the system approaches equilibrium. SAMs are usually prepared by immersing a substrate in the solution containing a ligand that is reactive toward the surface, or by exposing the substrate to the vapor of the reactive species. The most investigated SAMs are alkanethiolates $CH_3(CH_2)_nS^-$ on Au(111) surface [15], though the organosulfur compounds also absorb very strongly to silver, copper, platinum, mercury, iron, nanosize γ-Fe_2O_3 particles, colloidal gold particles, GaAs, and InP surfaces [16]. Alkanethiols absorb spontaneously on a gold surface from solution and form adsorbed alkanethiolates. The first step in creating the SAM is the absorption process where the sulfur atom chemisorbs to the Au(111). This is possible since sulfur compounds have a strong affinity to transition metal surfaces, such as Au, due to their ability to form multiple bonds with the surface metal clusters. Sulfur atoms bonded to the gold surface bring the alkyl chains into close contact; these contacts freeze out configurational entropy and lead to an ordered structure. The kinetics that drive this ordering are related to chain disorder, the dipole–dipole and van der Waals forces of the chain–chain interaction, and the surface mobility of the chains. The kinetics is faster for longer alkane chains, probably due to the increased van der Waals interactions [16]. Highly ordered SAMs of hexadecanethiolate on gold can be formed in a few minutes by immersing a gold substrate in a solution of hexadecanethiol (HDT) in ethanol (ca. 2 mM). In microcontact printing using a soft stamp, as described in Chapter 5, the transfer of thiol molecules onto a gold surface to become SAM takes only a few seconds [15].

The SAMs of thiolate compounds have many attractive features and offer several advantages over other SAM systems. They provide the highest structural

order by densely packing long alkanethiolate chains on the surface and forming crystalline or quasi-crystalline monolayers. They are easy to prepare, have density of defects low enough to be useful in many applications, good stability under ambient laboratory conditions, and most importantly, are resistant to acid erosion. For example, silver surfaces with octadecanethiolate monolayers could be kept in ambient conditions without tarnishing for many months, and copper surfaces coated with the same monolayer sustain nitric acid attack. Gold surface protected by a monolayer of HDT is 10^6 times slower in cyanide (gold etchant), as has been described in Chapter 6. Such a property makes the SAM film an excellent ultra thin resist or passivation layer for pattern transfer with nanometer dimension. SAMs are also very useful in surface engineering, that is, modification of surface properties such as wetting, adhesion, lubrication, corrosion, nucleation, protein adsorption, and cell attachment.

Once the gold surface is covered with organosulfur SAM which consists of long-chain polymer molecules, other organic molecules can be added to it. So the surface becomes functionalized. The SAM surface can be functionalized in such way that it has receptor-terminated alkanethiolates group which only recognizes certain other molecules, as schematically illustrated in Fig. 8.5, forming a molecular recognition system [17]. Specific recognition systems can be defined by different interfacial binding property of molecules. It is like a "lock and key" system. One receptor only recognizes a specific group of molecules [18], which is a very useful technique to construct sensors [19].

Not only the SAMs, the power of chemical synthesis also provides access to a potentially vast range of functionally and structurally diverse molecular building blocks, which can be linked through different types of relatively weak, noncovalent interactions (predominantly hydrogen bonds and metal–ligand interactions) to yield organized supramolecular architectures with tailor-made properties. Supramolecular engineering has become an active field in bottom-up approach of nanofabrication [20]. Although chemical synthesis is known to be able to form supramolecular crystals or supramolecular compounds in solution, which relies on molecular mobility in liquid phase and chemistry of intermolecular noncovalent bond, nanofabrication is more about constructing structures on a substrate surface. To construct supramolecular structure on a

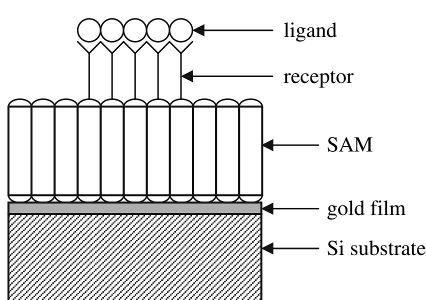


Fig. 8.5 Schematic of SAM-enabled molecular recognition

surface, the influences of the substrate atomic lattice and substrate electronic structure on noncovalent bonds also play critical roles. The parameters crucial for this type of self-assembly are the surface mobility of molecules, their lateral interactions, and their coupling to the surface atomic lattice. Though there are more constraints in assembly of molecules on a surface than in a solution, it provides the possibility to tune the surface property and to produce tailor-made molecular structures. One example is the construction of surface-supported metal–organic coordination networks, resulting from the interaction of organic linker molecules and transition metals at surfaces with the coordination behavior of benzenepolycarboxylic acid species and iron (Fe) atoms on a copper Cu(100) surface [21]. It involves sequential deposition of polyfunctional organic molecules and Fe atoms on an atomically clean Cu(100) substrate. By tuning the organic molecule-to-Fe concentration ratio and altering thermal annealing parameters, distinct molecular network architectures may be produced, which are stabilized by relatively strong lateral metal–organic coordination bonds [22].

8.3 Guided Self-Assembly

If thin films can be counted as 1D nanostructure, then pure self-assembly is quite useful in producing atomic or molecular layers of thin films, such as by MBE, CVD, or SAMs. Other than that, pure self-assembly or bottom-up approach cannot generate many useful nanostrutures, apart from forming lattice-like periodical arrays, as shown in Fig. 8.6, which may only have practical applications as photonic crystals [23, 24] or in magnetic storage [25, 26]. Self-assemblies have to combine with conventional top-down micro or nanofabrication technologies in order to form 2D or quasi 3D arbitrary pattern structures for wider applications; in other words, the self-assembly process has to be guided.

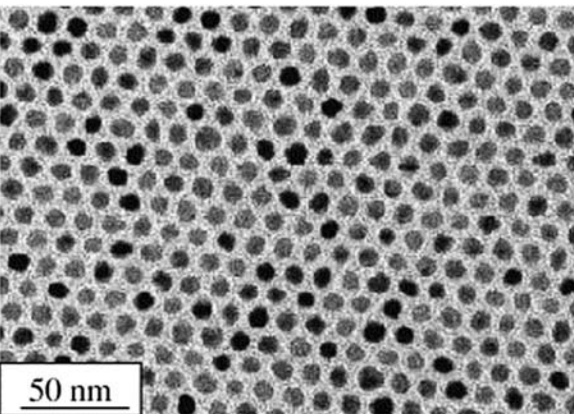


Fig. 8.6 Self-assembled layer of 6-nm diameter Fe–Pt magnetic nanoparticles (Reprint from [26] with permission)

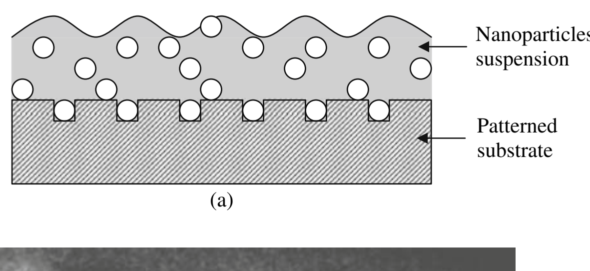
As described in the previous section, the lattice-like structure produced by self-assembly is strongly influenced by external energy and environmental conditions. These conditions and energy inputs can act a guide to direct the assembly developing toward a designed path. Here conventional fabrication can provide a template with designed pattern layout and self-assemblies are confined to the defined pattern areas on the template. The fabricated patterns can be topographical structures which create physical boundaries to prevent self-assembly from spreading all over the surface, or areas with specific surface energy which either facilitate or prohibit the self-assembly.

8.3.1 Surface Topography

Patterned substrate with surface topography can be quite effective to trap nanoparticles. Figure 8.7a illustrates schematically how nanoparticles are guided into pre-patterned areas when the substrate is immersed into a nanoparticle suspension. This is only a simplistic illustration. The real process of trapping the nanoparticles is much complicated. These nanoparticles are not physically falling into the traps, because the gravitation force is too weak for these nano-sized particles, but driven by capillary force of the suspension liquid. It only happens when the suspension liquid drop dries up by evaporation, as in the example of Fig. 8.2, and the surface topography pins the retreating liquid surface as it shrinks, bringing the nanoparticles into the trap. Figure 8.7b shows the scanning electron microscope (SEM) image of self-assembled nanoparticles in a groove. Once they aggregate in the trap, electrostatic force of the suspension in ionic solution or interparticle van der Waals force will further secure the assembly and their attachment to the substrate before the drying is completed. Instead of drying a liquid drop, nanoparticle suspension can also be dip-coated onto a topographical substrate; that is, the substrate is dipped into the suspension and then slowly pulled out, leaving a thin layer of liquid which contains nanoparticles on the surface. Uniform assembly of nanoparticles into patterned grooves can be achieved by balancing the pulling speed and evaporation rate of the thin liquid film [27].

8.3.2 Surface Wetting

It is apparent that the drying of liquid suspension is an inherent part of the nanoparticle self-assembly process. It is the capillary force during drying that drives the colloidal nanoparticles into aggregation. The above examples have shown that the capillary force can be used to intentionally guide the self-assembly process, provided that a suitable topographical template is used. On the other hand, the substrate surface can be chemically modified to stop or reduce the action of capillary force in some areas while it is active in other areas, so that colloidal particles can be selectively assembled. This can be achieved by



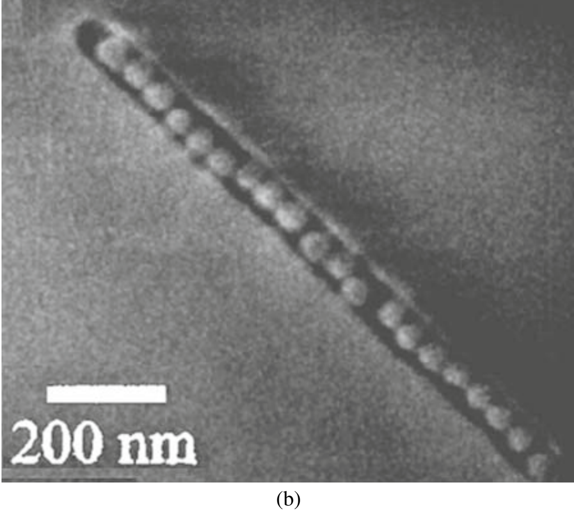


Fig. 8.7 Guided self-assembly of nanoparticles by surface topography. (**a**) Schematic illustration. (**b**) SEM image of assembled nanoparticles in a patterned groove (Reprint from [6] with permission)

changing the surface hydrophobicity. In this way, nanoparticles in a colloidal suspension will only assemble onto the hydrophilic surface area, not stick to the hydrophobic surface area. There are different ways to modify the hydrophobicity of a substrate surface. However, the key parameter is the contrast between hydrophobicity and hydrophilicity of adjacent surfaces, which determines the contrast of self-assembled patterns. The high hydrophobicity contrast can be achieved by patterned SAM. One of such SAM patterning processes is as follows: a silicon substrate is first coated with a 100-nm SiO_2 layer followed by a 12-nm TiO_2 layer on top; the desired geometrical features are patterned by standard photolithography of photoresist; the pattern is then locally etched to remove the TiO_2 by reactive-ion etching (RIE); the SiO_2/TiO_2 pattern is then immersed in an aqueous solution of 0.5-mM ammonium dodecyl phosphate ($CH_3(CH_2)11PO_4(NH_4)_2$) ($DDPO_4$) for 48 h and rinsed with high-purity water. This produces a well-defined hydrophobic self-assembled alkane phosphate monolayer (SAM) specifically on the TiO_2 area while leaving the SiO_2 areas entirely uncovered. By this process, a hydrophobicity contrast pattern is created

with the hydrophobic SAM covering the TiO_2 regions (advancing contact angle = 110°) and the hydrophilic SiO_2 areas (advancing contact angle <6°) [28].

A much more straightforward chemical patterning of surface is by microcontact printing (μCP). Microcontact printing technique has been discussed in length in Chapter 5 (Section 5.6.2). The technique relies on the rapid formation of SAMs in the areas that are in contact with a rubber stamp. This is only possible with alkanethiols on gold. Alkanethiols solution serves as the ink. The stamps are made of polydimethylsiloxane (PDMS) elastomer casting against photolithographically patterned masters, and they form a conformal contact with the substrate. Once in contact, a monolayer of alkanethiol molecules is diffusively transported onto the surface. Different SAMs have different hydrophobicities, which can be selectively coated onto a surface. Here is an example of patterning procedure: a PDMS stamp inked with 1-mM solution of $HS(CH_2)_{15}COOH$ in ethanol is brought into contact with gold surface for 1 min, and rinsed with ethanol to produce pattern areas covered with a carboxylic acid-terminated SAM on the substrate. The substrates are subsequently immersed in a 1-mM solution of $HS(CH_2)_{17}CH_3$ in ethanol for 1 h to cover the remaining surface with a methyl-terminated SAM. The homogeneous SAM of $HS(CH_2)_{15}COOH$ has an advancing contact angle of 31° with water at pH 2; the homogeneous SAM of $HS(CH_2)_{17}CH_3$ has an advancing contact angle of 103° with water at pH 2. This hydrophobicity contrast is sufficient to produce selective self-assembly of PSs ranging from 200 nm to 10 μm. Ordered arrays of particles have been created spontaneously by evaporative deposition of colloidal suspension drops whose diameters are large compared to the length scale of the underlying wetting pattern [29].

The fundamental limit of μCP is resolution. As has been discussed in Chapter 5, soft stamp may deform when pressed down to make contact with substrate surface, and more importantly the ink can diffuse. As soon as the stamps are in contact with the surface, the alkanethiols start to diffuse rapidly across the gold surface, thereby broadening SAM-protected features. Furthermore, ink can also transfer through vapor phase from those areas where the stamps are not in contact with the surface [30]. On the other hand, electron beam (e-beam) lithography is known to have very high resolution and the versatility to pattern any shapes. In the same way as energetic electrons break the molecular chain of polymer resists, electrons can also modify the molecular structure of SAMs. Electron-induced reactions of aliphatic and aromatic SAMs are similar to those in positive-tone and negative-tone resists: destruction of SAMs of hydrocarbons (alkanethiols, alkylsilanes) by electron irradiation makes this group of SAMs act like positive-tone resists. For SAMs of aromatic biphenyls, the electrons induce a C–H cleavage and the process is then followed by cross-linking between neighboring phenyl units. During this process, the cross-linked molecules maintain their preferred orientation, so that the electrons generate a molecularly well-ordered and cross-linked layer that can be used as a negative-tone resist. Electrons can also generate chemical nanostructures in aromatic SAMs. Upon irradiation, the terminal

nitrogroups of nitrobiphenylthiol SAMs are converted into amino groups, while the underlying aromatic substrate is dehydrogenated and cross-linked. The chemical conversion of the nitrogroup can be used to create surface binding sites for specific immobilization of molecules. The electron modification of SAMs is unfortunately very inefficient. Electron exposure dose up to 50,000 $\mu C\ cm^{-2}$ is needed in order to induce noticeable changes in the SAM structure [31].

8.3.3 Electrostatic Force

Compared with chemical, surface tension, and Van der Waals interactions, which is of short range, electrostatic interaction is much long ranged. Two oppositely charged particles suspended in a fluid will attract to each other. Similarly, charged particle in a suspension fluid will attract to electrodes which is oppositely charged. This is the basis for electrostatic force-assisted self-assembly. Substrate with uniform surface charge can be made by immersing clean gold surfaces in ethanol solution of alkanethiols for 10 min. The surface can be either positively charged or negatively charged, depending on the type of thiol groups in the solution. Patterned surface charge can be made by μCP with different thiol compounds as the ink [32], or contact electrification [33]. To maintain a stable and controllable surface charge, it is best to use patterned electrodes with externally applied field. Polystyrene spheres, which are the particles mostly used for controlled self-assembly, can also be made to carry electric charge. Carboxyl-terminated PS can be made to carry negative charge and amidine-terminated PS can be made to carry positive charge. Similarly, gold nanoparticles can also be made to carry surface charge by coating with polymers (hexadimethrine bromide for negatively charged gold, sodium polyvinylsulfonate for positively charged gold) [32]. For electric force-assisted self-assembly, the oppositely charged nanoparticles are attracted to the charged substrate surface in the same way as electrophoretic deposition. However, they may not assemble into orderly packed colloidal crystalline structure, because these particles repel each other due to the same type of electric charge they carry. With an externally controllable electrical field, an orderly packing can be formed, for example, by raising the applied voltage [34]. The exact nature of the assembly is not clear, but is believed to be due to some electrohydrodynamic [35] or electroosmotic [36] effect near the electrode surface at high field. The behavior of charged nanoparticles can also be altered by UV illumination of a transparent electrode surface, which also helps the formation of colloidal crystal structures. Figure 8.8 is a SEM image of assembled PSs by UV-assisted electrophoretic deposition. The assembly is not perfect as there are only localized 2D crystalline lattice (ordered) structures within the patterned area. Each of the crystalline domain is around 10–20 μm. However, the overall goal of selective assembly of PS onto pre-patterned area has been achieved, as seen from the inset image.

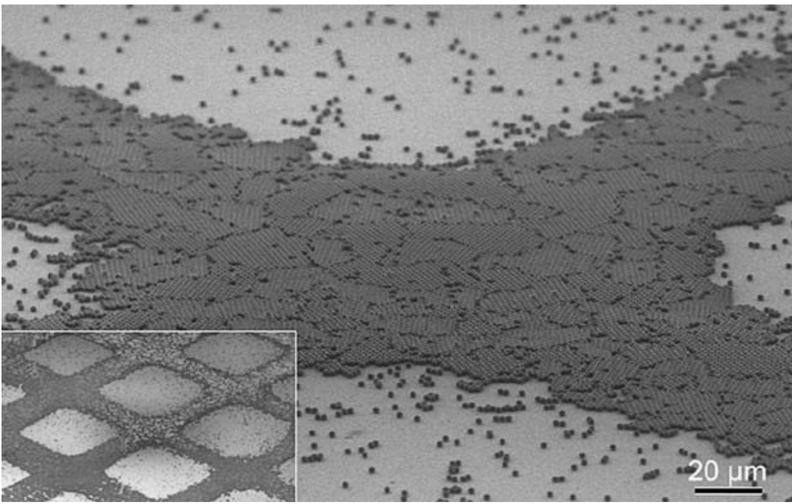


Fig. 8.8 Two-dimensional pattern formed by self-assembled polystyrene spheres, the inset image shows the overall appearance of the pattern (~600 μm) (Reprint from [37] with permission)

If not using an external electrical field, then capillary force can be utilized to help the assembly of charged colloidal particles onto a oppositely charged surface. To allow the capillary force to play a role, the assembly proceeds in two stages: the first stage is the attraction of charged particle to the charged substrate surface in a liquid suspension; the second stage is to take out the substrate from the suspension and let it dry. The charged substrates are prepared by coating 2 nm of Ti and 20 nm of Au on glass slides. PDMS stamps with various relief structures are "inked" with 10 mM of $HS(CH_2)_{15}CO_2H$ in ethanol and brought into conformal contact with the gold surface for 10 s, and then rinsed with ethanol; the noncontact areas are derivatized with a 10-mM solution of either $HS(CH_2)_{11}N(CH_3)_3^+Cl^-$ or $HS(CH_2)_{15}CH_3$ in ethanol by immersion for 1 min, resulting in patterned substrates bearing anionic and cationic regions or anionic and neutral regions. The charged substrate is then immersed in a colloidal suspension upside down to avoid the nonspecific sedimentation induced by gravity. After 2 min, the substrate is taken out, and the excess suspension is carefully rinsed from the surface with water. The second stage is capillary-driven assembly through water evaporation. During water evaporation stage, the attached colloids start to rearrange when the thickness of the water layer becomes comparable to the particle size. The particles within each charged region on the substrate move toward the center of the region, pushed by capillary force of low-level water as illustrated in Fig. 8.9, and form a dense cluster [38]. If the diameter of charged region is comparable to the particle size, single particle array can be assembled. In this case, the capillary force can fix the particle right to the center of charged region.

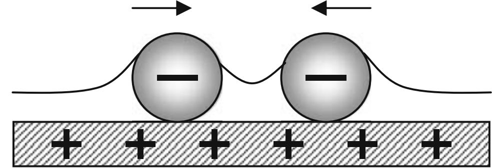


Fig. 8.9 Illustration of capillary force pushing charged particles together

8.3.4 Magnetic Force

Similar to electrostatic force, the magnetic force is also long-ranged. Magnetic particles can be assembled onto a magnetized surface, which is the basis for magnetically assisted self-assembly process. The advantage of magnetically assisted self-assembly is that information can be recorded remotely and stored in the same manner as that of magnetic disk data storage for a long period of time. The assembly processes guided by an appropriately magnetized surface can be accomplished without simultaneously applying and holding potential at numerous locations and without the need for complex electrical connections. In particular, unlike electric fields, magnetic field has practically no effect on most molecules or chemical processes. In addition, while materials may become polarized by electric fields, magnetic fields have greater selectivity and can permeate uninhibited through most mediums, particularly through liquids and gases.

Magnetically assisted self-assembly mainly refers to assembly of superparamagentic colloidal particles whose magnetization is linearly related to the applied magnetic field; that is, no hysteresis is exhibited in this type of materials. Superparamagnetic colloidal particles remain stably suspended in the absence of magnetic fields because they are unmagnetized, and thus experience virtually no magnetic interactions with other particles. Particles made of ferromagnetic materials becomes superparamagnetic when their size is smaller than the superparamagnetic critical size, which is roughly 10 nm in diameter [39]. Superparamagnetic colloidal particles are typically 100 nm–10 μm in diameter, composed of superparamagnetic nanoparticles encapsulated or dispersed inside silica or various polymers, such as polystyrene or polyvinyldibenzene. Increasing the magnetic content of the particles, by encapsulating more magnetic nanoparticles inside each polymer bead, results in only an increased magnetic response without altering its superparamagnetic properties.

The trapping or assembly of magnetic particles is the combined action of external and local magnetic fields, as illustrated in Fig. 8.10. Superparamagentic colloidal particles suspended in a colloidal liquid are magnetized by an external magnetic field. These particles become magnetic dipoles aligning to the external magnetic field lines. They are also subject to the local magnetic force generated by patterned magnetic elements on the template, which either attracts or repels these magnetic dipoles. These magnetic particles are then selectively assembled into the template patterns, or trapped by the local magnets. Since each magnetic element on the template can be individually magnetized, just as the data recording on a magnetic disk, the assembling of magnetic particles can be programmed to form nonperiodic patterns [40].

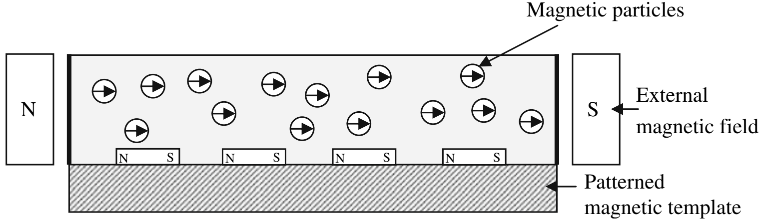


Fig. 8.10 Schematic of magnetically assisted self-assembly of magnetic particles

There are some issues related to magnetic assembly of particles. First, in order to achieve precise assembly of magnetic particles, that is, controlled number of particles per trapping site, the size and separation of magnetic elements on the template has to match the particle size. For example, the trapping probability of Ni magnetic nanowires was studied on magnetic templates with different gaps of nanomagnet elements and different distances between the two nanomagnets. It was found that Ni magnetic nanowires with a length of 10 μm have a significantly higher probability of getting trapped in the nanomagnet gaps with a width of 10 μm compared to the nanomagnet gaps with a width of 5 μm. The matched dimension of the magnetic nanowire and the gap of a nanomagnet pair leads to lower magnetostatic energy compared to an unmatched dimension between the nanowire and the gap [41]. Alternatively, a physical trap (micro well) can be patterned above the magnetic element, as schematically shown in Fig. 8.11, to ensure that only one particle is trapped. Another issue is the size of magnetic particles. For magnetic particles bigger than 10 μm, gravitational force is no longer negligible. The movement of a particle is not only controlled by magnetic force but also by hydrodynamic drag force and gravitational force. For magnetic particles smaller than 100 nm, the magnetic force becomes too weak because of low content of ferromagnetic material in each particle. In this case, other short-range forces, such as surface tension and van der Waals forces, become comparable to the magnetic force. The precise control of magnetic particle movement by local magnetic elements becomes difficult [42]. Therefore, for the best manipulation of magnetic particles, the appropriate particle size is in the range of 100 nm–10 μm.

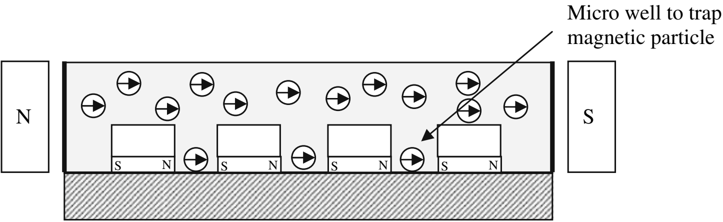


Fig. 8.11 Schematic of patterned physical trap to assist assembly of magnetic particles

Most of the magnetically assisted self-assembly of magnetized particles are in the head-to-tail configuration and tend to form chain-like clusters, because of the dipolar nature of magnetized particles. It is possible to assemble lattice-like 2D structures by combining both magnetic force and surface tension force. As shown in Fig. 8.2, nanoparticles can self-assemble into superlattice structure by a drying process; that is, evaporation of a water drop containing nanoparticles results in close packing of the nanoparticles. If the nanoparticles are superparamagnetic and external magnetic field is introduced to assist the assembly process, a superlattice structure consisting of magnetic particles can be formed. It is found that the drying dispersion alone can only promote short-range ordered monolayer packing of the magnetic particles. With the assistance of a weak external magnetic field for a very short time (<2 min) during the initial stage of dry dispersion assembly process, highly ordered and defect-free superlattices with a typical dimension of up to several micrometers in the x–y plane and a thickness of a few particle monolayers have been formed [43]. A continuous application of external magnetic field results in large 3D superlattice structures. Although the magnetic-field-induced dipole–dipole attraction between the individual superparamagnetic particles is responsible for the alignment, other attractive forces, such as van der Waals attraction and capillary force, also play crucial parts in the growth of large, highly ordered 3D superlattices.

8.4 Building Blocks of Future Nanosystems

The self-assembly processes introduced so far in this chapter rely solely on intermolecular or interparticle force to connect the basic units together. The structures by this type of action are lattice-like ordered structures. To turn this type of nanostructures into functional devices and systems, top-down conventional nanofabrication technologies have to be employed, as described in the previous section, to "guide" the self-assembly processes. Nature is far more intelligent in construction of complex living organisms from a few basic building blocks. For example, 99% of human biomass is made of just four elements, hydrogen, oxygen, carbon, and nitrogen [4]. Instead of using the brute force of lithography and patterning processes of top-down nanofabrication, learning from nature and harnessing the growth process of living organisms would be the ultimate goal of nanofabrication for constructing functional nanosystems. Though there is a long way to go before reaching the goal, work is already underway to use some basic building blocks provided by nature to assist in the construction of functional nanosystems. The most noticeable example introduced in this section is the use of DNA molecules as scaffolds from which other nanostructures are constructed. Carbon nanotube is another example. Though CNTs are man-made nanostructures, as basic building blocks CNTs are found in many different nanodevices and applications. Block copolymers represent a group of polymer materials which offers endless possibilities to assemble into

different shapes and forms. Porous alumina is another example of naturally occurring ordered nanostructure, providing a nice template from which other nanostructures and systems can be built. With the advances in chemical and biological research, more of such building blocks will find their way into man-made nanosystems.

8.4.1 DNA Scaffold

Deoxyribonucleic acid (DNA) is a long chain of polymer made from repeating units called nucleotides. Biologically, DNA contains the genetic instructions used in the development and functioning of all known living organisms. DNA becomes an attractive building block in nanofabrication because of its unique physical and chemical properties. Although DNA consists of only four bases which are pairwise ordered and only 2–3 nm wide, the intertwined chain, in the form of double-helix structure, contains millions of molecules and is up to centimeters long. DNA chain is normally in tangle or coil form. It can be stretched to full length by either a receding liquid meniscus [44] or optical tweezers [45]. Figure 8.12 shows schematically how the DNA strand can be stretched by a receding meniscus. The phenomenon has been called molecular combing, which is a simple technique to untangle the DNA strand and align it to a surface [46]. Other stretching techniques include electrophoretic, dielectrophoretic, or hydrodynamic stretching [47]. DNA molecular chain is remarkably elastic. The force required to stretch DNA beyond its natural length is approximately 50 pN [48]. The DNA double helix is stabilized by hydrogen bonds between the bases attached to the two strands, while hydrogen bonds is also one of the intermolecular forces described previously for some of the molecular self-assembly processes. Therefore, DNA chains can be "tricked" to attach to variety of other organic and inorganic molecules to form other functional nanostructures. In this way, DNA itself becomes a template or a scaffold from which other nanosystems are built.

One of the early applications of DNA-based self-assembly was DNA-assisted self-assembly of gold nanoparticles [49]. Gold nanoparticles treated with both oligonucleotides and alkane thiols become functionalized. In a linking DNA solution these surface-functionalized gold nanoparticles can recognize each other and become aggregated into clusters. Unlike the conventional thiol-gold absorption, the DNA-based assembly is reversible. Upon heating to above 42°C the assembly can be reversed. Once the gold nanoparticles are assembled with DNA molecules, the new nanostructures can be easily characterized by other physical means such as transmission electron microscope or atomic force microscope. It also provides a new way of detecting DNA molecules. For example, gold nanoparticles assembled into DNA molecules can change the surface conductance, which can be used to enhance the sensitivity of DNA or protein detection [50].

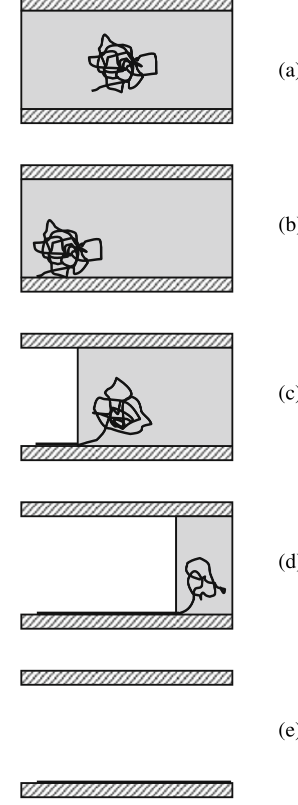


Fig. 8.12 Stretching DNA strand by a receding meniscus: (**a**) Initially coiled DNA in liquid. (**b**) Graft one end of the strand onto the surface. (**c**) and (**d**) Receding air–liquid interface (meniscus) stretches the strand. (**e**) stretched DNA is left dry in air

A stretched DNA strand is a natural nanowire with diameter of ~2 nm, which is unattainable with conventional nanofabrication techniques. Yet it is long enough to connect with conventionally fabricated nanostructures. To use a stretched DNA as part of a nanoelectronic device, two issues have to be solved: one is the precise positioning of DNA and another is to make the DNA conductive. Anchoring of DNA onto pre-patterned electrodes can be done in conjunction with hydrodynamic stretching process. The anchoring is possible because DNA tethered at one end can be stretched by the shear force generated by the momentum gradient between an applied fluid flow and the DNA chain. The end-to-end distance of stretched DNA depends mainly on the velocity of flow and viscosity. DNA is fixed to a gold electrode surface by Au–thiol coupling mechanism if the ends of DNA are ligated to two thiol-capped oligonucleotides. After one end of DNA was immobilized to one electrode, a flow at velocity of 50–100 μm s^{-1} was applied. The other end of stretched DNA molecules became bound to the other electrode by Au–thiol coupling [47]. This is the unique feature of DNA-based self-assembly. It has the property of molecular recognition that only the DNA nanowires capped with oligonucleotides can be positioned and immobilized onto the electrode surface.

Although DNA has the appropriate molecular-recognition and mechanical properties, the poor electrical characteristics prevent its direct use in electrical circuits. Metallization of DNA nanowires can be realized, again, by utilizing the molecular recognition feature of DNA. Ag ions are found to be able to bind to DNA by Ag–Na ion exchange [51]. This process is highly selective and only applies to the DNA template. The silver ion-exchanged DNA is then reduced to form nanometer-sized metallic silver aggregates bound to the DNA skeleton. These silver aggregates are subsequently further "developed", much as in the standard photographic procedure, using an acidic solution of hydroquinone containing silver ions under low light conditions. The silver aggregates on the DNA act as catalysts and significantly accelerate the process. The final metallized DNA nanowire measures about 100-nm in diameter and 12-μm long between two electrodes. The metallized nanowire could not be made thinner because the silver aggregates attached to the DNA skeleton have grain size of about 30–50 nm [51]. This is the drawback of catalytic deposition of metals. Thinner DNA-based nanowires can be made by direct physical evaporation of metallic atoms onto the DNA skeleton. Sub-10 nm thin nanowires have been made by vacuum deposition of a thin layer of gold onto a DNA template [52].

The DNA can also provide a scaffold to localize and align CNTs. A key issue in the use of CNTs in nanofabrication is the need for their controlled placement at well-defined positions on surfaces. Previous efforts to control localization of nanotubes or nanorods have included the use of microfluidics, applied electric field, or patterning using side edges of multilayer films. DNA is an attractive candidate template, since it possesses a small diameter ($\sim$2 nm), controllable length that can exceed 10 μm, and the ability to use molecular recognition in forming specific base pairs. As aforementioned, DNA is relatively easy to locate and immobilize onto an electrode surface. Once in position, the DNA can then be used as a template to further assemble CNTs. The key process is to immerse the DNA-templated substrate into 1-pyrenemethylamine hydrochloride (PMA) solution [53]. With the PMA treatment, 63% of single-walled CNTs in the CNT suspension were able to anchor along the DNA strand. Otherwise, only $\sim$10% of the nanotubes were aligned. The DNA-assisted localization of CNTs can be employed to make CNT-based field-effect transistors. Figure 8.13 shows one section of DNA where single-walled CNT (Fig. 8.13a) and a bunch of single-walled CNTs (Fig. 8.13b) are anchored to the DNA strand which has been metallized on both sides of the DNA section by catalytic deposition [54]. The thin long DNA strands can also be woven into mesh-like structures by molecular recognition process. The DNA mesh is then used as a mask and subsequent deposition of gold transfers the DNA mesh into nanoscale grooves in the gold film [55]. The process has been given the name of "molecular lithography", because the mask pattern is not made from polymer resists but from self-assembly of DNA molecules. Another name is DNA shadow nanolithography, in which stretched DNA works as a shadow mask to create trenches of less than 30-nm wide on a silicon substrate [56].

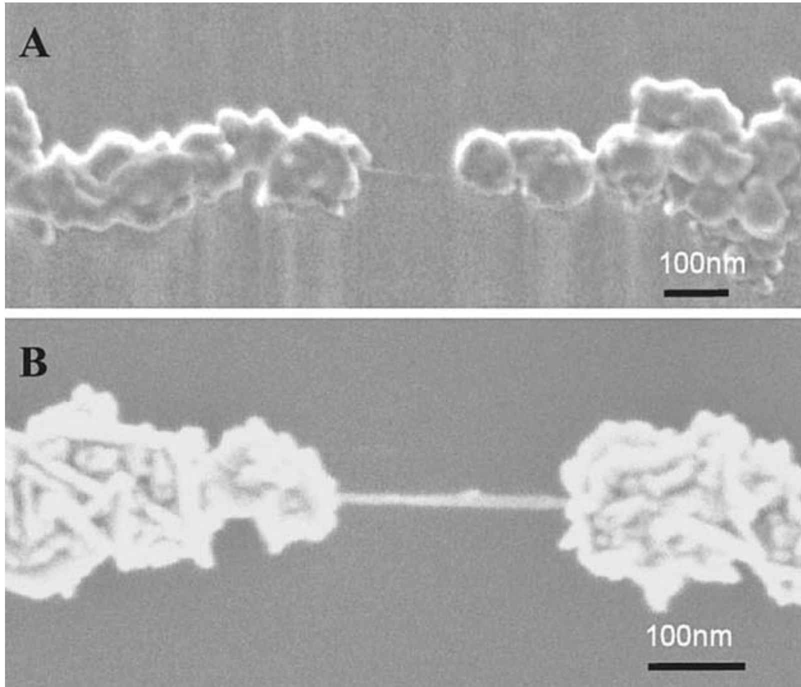


Fig. 8.13 SEM images of metallized DNA with the middle section having CNT anchored (**a**) single-walled CNT, (**b**) a rope of single-walled CNTs (Reprint from [54] with permission)

Apart from its biological properties, DNA has many other surprising properties which extend its application far beyond merely a genetic code carrier. Its thin and long strands can be twisted, rolled, as well as stretched. Polyhedral DNA nanostructures with the connectivity of a cube of truncated and regular octahedral and tetrahedral have been made [57]. Molecular motors and nanomachines can be made from DNA, as a recent review article revealed [58]. The construction and operation of DNA nanomachines are primarily chemical synthesis processes, which is beyond the scope of this book.

8.4.2 Carbon Nanotubes

Like the DNA strands, CNTs are thin and long nanostructures. Single-walled carbon nanotubes (SWNT) have diameter in the order of 1 nm. Even multiwalled carbon nanotubes (MWNT) are only a few tens of nanometer thick. They are, however, of the length up to a few millimeters. Since reported in 1991 [59], CNT has become the single mostly investigated object in nanotechnology. The tremendous interests in CNTs attribute to their amazing properties. Carbon nanotubes are stronger than high-carbon steel, conduct electrical current

density 1000 times greater than silver and copper, and have thermal conductivity nearly 20 times higher than copper [60]. Unlike the DNA, CNTs are man-made nanostructures. Several techniques have been developed to manufacture CNTs, including arc discharge, laser ablation, and CVD [61]. CNTs can now be produced in large quantity commercially. Manufacturing of CNTs is not the type of nanofabrication covered by this book. However, to turn CNTs into useful and functional nanodevices, there are many nanofabrication-related issues.

One of the key issues is how to place a CNT to where it is needed. Chemically synthesized CNTs come in free forms. They are not like DNAs which have molecular recognition property. To connect a CNT to electrodes, the commonly used method is to let the electrodes find the CNT, not the CNT find the electrodes, that is, dispersing CNT suspension solution over the surface of substrate with pre-patterned electrodes and hoping one of the CNTs will land between a pair of electrodes [62]. The process is quite random and the probability of a CNT landing between electrodes is low ($<10\%$), because the CNT has to be in the right place with the right length in order to bridge the gap of a pair of electrodes. There is no guarantee that the CNT landed on the electrodes is the one with right property, since the property of chemically synthesized CNTs is quite dispersed. Compounding to the difficulty is that CNTs are quite easy to rope together into bundles of 10–1000 CNTs by van der Waals force. A long period of agitation in ultrasonic batch is necessary to sufficiently disperse the CNTs in suspension solution. A slightly better approach is to locate a CNT first under scanning electron microscope (SEM) and then fabricate the electrodes attaching to the CNT. This can be done with focused ion beam (FIB)-assisted deposition of Pt [63]. Fortunately, a dual-beam FIB system is equipped with SEM function, which enables the observation and in situ metal deposition to be performed simultaneously. Figure 8.14 is the SEM picture of a MWNT connected to a pair of electrodes which are made by FIB-assisted Pt deposition [64]. In principle, it is also possible to make the contact electrodes by e-beam lithography. This process involves observation of dispersed CNTs on a pre-marked substrate. Once the desired CNT is found and its exact location is determined with the help of alignment marks on the substrate, e-beam exposure of resist and subsequent lift-off of metal should be able to place the patterned electrodes in the CNT.

The above methods suffer a common drawback, for example, low efficiency. It is a tedious process to connect one CNT, which is impossible to scale up for a large number of CNTs. Though DNA can help to align CNTs, as described in the previous section, stretching and hybridization of DNA to electrodes is still a manual process. It would be ideal if the CNTs can self-assemble to electrodes. This is possible by functionalizing the surface with certain SAMs. It has been found that CNTs adhere to -NH_2-terminated SiO_2 surface but not to -CH_3-terminated SiO_2 surface [65]. This has led to the possibility of selectively placing CNTs onto a substrate surface. The -CH_3-terminated surface can be prepared by covering SiO_2 surface with a trimethylsilyl (TMS) monolayer.

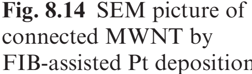
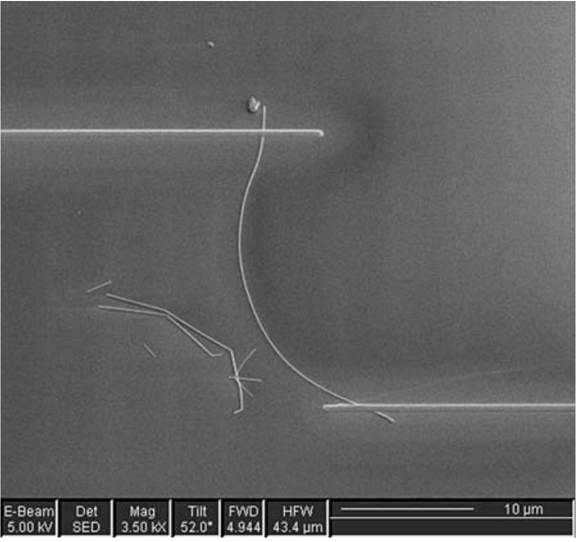


Fig. 8.14 SEM picture of connected MWNT by FIB-assisted Pt deposition



The monolayer is then patterned by AFM or e-beam to expose the underlying SiO_2 which is filled by aminopropyltriethoxysilane (APTS) to create the -NH_2-terminated sites. Individual SWNTs are deposited on the chemically patterned surfaces by soaking the substrate in CNT suspension for 10 min, and then rinsed with methanol immediately after being taken out of the CNT suspension and blown dry with nitrogen. Figure 8.15 shows an anchored SWNT between the

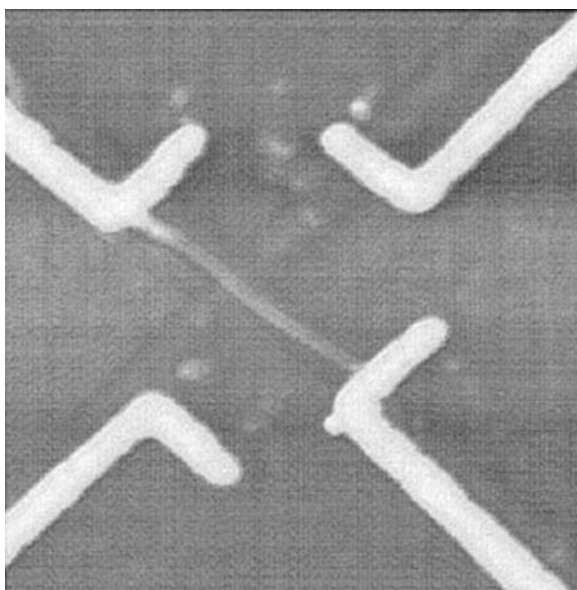


Fig. 8.15 Selective positioning of SWNT between electrodes by chemical patterning of surface (Reprint from [65] with permission)

patterned electrodes by this method. Alternatively, the surface can be functionalized by APTS only and then cover the functionalized surface with PMMA resist. By e-beam lithography of PMMA, certain areas of the surface can be opened to allow CNTs to anchor to. This also enables selective positioning of CNTs [66].

Instead of randomly locating a CNT to electrodes, CNTs can be synthesized onto the electrodes. An array of electrical contacts is fabricated first. The substrate is then placed in a CVD chamber to grow CNTs. The electrical contacts are covered with catalysts from which CNTs are synthesized. CNTs have been successfully synthesized from islands of catalysts with 90% of them as individual SWNTs with diameters ranging from 1 nm to 3 nm [67]. These SWNTs are straight and some of them are long enough to bridge the 10-μm gap between the catalyst islands. However, there is no control on the orientation and density of CNT growth. The distributions of synthesized SWNTs are quite random, though they all originated from these catalyst islands. The random orientation of CNT growth can be improved by applying electrical field between the adjacent electrodes. Highly aligned suspended SWNTs can be grown under electric fields of ~0.5–2 V μm^{-1} [68]. Figure 8.16 shows a SEM image of controlled growth of SWNTs under applied electrical field. Though the density of SWNTS is still not controllable, the probability of SWNTs bridging the gap of electrodes is greatly improved. The main constraint in this process is the compatibility of patterned electrodes with the CVD growth conditions

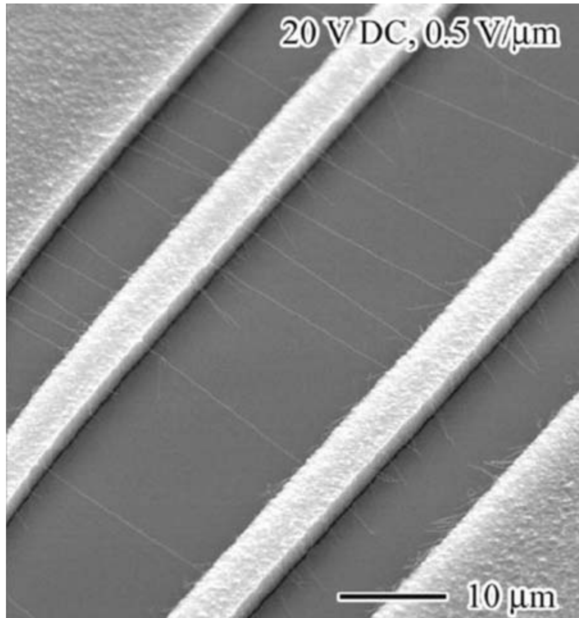


Fig. 8.16 Oriented synthesis of SWNTs with applied electrical field (Reprint from [68] with permission)

[69]. In the above example, the CVD growth of CNTS was conducted in the flow of methane gas at 1000°C temperature. This certainly is not a desirable condition if the CNTs are to be grown onto a substrate with CMOS circuitry. Recently a low-temperature CVD process has been developed. SWNTs can be grown at 350°C temperature [70], which opens the possibility of integrating CNTs with complementary metal-oxide semiconductor (CMOS)-based ICs.

As CVD synthesis of CNTs only takes place at the location of catalysts, it is possible to precisely control the location of CNT growth by lithographically patterning the catalysts. Nickel as the catalyst material can be patterned by e-beam lithography into array of dots. Vertically aligned CNTs are grown on each of the Ni dot by plasma-enhanced CVD (PECVD) process at 700°C substrate temperature at ~150 nm min^{-1} growth rate [71]. If the Ni dot is sufficiently small (~100 nm) single CNT can be grown on each dot [72]. Small Ni dots can also be created by high-temperature (~700°C) annealing of Ni thin film, which causes the film to break up into nanoparticles, though the dots formed in this way are no longer as regular array. Figure 8.17 shows a single CNT grown inside a gated cavity, which serves as a field electron emitter [73]. Controlled vertically aligned growth of CNTs has provided a promising application in IC manufacturing to replace copper as interconnects through via holes. As has been introduced in Chapter 6 of this book (Section 6.2.4), current interconnects through via holes in ICs are accomplished by the so-called damascene process whereby copper is filled into deep via holes to connect different layers of circuits. CNTs, as aforementioned, have much better electrical conductivity than copper. By vertically growing a bundle of CNTs in a via hole, it can potentially replace copper as interconnects in ICs [73]. A long-term goal is to produce CNT-based transistors to replace current CMOS-based ICs. CNT plays a key role in many other applications where its physical, chemical, and biological properties are of main interest rather than the nanofabrication aspect of it.

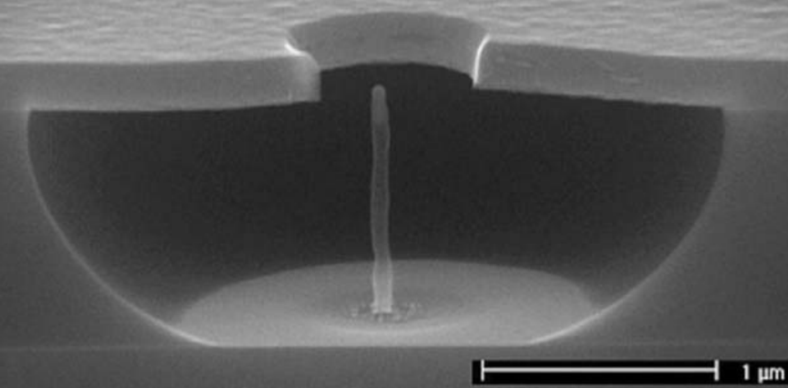


Fig. 8.17 Single CNT grown inside a gated silicon cavity as a field electron emitter (Reprint from [73] with permission)

8.4.3 Block Copolymers

Nanofabrication is concerned with making shapes and structures in nanometer scale. In addition to top-down patterning techniques, a series of bottom-up self-assembly methods have been introduced in previous sections. Some common features exist in these methods: individual units such as molecules or colloidal nanoparticles are mobile; lattice-like nanostructures are self-assembled from the dispersed individual units by noncovalent forces. Nanostructures formed from block copolymers, however, represent a different category of self-assembly process. The structuring process takes place within bulk or thin film of copolymer materials.

Block copolymers consist of chemically dissimilar polymer chains covalently linked to form a single molecule [74]. Under certain conditions, the dissimilar blocks tend to segregate into different spatial domains, called microphase separation, similar to oil–water separation. On the other hand, the spatial separation is constrained by the chemical connectivity of the blocks. As a result of these competing effects, self-organized periodic structures are created on a nanometer scale. Formation of any particular microdomain structure depends on relative volume ratio between blocks and chain architecture as well as the persistence lengths of the respective blocks. In the simplest case of noncrystalline flexible coil AB diblock copolymers, the composition of the AB diblock (i.e., the volume fraction of block A) controls the geometry of the microdomain structure. Table 8.1 gives some possible spatial microdomain structures due to different volume fractions of A block in the diblock copolymer. More complex spatial microdomain structures can be formed with triblock copolymers. The

Table 8.1 Possible spatial microdomain structures for diblock copolymer with different volume fractions of one of the blocks (Reprint from [75] with permission)

Nature of patterns	Spheres (SPH) (3D)	Cylinders (CYL) (2D)	Double gyroid (DG) (3D)	Double diamond (DD) (3D)	Lamellae (LAM) (1D)
Space group	$Im\bar{3}m$	$p6mm$	$Ia\bar{3}d$	$Pn\bar{3}m$	pm
Blue domains: A block					
Volume fraction of A block	0-21%	21-33%	33-37%		37-50%

nanoscale spatial structures that can be formed from block copolymers are almost endless. The key is how to control the formation of a preferable spatial microdomain structure.

Block copolymers, if let alone, will form some sort of ordered structures due to the aforementioned competing effect and microphase separation. However, the microdomains composed of the different blocks and having sizes of several tens of nanometers are formed randomly and grow as a polygranular texture, with periodic ordering maintained only over distances of several tens of lattice constant (i.e., a typical grain size of only 1–2 μm). To achieve long-range ordered microdomain separation, both internal and external influences to the formation process are necessary. The internal control includes the choice of molecular weight, composition, and chain architecture. Table 8.1 has demonstrated how the microdomain structures vary simply due to the content of one of the blocks. The external control is via mechanical flow field [76], electric field [77], or by the influence of surface interactions [78], chemically patterned surface [79], or temperature gradient to create a preferred orientation or alignment of microdomains [75].

Thermal annealing the block copolymer to above its glass transition temperature is the most commonly used external influence. At high temperature, the mobility of block polymer chains is enhanced, which facilitates the interaction between different blocks. However, uniform heating does not produce fully ordered structures. A zone heating which creates a directional temperature gradient can help the block copolymers to develop in ordered state.

The control of mechanical flow field is to create a directional compression and stretching of melting block copolymers, which helps to develop directional alignment of microdomains. The shear gradient developed as molten block copolymer is squeezed through an orifice into a heated glass capillary can result in an axially well-ordered cylindrical microdomain structure with near "single crystal" texture [80].

Block copolymers comprised of blocks with different dielectric constants have been found to orient under an applied electric field, where the employed field strength ranges from 10^6 V m^{-1} to 10^8 V m^{-1}. Microdomain orientation using an electric filed is an alternative approach in order to circumvent limitations in the flow field control due to high viscosities typical for high-molecular-weight copolymers [81].

Patterned substrate surface with topographic structures can be used as a template to direct the self-assembly of block copolymers [82]. Figure 8.18 is the self-assembly of PS-PFS (polystyrene-polyferrocenyldimethylsilane) block copolymer inside patterned grooves. The long vertical-walled grooves of 80-nm depth were made using e-beam lithography and plasma etching into silicon dioxide film. The 1.5 wt.% PS-PFS block copolymer in toluene solution was spin-coated onto the grooved substrate and then annealed at 140°C for 48 h to obtain a monolayer of spherical PFS domains in a PS matrix within the substrate grooves [83]. The surface can also be chemically modified to direct the assembly of block copolymers into predefined patterns [84].

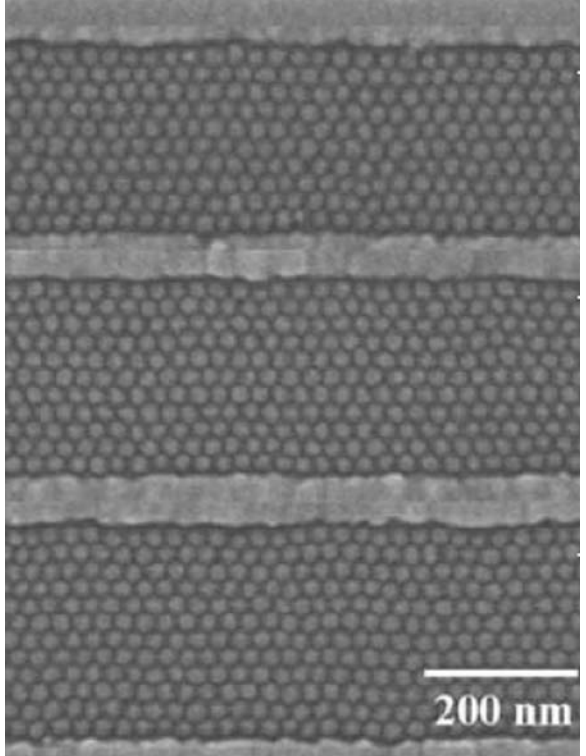


Fig. 8.18 Spherical domains assembled from PS–PFS block copolymer inside patterned SiO_2 grooves (Reprint from [82] with permission)

The periodical nature of self-assembled block copolymers can find many applications, such as serving as photonic crystals, nanostructured networks and membranes, medium for high-density data storage, and templates to assemble nanoparticles and lithographic masks [75]. Using spin-coated block copolymer thin films of well-ordered spherical or cylindrical microdomains as the etching mask, dense, periodic arrays of holes and troughs have been fabricated in silicon, silicon nitride, and germanium. For example, hexagonally ordered polyisoprene (PI) spheres or parallel cylinders of polybutadiene (PB) were formed in a polystyrene (PS) matrix. The PI spheres or PB cylinders were then degraded and removed to produce a PS mask for pattern transfer by fluorine-based RIE. Domains approximately 20-nm wide, 20-nm deep, and spaced 40-nm apart, equivalent to density of 3×10^{12} holes cm^{-2}, were obtained on a 3″ wafer [85]. Another example is the self-assembly of PS:PMMA (polystyrene:polymethylmethacrylate) diblock copolymer as an etching mask [86]. The PS:PMMA copolymer from a dilute solution (1 wt. %) in toluene is spin-coated onto a silicon wafer. Annealing the diblock copolymer film at 180°C for 1 h facilitates microphase separation and hence formation of PMMA hexagonal cylinder array in the PS matrix. The PMMA cylindrical domains are then removed by using a blanket UV (wavelength of 254 nm) exposure for 5 min

and followed by an aqueous development in acetic acid. Hexagonal pore lattice with average diameter of 20 nm and separation of 40 nm is produced. The porous PS film can be used as etching mask to transfer into a thin layer of SiO_2, and then the SiO_2 pattern serves as etching mask to further transfer into silicon by RIE.

The aligned block polymers not only can serve as an equivalent polymer resist structures for subsequent pattern transfer, they can also act as carriers to load metallic materials to turn into metallic nanostructures. It is found that polystyreneblock-poly(2-vinylpyridine) (PS-b-P2VP) block polymer can absorb metal ions in an anionic metallic solution [87]. The reason is that the P2VP blocks in the PS-b-P2VP copolymer become protonated in an acidic environment (1% HF or HCl), carrying positive charges. The anions (negatively charged metal ions) are then attracted to the cationic P2VP blocks and become embedded in the blocks. After loading the metal ions, the P2VP block polymer can be eliminated by oxygen plasma etching, leaving the metallic nanostructures in the shape of what it was as the P2VP blocks on a substrate. Gold nanowires have been made by this process. The microphase separation of P2VP blocks from PS-b-P2VP block copolymers is formed by spin-coating the PS-b-P2VP onto a silicon substrate, followed by annealing at 230°C for 24 h. The silicon substrate has been pre-patterned by e-beam lithography and dry etching to form nanoscale trenches about 35-nm deep. The P2VP blocks are aligned in the trenches as nanowires after the annealing. Immersion of P2VP-patterned silicon substrate in 10-mM HAuCl4 + 0.9% HF solution for 5–10 min then generates the $[AuCl_4]^-$-loaded P2VP nanowires. An oxygen plasma treatment afterwards turns the P2VP nanowires into gold nanowires [87]. Through various physical and chemical manipulations, block copolymer lithography can be an alternative nanostructuring technique to make large area of polymer patterns at sub-100 nm scale without conventional electron or photon exposure of polymer resists [88].

8.4.4 Porous Alumina

As has been stated many times in previous chapters, high-density and high-aspect-ratio nanostructures are the most difficult to fabricate by any conventional top-down nanofabrication techniques. One approach to overcoming the difficulty is to utilize naturally occurring nanostructures as hosts for nanofabrication. DNA molecule is one of such examples which has aspect ratio of many thousands to one. Researchers have also made use of cicada wings as nanoimprint stamps to replicate high-density nanometer holes [89]. Another very neat technique for producing high-density and high-aspect-ratio nanostructures is to anodize aluminum.

Anodization is an electrochemical process whereby certain metals (aluminum, niobium, tantalum, titanium, tungsten, zirconium) can be oxidized electrochemically. During anodization, the metal workpiece is the anode in an

electrolyte bath. The cathode is any electrical conductor that does not react with the electrolyte. Once connected to a power supply, electrons attracted to the anode allow ions at the metal surface to react with water molecules in the electrolyte to form an oxide layer on the metal and hydrogen gas is released from the cathode side. Normally the electrolyte is selected in which the oxide is insoluble so that a barrier oxide is formed, which can serve as protection coating from corrosion or an electrical insulation layer. If the electrolyte is a diluted acid, the oxide layer will be partially dissolved and the oxidation surface becomes porous. This is how porous alumina is produced.

It has been known for a long time that anodization of aluminum in acidic electrolyte can produce porous aluminum oxide (Al_2O_3, alumina) surface finish. It is the primary technique to produce durable coloring of aluminum surface in architectural applications. Porous alumina thin membranes also serve as filters, with the trade name ANOPOR^{TM}, having pore size from 200 nm down to 20 nm and membrane thickness of 60 μm [90]. Formation sequence of porous alumina is schematically illustrated in Fig. 8.19. When the current starts to pass through the aluminum anode, a barrier layer of nonporous alumina begins to form (step-I). During anodization of aluminum in an acidic electrolyte, there are two competing processes taking place: one is the oxidation to form alumina layer and another is the dissolution of alumina in acid. If there is no dissolution of alumina, the oxidation process will eventually stop because the surface becomes insulating as the oxide layer becomes thicker. As this layer grows thicker, the electrical field across the oxide layer increases due to the increase of resistance. The high electrical field causes small fractures on the top of the oxide layer (step-II). Nanopores are formed initially from these fractures in the alumina surface, where acid penetrates and dissolves alumina and some cracks develop further into large pores (step-III). Because the starting points of alumina dissolution are random, the nanopores initially formed are also random. However, as the oxidation and acid dissolution progress, a balance between oxidation and dissolution is reached and ordered hexagonal structures start to emerge (step-IV). The formation of porous surface also decreases the resistance, which produces a voltage decrease when working at constant current density. Finally, a constant voltage is maintained due to the balance of oxidation–dissolution at the bottom of the pores.

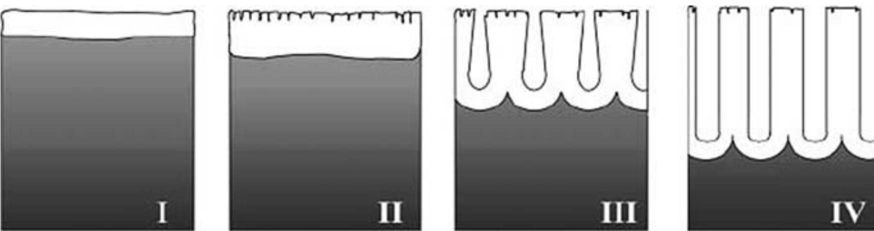


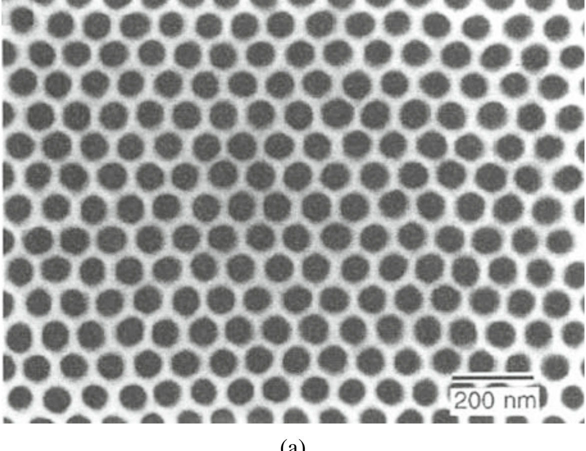
Fig. 8.19 Sequence of porous alumina formation (Reprint from [91] with permission)

Figure 8.20a shows highly ordered hexagonal holes formed in alumina after 160 h anodization of pure aluminum in 0.3-M oxalic acid and 40 V at 0°C temperature [92]. Figure 8.20b is the cross-section view of the porous alumina, showing very dense and high-aspect-ratio honeycomb structures. High voltage and high electrolyte temperature, which result in high anodization current density, will cause the formation of nanopores faster. The growth rate of porous alumina as high as 10 μm min^{-1} was reported at voltage of 195 V and current density of 150 mA cm^{-2} [93]. However, the pore size also grows bigger at higher current density. Another way to speed up the nanopore formation process while maintaining small pore size is to anodize in two steps: first anodization is conducted at the top 1-μm thickness of aluminum. The top alumina is then removed by immersing in 4% (vol.)-CrO_3 + 10% (vol.)H_3PO_4 for 12 h at 25°C. While the nanopores in the initial anodization are random, some regularity of pore structure starts to emerge at the interface between the alumina and aluminum. The regular surface texture is then used as the starting point for second anodization. During the second anodization ordered hexagonal nanopores are being formed with growth rate of 1–2 μm min^{-1} [94]. The pore structures can be further adjusted by pore widening or pore shrinking. For example, pore widening can be performed in phosphoric acid at 35°C and pores shrinking can be accomplished by converting alumina to aluminum oxy-hydroxide (Al(O)OH) in boiling water. Since the nanopores initiate from surface textures, the surface can be deliberately textured by a conventional patterning technique, for example, by nanoimprint. Nanopores will be formed with controlled regularity and density [95]. However, the anodization condition has to be tuned to match the lattice constant of surface pattern. Otherwise, no regular nanopores will be formed.

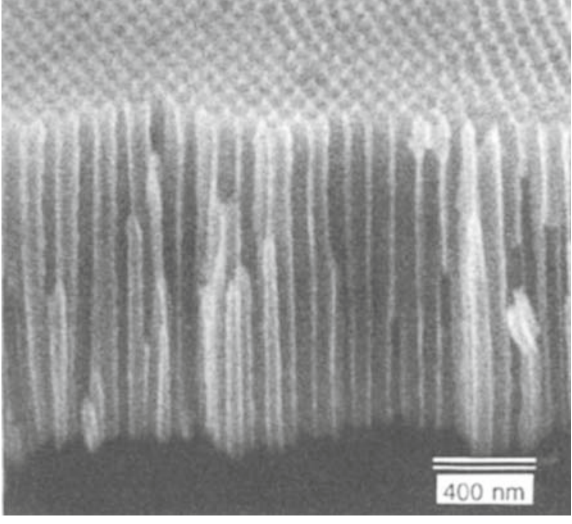
Although ordered porous alumina can be directly used as nanopore-sized filters or as photonic crystals [95], in most cases they are used as templates from which other nanostructures are created. For example, nanowires or nanorodes can be produced by filling up the nanopores with metals [96, 97]. The porous alumina templates have also been used to assemble nanoparticles, whereby nanoparticles are packed into the nanopores by electrophoretic force. Figure 8.21a shows regular nanopores created by two-step anodization of aluminum film and Fig. 8.21b shows nanoparticles packed into the nanopores by electrophoretic force.

Since porous alumina can be formed in aluminum films, the nanopore template can be created by first depositing an aluminum film onto a semiconductor substrate material such as silicon and then turning the aluminum film into porous alumina by anodization. Magnetic nanodots arrays have been made on silicon surface by the porous alumina template [99]. In this process, the porous alumina template is rather like a resist mold. After deposition of magnetic thin film to fill up the nanopores, the alumina template

Fig. 8.20 Hexagonal nanopore array formed by aluminum anodazation. (**a**) *Top* view. (**b**) *Cross-sectional* view (Reprint from [92] with permission)



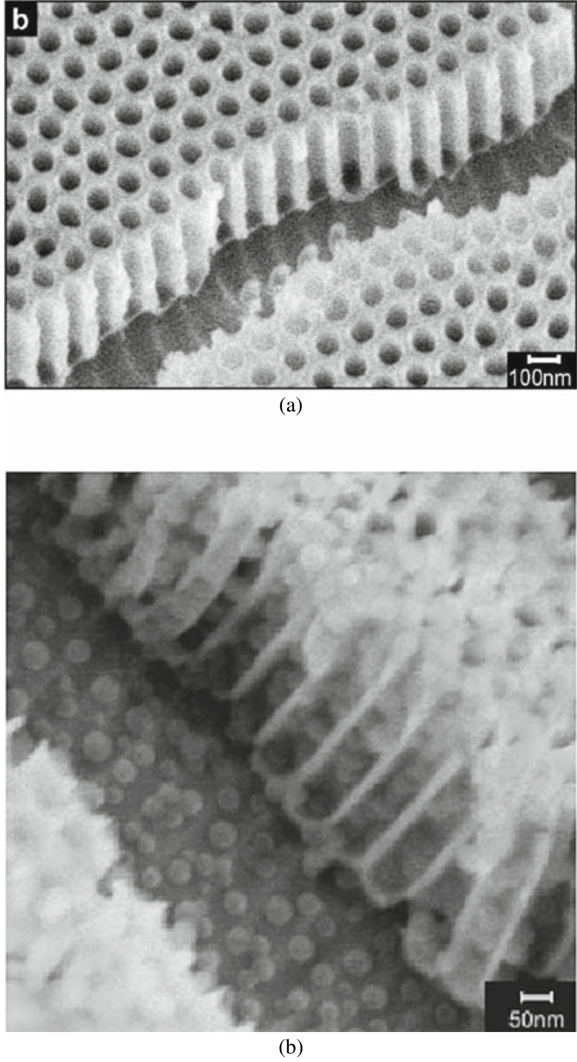
(a)



(b)

is dissolved in either 10 wt.% NaOH solution at room temperature for 2 min or in a mixture of 6.0 wt.% phosphoric acid + 1.8 wt.% chromic acid at 60°C for 15 min [99], just like a lift-off process, leaving the magnetic dots on the silicon substrate.

One must be aware that porous alumina is an insulator. As shown in Fig. 8.19, each nanopore is surrounded by Al_2O_3. Figure 8.22 is the SEM image of a cross-section of porous alumina. It can be seen that the nanopores are terminated by a layer of alumina. If the porous alumina is made on a conducting substrate surface, any metals deposited into the nanopores will



(a)

(b)

Fig. 8.21 (**a**) Cross-section of ordered nanopores in alumina. (**b**) Packed nanoparticles in the nanopores by electrophoretic force (Reprint from [98] with permission)

not be in contact with the substrate surface, unless the barrier layer is removed prior to the metal deposition. If the metal filling is done by electroplating, the barrier layer must be removed in order to expose the underlying conducting surface. The barrier layer can be dissolved in 5% phosphoric acid. However, the acid dissolution also reduces the thickness of porous alumina layer and widens the pore size.

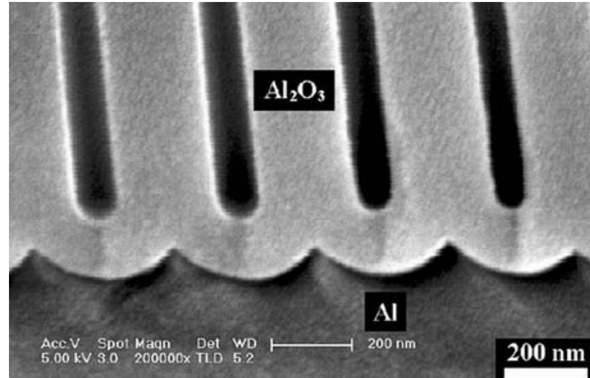


Fig. 8.22 Cross-section of nanopores in alumina (Reprint from [93] with permission)

References

1. Whitesides, G.M., J.P. Mathias, and C.T. Seto, *Molecular self-assembly and nanochemistry: a chemical strategy for the synthesis of nanostructures*. Science, 1991. **254**: p. 1312.
2. Whitesides, G.M., *Self Assembly and Nanotechnology*, in *Fourth Foresight Conference on Molecular Nanotechnology*. 1995.
3. Whitesides, G.M. and B. Grzybowski, *Self-assembly at all scales*. Science, 2002. **295**: p. 2418.
4. Zhirnov, V. and D.J.C. Herr, *New frontiers: self-assembly and nanoelectronics*. IEEE Computer, 2001. **34**(1): pp. 34–43.
5. Andres, R., et al., *The design, fabrication and electronic properties of self-assembled molecular nanostructures*, in *The Handbook of Nanostructured Materials and Nanotechnology*, H.S. Nalwa, Editor. 1998, Academic Press: San Diego.
6. Liddle, J.A., Y. Cui, and P. Alivisatos, *Lithographically directed self-assembly of nanostructures*. J. Vac. Sci. Technol., 2004. **B22**(6): p. 3409.
7. Libbrecht, K., *The enigmatic snowflakes*. Physics World, 2008. **21**(1): p. 19.
8. Denkov, N.D., et al., *Mechanism of formation of two-dimensional crystals from latex particles on substrates*. Langmuir, 1992. **8**: pp. 3183–3190.
9. Syms, R.R.A., et al., *Surface tension-powered self-assembly of microstructures – the state of the art*. J. Microelectromech. Syst., 2003. **12**: pp. 387–417.
10. Syms, R.R.A., *Surface tension powered self-assembly of 3-D microoptomechanical structures*. J. Microelectromech. Syst., 1999. **8**(4): pp. 448–455.
11. Striemer, C.C., et al., *Charge- and size-based separation of macromolecules using ultrathin silicon membranes*. Nature, 2007. **445**: p. 05532.
12. Aurongzeb, D., *Low index faceting of Pt nanostructures on oxide surface with potential application for fuel cells*. J. Appl. Phys., 2007. **102**: p. 064302.
13. Alaca, B.E., H. Sehitoglu, and T. Saif, *Guided self-assembly of metallic nanowires and channels*. Appl. Phys. Lett., 2004. **84**(23): p. 4669.
14. Pease, L.F., et al., *Self-formation of sub-60-nm half-pitch gratings with large areas through fracturing*. Nat. Nanotechnol., 2007. **2**(9): pp. 545–548.
15. Xia, Y. and G.M. Whitesides, *Soft lithography*. Angew. Chem. Int. Ed., 1998. **37**: pp. 550–575.
16. Ulman, A., *Formation and structure of self-assembled monolayers*. Chem. Rev., 1996. **96**: pp. 1533–1554.

17. Motesharei, K. and D.C. Myles, *Molecular recognition on functionalized self-assembled monolayers of alkanethiols on gold*. J. Am. Chem. Soc., 1998. **120**: pp. 7328–7336.
18. Jeoung, E., J.B. Carroll, and V.M. Rotello, *Surface modification via 'lock and key' specific self-assembly of polyhedral oligomeric silsequioxane (POSS) derivatives to modified gold surfaces*. Chem. Comm., 2002. **14**: pp. 1510–1511.
19. Wenz, G. and P. Liepold, *Self-assembly of biotin and thio-functionalized carboxymethyl celluloses on gold and molecular recognition of streptavidin detected by surface plasmon resonance*. Cellulose, 2007. **14**(2): pp. 89–98.
20. Barth, J.V., G. Costantini, and K. Kern, *Engineering atomic and molecular nanostructures at surfaces*. Nature, 2005. **437**: p. 671.
21. Lin, N., et al., *Supramolecular engineering of metal-organic networks at surfaces*, in *Proceedings of the 12th International Conferences Scanning Tunneling Microscopy/Spectroscopy and Related Techniques*, P.M. Koenraad and M. Kemerink, Editors. 2003. p. 144.
22. Langner, A., et al., *Self-recognition and self-selection in multicomponent supramolecular coordination networks on surfaces*. Proceedings of the National Academy of Sciences of the United States of America, 2007. **104**(46): pp. 17927–17930.
23. Vlasov, Y.A., et al., *On-chip natural assembly of silicon photonic bandgap crystals*. Nature, 2001. **414**: p. 289.
24. Velikov, K.P., A. Moroz, and A. van Blaaderen, *Photonic crystals of core-shell colloidal particles*. Appl. Phys. Lett., 2002. **80**(1): p. 49.
25. Sun, S., et al., *Monodisperse FePt nanoparticles and ferromagnetic FePt nanocrystal superlattices*. Science, 2000. **287**: pp. 1989–1992.
26. Anders, S., et al., *Lithography and self-assembly for nanometer scale magnetism*. Microelectronic Engineering, 2002. **61–62**: pp. 569–575.
27. Juillerat, F., et al., *Fabrication of large-area ordered arrays of nanoparticles on patterned substrates*. Nanotechnology, 2005. **16**: pp. 1311–1316.
28. Fan, F. and K.J. Stebe, *Assembly of colloidal particles by evaporation on surfaces with patterned hydrophobicity*. Langmuir, 2004. **20**: pp. 3062–3067.
29. Huwiler, C., et al., *Self-assembly of functionalized spherical nanoparticles on chemically patterned microstructures*. Nanotechnology, 2005. **16**: pp. 3045–3052.
30. Huck, W.T.S., *Self-assembly meets nanofabrication: recent developments in microcontact printing and dip-pen nanolithography*. Angew. Chem. Int. Ed., 2007. **46**: pp. 2754–2757.
31. Geyer, W., et al., *Electron induced chemical nanolithography with self-assembled monolayers*. J. Vac. Sci. Technol., 2001. **B19**(6): p. 2732.
32. Tien, J., A. Terfort, and G.M. Whitesides, *Microfabrication through electrostatic self-assembly*. Langmuir, 1997. **13**: pp. 5349–5355.
33. McCarty, L.S., A. Winkleman, and G.M. Whitesides, *Electrostatic self-assembly of polystyrene microspheres by using chemically directed contact electrification*. Angew. Chem. Int. Ed., 2007. **46**: pp. 206–209.
34. Trau, M., D.A. Saville, and I.A. Aksay, *Field-induced layering of colloidal crystals*. Science, 1996. **272**: pp. 706–709.
35. Trau, M., D.A. Saville, and I.A. Aksay, *Assembly of colloidal crystals at electrode interfaces*. Langmuir, 1997. **13**: pp. 6375–6381.
36. Solomentsev, Y., M. BoÈhmer, and J.L. Anderson, *Clustering and pattern formation during electrophoretic deposition: a hydrodynamic model*. Langmuir, 1997. **13**: pp. 6058–6068.
37. Hayward, R.C., D.A. Saville, and I.A. Aksay, *Electrophoretic assembly of colloidal crystals with optically tunable micropatterns*. Nature, 2000. **404**: p. 56.
38. Aizenberg, J., P.V. Braun, and P. Wiltzius, *Patterned colloidal deposition controlled by electrostatic and capillary forces*. Phys. Rev. Lett., 2000. **84**(13): p. 2997.
39. Lee J., Isobe T., and M. Senna, *Preparation of ultrafine Fe3O4 particles by precipitation in the presence of PVA at high pH*. J. Colloid and Interface Sci., 1996. **177**: pp. 490–494.
40. Yellen, B.B. and G. Friedman, *Programmable assembly of colloidal particles using magnetic micro-well templates*. Langmuir, 2004. **20**: pp. 2553–2559.

41. Liu, M., et al., *Self-assembled magnetic nanowire arrays*. Appl. Phys. Lett., 2007. **90**: p.103105.
42. Yellen, B.B., G. Friedman, and A. Feinerman, *Analysis of interactions of nanoparticles with magnetic templates*. J. Appl. Phys., 2002. **91**(10): pp. 8552–8554.
43. Ahniyaz, A., Y. Sakamoto, and L. Bergstrom, *Magnetic field-induced assembly of oriented superlattices from maghemite nanocubes*. Proceedings of the National Academy of Sciences of the United States of America, 2007. **104**(45): pp. 17570–17574.
44. Bensimon, D., et al., *Stretching DNA with a receding meniscus: experiments and models*. Phys. Rev. Lett., 1995. **74**(23): p. 4754.
45. Wang, M.D., et al., *Stretching DNA with optical tweezers*. Biophys. J. 1997. **72**: pp. 1335–1346.
46. Bensimon, A., et al., *Alignment and sensitive detection of DNA by a moving interface*. Science, 1994. **265**: p. 2096.
47. Gu, Q., et al., *DNA nanowire fabrication*. Nanotechnology, 2006. **17**: pp. R14–R25.
48. Austin, R.H., et al., *Stretch genes*. Physics Today, 1997. **50**(2): pp. 32–38.
49. Mirkin, C.A., et al., *A DNA-based method for rationally assembling nanoparticles into macroscopic materials*. Nature, 1996. **382**: p. 607.
50. Chien, J.H., P.H. Chen, and L.S. Kuo, *Protein detection using a radio frequency biosensor with amplified gold nanoparticles*. Appl. Phys. Lett., 2007. **91**: p.143901.
51. Braun, E., et al., *DNA-templated assembly and electrode attachment of a conducting silver wire*. Nature, 1998. **391**: p. 775.
52. Quake, S.R. and A. Scherer, *From micro- to nanofabrication with soft materials*. Science, 2000. **290**: p. 1536.
53. Xin, H. and A.T. Woolley, *DNA-templated nanotube localization*. J. Am. Chem. Soc., 2003. **125**: pp. 8710–8711.
54. Keren, K., R.S. Berman, and E. Buchstab, *DNA-templated carbon nanotube field-effect transistor*. Science, 2003. **302**: p. 1380.
55. Deng, Z. and C. Mao, *Molecular lithography with DNA nanostructures*. Angew. Chem. Int. Ed., 2004. **43**: pp. 4068–4070.
56. Becerril, H.A. and A.T. Woolley, *DNA shadow nanolithography*. Small, 2007. **3**(9): pp. 1534–1538.
57. Goodman, R.P., et al., *Rapid chiral assembly of rigid DNA building blocks for molecular nanofabrication*. Science, 2005. **310**: p. 1661.
58. Bath, J. and A.J. Turberfield, *DNA nanomachines*. Nat. Nanotechnol., 2007. **2**: p. 275.
59. Iijima, S., *Helical microtubules of graphitic carbon*. Nature, 1991. **354**: pp. 56–58.
60. *Carbon nanotube*, definition provided in *Wikipedia* (web link: http://www.wikipedia.org/).
61. Jung, Y.J., *Controlled synthesis of carbon nanotubes using chemical vapor deposition methods*, in *Nanomanufacturing Handbook*, A. Busnaina, Editor. 2007, CRC Press.
62. Martel, R., et al., *Single- and multi-wall carbon nanotube field-effect transistors*. Appl. Phys. Lett., 1998. **73**(17): p. 2447.
63. Dovidenko, K., et al., *FIB-assisted Pt deposition for carbon nanotube integration and 3-D nanoengineering*. Mat. Res. Soc. Proc., 2003. **H7.7**: p. 739.
64. Jin, A., et al., *Application of dual-beam system in material science and technology*. Chinese Journal of Nonferrous Metals, 2004. **Z2**: p. 56.
65. Liu, J., et al., *Controlled deposition of individual single-walled carbon nanotubes on chemically functionalized templates*. Chem. Phys. Lett., 1999. **303**: pp. 125–129.
66. Valentin, E., et al., *High-density selective placement methods for carbon nanotubes*. Microelectronic Engineering, 2002. **61–62**: pp. 491–496.
67. Kong, J., et al., *Synthesis of individual single walled carbon nanotubes on patterned silicon wafers*. Nature. 1998. **395**: p. 878.
68. Zhang, Y., et al., *Electric-field-directed growth of aligned single-walled carbon nanotubes*. Appl. Phys. Lett., 2001. **79**(19): p. 3155.
69. Tsakalakos, L., *Manufacturing electrical contacts to nanostructures*, in *Nanomanufacturing Handbook*, A. Busnaina, Editor. 2007, CRC Press.

70. Cantoro, M., et al., *Catalytic chemical vapor deposition of single-wall carbon nanotubes at low temperatures*. Nano Lett., 2006. **6**: pp. 1107–1112.
71. Merkulov, V.I., et al., *Patterned growth of individual and multiple vertically aligned carbon nanofibers*. Appl. Phys. Lett., (2000). **76**(24): p. 3555.
72. Teo, K.B.K., et al., *Uniform patterned growth of carbon nanotubes without surface carbon*. Appl. Phys. Lett., 2001. **79**(10): p. 1534.
73. Robertson, J., *Growth of nanotubes for electronics*. Materialstoday, 2007. **10**(1–2): p.36.
74. Ruzette, A.-V. and L. Leibler, *Block copolymers in tomorrow's plastics*. Nature Mater., 2005. **4**: p. 19.
75. Park, C., J. Yoon, and E.L. Thomas, *Enabling nanotechnology with self assembled block copolymer patterns*. Polymer, 2003. **44**: pp. 6725–6760.
76. Xu, H., et al., *Flow-enhanced epitaxial ordering of brush-like macromolecules on graphite*. Langmuir, 2006. **22**(3): pp. 1254–1259.
77. Morkved, T.L., et al., *Local control of microdomain orientation in diblock copolymer thin films with electric fields*. Science, 1996. **273**(5277): pp. 931–933.
78. Hahm, J. and S.J. Sibener, *Cylinder alignment in annular structures of microphase-separated polystyrene-b-poly(methyl methacrylate)*. Langmuir, 2000. **16**(11): pp. 4766–4769.
79. Edwards, E.W., et al., *Long-range order and orientation of cylinder-forming block copolymers on chemically nanopatterned striped surfaces*. Macromolecules, 2006. **39**(10): pp. 3598–3607.
80. Keller, A., E. Pedemonte, and F.M. Willmouth, *Macro-lattice from segregated amorphous phases of a three block copolymer*. Nature, 1970. **225**: p. 538.
81. Ashok, B., M. Muthukumar, and T.P. Russell, *Confined thin film diblock copolymer in the presence of an electric field*. J. Chem. Phys., 2001. **115**: p. 1559.
82. Cheng, J.Y., et al., *Templated self-assembly of block copolymers: effect of substrate topography*. Adv. Mater., 2003. **15**: pp. 1599–1602.
83. Cheng, J.Y., A.M. Mayes, and C.A. Ross, *Nanostructure engineering by templated self-assembly of block copolymers*. Nature Mater., 2004. **3**: p. 823.
84. Stoykovich, M.P., et al., *Directed assembly of block copolymer blends into nonregular device-oriented structures*. Science, 2005. **308**: p. 1442.
85. Harrison, C., et al., *Lithography with a mask of block copolymer microstructures*. J. Vac. Sci. Technol., 1998. **B16**: p. 544.
86. Black, C.T., et al., *Highly porous silicon membrane fabrication using polymer self-assembly*. J. Vac. Sci. Technol., 2006. **B24**: p. 3188.
87. Chai, J., et al., *Assembly of aligned linear metallic patterns on silicon*. Nat. Nanotechnol., 2007. **2**(8): pp. 500–506.
88. Cheng, J.Y., et al., *Fabrication of nanostructures with long-range order using block copolymer lithography*. Appl. Phys. Lett., 2002. **81**(19): pp. 3657–3659.
89. Zhang, G., et al., *Cicada wings: a stamp from nature for nanoimprint lithography*. Small, 2006. **2**(12): pp. 1440–1443.
90. Hitchon, W.N.G., *Plasma Processes for Semiconductor Fabrication*. Cambridge Studies in Semiconductor Physics and Microelectronic Engineering. 1999, Cambridge University Press.
91. Burgos, N., M. Paulis, and M. Montes, *Preparation of Al_2O_3/Al monoliths by anodisation of aluminium as structured catalytic supports*. J. Mater. Chem., 2003. **13**: pp. 1458–1467.
92. Masuda, H. and K. Fukuda, *Ordered metal nanohole arrays made by a two-step replication of honeycomb structures of anodic alumina*. Science, 1995. **268**: p. 1466.
93. Li, Y., et al., *Fabrication of highly ordered nanoporous alumina films by stable high-field anodization*. Nanotechnology, 2006. **17**: pp. 5101–5105.
94. Myung, N.V., et al., *Alumina nanotemplate fabrication on silicon substrate*. Nanotechnology, 2004. **15**: pp. 833–838.
95. Choi, J., et al., *Perfect two-dimensional porous alumina photonic crystals with duplex oxide layers*. J. Appl. Phys., 2003. **94**(8): p. 4757.

96. Huber, C.A., et al., *Nanowire array composites*. Science, 1994. **263**: p. 800.
97. Losic, D., et al., *Fabrication of gold nanorod arrays by templating from porous alumina*. Nanotechnology, 2005. **16**: p. 2275.
98. Gultepe, E., et al., *High-throughput assembly of nanoelements in nanoporous alumina templates*. Appl. Phys. Lett., 2007. **90**: p. 163119.
99. Li, C.-P., et al., *Fabrication and structural characterization of highly ordered sub-100-nm planar magnetic nanodot arrays over 1 cm^2 coverage area*. J. Appl. Phys., 2006. **100**: p. 074318.

Subject Index

A
Aberrations, 85–89
Additive nanofabrication
  Dip-pen nanolithography, 145–147
  field-induced deposition, 143–145
Additive pattern transfer, 212–225, 229
Advancing contact angle, 307
Aerial image, 24, 31, 34, 37, 41, 51
AFM, *see* Atomic force microscope (AFM)
ALD, *see* Atomic layer deposition (ALD)
Alignment, 172–173
Alkaline etching of silicon, 237–238
Alkanethiol, 140
Anisotropic wet etching, 236–238
Anisotropy, 234–236, 242, 244–245, 248, 250, 258, 273
Anodization, 324–326
Antistiction layer, 170
Apertureless glass fiber, 139
Application specific integrated circuits (ASIC), 36, 67
ARDE, *see* Aspect ratio dependent etching (ARDE)
ASIC, *see* Application specific integrated circuits (ASIC)
Aspect ratio dependent etching (ARDE), 254–255
Atomic force microscope (AFM), 129
Atomic layer deposition (ALD), 2
Attenuated PSM, 36–37
AZPN114, 113–114

B
Backscattering, 92, 94, 99, 103, 105, 108, 111, 291
Biasing, 40–41
Bilayer resist system, 47, 151, 175, 220, 222
Binary photomask, 34
Birefringence, 14–15
Block copolymers, 312, 321–324
Bosch process, 250–252
Bottom-up approach, 295–296, 304
Brewster's angle, 26–27

C
Capillary force, 202–206
Carbon dioxide snow, 223
Carbon nanotubes (CNT), 1–2, 295, 312, 316–320
CD-SEM, 52
Chain secession, 107
Charged beams, nanofabrication by, 77–79
  charged particles projection lithography, 120–124
  focusing charged particle beam
    aberrations, 85–89
    charged particle optics, 79–81
    sources, 81–84
  ion sputtering and redeposition, 118–120
  resist materials and processes, 106–107
    contrast of, 112–114
    resolution enhancement processes, 114–118
    sensitivity of, 107–111
  scattering and proximity effect, 89
    electron scattering, 89–94
    ion scattering, 105–106
    low-energy e-beam lithography, 101–105
    proximity effect and correction, 94–99
    secondary electrons, effect of, 99–101
Charged particle optics, 79–81
Charged particles projection lithography, 120–124
Chemical amplification, 44–45
Chemically amplified (CA) resists, 7, 113
Chemical mechanical polishing (CMP), 26, 53, 280

Chemical self-assembly, 297
Chemical vapor deposition (CVD), 145, 213, 229, 269, 297
Chromatic aberration, 85–88, 104
Chromeless phase lithography (CPL), 37–39
CMP, *see* Chemical mechanical polishing (CMP)
CNT, *see* Carbon nanotubes (CNT)
Cold developer, 115–116
Colloidal particles, 295, 305, 309–310
Colloidal self-assembly, 297
Combined nanoimprint and photolithography (CNP), 189–192
Contrast, 24, 26, 32, 37, 44–47, 113–114, 117–118, 190–191, 200, 226, 287, 306–307
  of resist materials, 112–114
Copolymer P(MMA-MAA), 220
Cost of ownership (CoO), 20, 186
Covalent bonds, 298
CPL, *see* Chromeless phase lithography (CPL)
Critical dimension (CD), 25, 48, 51–52, 161
Cross-linking, 7, 107, 307
Cross-over, 82, 87
Cryogenic etching, 249, 252
CVD, *see* Chemical vapor deposition (CVD)

D

Damascene process, 145–147, 229–230, 235, 320
Deep reactive ion etching (DRIE), 250
Deep UV, 7, 14–15, 18, 20, 26, 29, 32, 44, 47, 268, 284
Demolding, 163, 169–172, 177, 181, 184–185, 192–194, 251
Deoxyribonucleic acid (DNA), 313–317
Depth of focus (DOF), 24–27, 32, 50
Descum, 236
Design for manufacturing (DFM), 53–54
DFM, *see* Design for manufacturing (DFM)
Diamond-like carbon (DLC), 169
Diffraction, 9–10
  angle, 10
Digital Micromirror Devices (DMD), 69
Dimensional subtraction and addition
  lateral addition, 278–280
  lateral subtraction, 272–277
  nanosphere lithography, 283–285
  vertical subtraction, 280–283
Dip-pen nanolithography (DPN), 145–147, 235
Distributed Bragg reflectors, 17
DLC, *see* Diamond-like carbon (DLC)
DNA scaffold, 313–316
DNA strand, 314–316
DOF, *see* Depth of focus
Dose correction, principle of, 95
Double processing, 54–57, 66, 286
Double refraction, *see* Birefringence
DPN, *see* Dip-pen nanolithography (DPN)
DRIE, *see* Deep reactive ion etching (DRIE)
Dry etching, 182, 234, 238, 324
Dry ice, *see* Carbon dioxide snow

E

E-beam direct write, 121–122
E-beam projection lithography (EPL), 121–123
EDP, *see* Ethylenediamine Pyrocatechol (EDP)
Electrochemical etching, 147–148
Electroless plating, 226
Electron beam (e-beam) lithography, 129, 161, 307
Electron energy, 108
Electron optics, 77, 79
Electron projection, 27
Electron scattering, 89–94
Electrophoretic deposition, 308
Electroplating, 164–165, 211–212, 226–229, 328
Electrostatic force, 308–310
  assisted self-assembly, 308
Embedded PSM, 36–37
ENFOL, *see* Evanescent near-field optical lithography (ENFOL)
EPL, *see* E-beam projection lithography (EPL)
Etch(ing)
  barrier, 184–185
  parameters
    anisotropy, 234
    selectivity, 234
  resistance, 44, 50–53, 109–110, 136, 176, 178, 238
  selectivity, 50, 136, 181, 183, 185, 244, 249, 260
  See also Isotropic wet etching
Ethylenediamine Pyrocatechol (EDP), 238
Evanescent near-field optical lithography (ENFOL), 57–58
Evanescent wave, 57–58, 61–62, 133, 139
Excimer lasers, 7, 14, 44

Exposure of resists
  field electron emission, 133–135
  by NSOM, 139–140
  by STM, 135–139
Extreme UV (EUV), 16–20
  lithography (EUVL), 3, 16–20, 27, 56–57, 123, 186
  mask, 16, 18
  optics, 17–19
  resists, 18–20
  source, 16–20

F
Faceting, 261
FIB, *see* Focused ion beam (FIB)
Field electron emission, 133–135
Field emission (FE) sources, 82
Field evaporation, 84, 143–145, 148
Field-induced decomposition, 148–149
Field-induced deposition, 143–145
Fluororesin, 179
Focused ion beam (FIB), 18, 77–78, 87–89, 118, 120, 211, 232, 258, 267, 276–277, 279–280, 291–292, 317–318
  trimming, 276
Focusing charged particle beam
  aberrations, 85–89
  charged particle optics, 79–81
  sources, 81–84
Folwer-Nordheim theory, 134
Footing, 257
Forward scattering, 90, 92, 95, 99, 103
Free radicals, 241, 249

G
Gas assisted sputtering, 119
Glass transition temperature (T_g), 163, 166, 169, 174, 194, 213, 322
G-line, 7, 14, 26, 32, 44, 58, 59
Glow region, 239
Gold dot, 144
Grey-scale mask, 39
Guided self-assembly, 296, 304–306

H
Half pitch, 31–32, 47, 55–57, 64–66, 112, 283, 286
Half-tone PSM, 36–37
Halogen compounds, volatility of, 242
Halogen gas, 241–242
Hard masking (HM), 51
Highly oriented pyrolytic graphite (HOPG), 148
High-throughput SPL, 153–156
HM, *see* Hard masking (HM)
Holographic lithography, 63
HOPG, *see* Highly oriented pyrolytic graphite (HOPG)
Hot embossing, 162
HSQ, *see* Hydrogen silsesquioxane (HSQ)
HSQ/PMMA resist, 175–176
Humidity, 141–142, 145–147
Huygen's Principle, 10
Hydrane HS2550, 174
Hydrogen bonds, 313
Hydrogen silsesquioxane (HSQ), 109, 165, 174, 177, 224, 275–276
Hydrophilicity, 306
Hydrophobicity, 153, 306–307

I
IC manufacturing, 2–4, 7–8, 14–15, 19–20, 28–29, 32, 36–37, 42–44, 48, 50, 52, 54–57, 66–68, 70, 112, 121, 123, 154, 156, 161, 177, 184, 186, 229, 234, 238, 268, 286, 320
IIL, *see* Imaging interferometric lithography (IIL)
I-line, 7, 14, 26, 33, 44, 47, 64
ILT, *see* Inverse Lithography Technology (ILT)
Imaging interferometric lithography (IIL), 67
Immersion imaging, 27–28
Immersion optical lithography, 15, 27–29, 186, 286
Imprinting proximity effect, 167
Indirect nanofabrication, 267–268
  dimensional subtraction and addition
    lateral addition, 278–280
    lateral subtraction, 272–277
    nanosphere lithography, 283–285
    vertical subtraction, 280–283
  multistep processing, 285–287
  sidewall lithography, 268–272
  super resolution patterning, 287–292
Inductively coupled plasma (ICP), 249–253
  etcher, 178
  system, 249
Integrated circuit (IC), 2, 4, 7, 25, 53, 78, 161, 226, 229, 279
Interferometric alignment, 188
Interferometric lithography (IL), 8, 63–67, 289

Interferometric optical lithography, 62–66
Intermediate Polymer Stamp, 179–180
International Technology Roadmap for Semiconductors (ITRS), 2, 32, 48, 161
Inverse Lithography Technology (ILT), 42
Ion beam
  milling, 241, 258–262, 292
  sculpting, 279
Ion channeling effect, 258
Ion-induced damage effect, 258
Ion milling, 241, 258–262, 292
Ion optics, 79, 124
Ion projection, 27
Ion projection lithography (IPL), 3, 78, 106, 123
Ion scattering, 105–106
Ion sputtering, 118–120, 220, 238, 241, 244–245, 258–260
  and redeposition, 118–120
Isotropic wet etching, 234–236
K_1 factor, 13, 30–32, 38, 51, 53–56, 227
Knudsen's cosine law, 213, 215, 217

L

Laser direct write, 42, 68
LEEPL, *see* Low-energy e-beam lithography
LER, *see* Line edge roughness (LER)
Lift-off, 135–137, 151, 164, 168, 170, 175, 191, 211–213, 218–220, 222–226, 231, 261, 317, 327
Lift-off resist (LOR), 222
LIGA, 226
Line edge roughness (LER), 18–19, 44, 48–49, 58, 62, 110, 116, 185
Liquid metal ion sources (LMIS), 77, 83, 144
LMIS, *see* Liquid metal ion sources (LMIS)
Loading effect, 254–255
Local oxidation
  lithography, 140–143
  process, 129
Long-throw deposition, 217–218
LOR resist, 222–223
Low-energy e-beam lithography, 101–105, 137
Low-energy e-beam projection lithography (LEEPL), 105

M

MAAL, *see* Micro-aspiration-assisted lithography (MAAL)
Magnetically assisted self-assembly, 310–312
Magnetic force, 310–312
Magnetron, 213, 216–218
MAPPER, 123
Mask(s)
  grey-scale, 39
  hard masking (HM), 51
  optical proximity correction (OPC), 235
  pattern transfer by stencil, 231–233
  phase-shifting masks (PSM), 34–38, 199
  scattering, 121
  stencil, 105–106, 121–124, 231–233, 281
  X-ray, 4, 20–22, 227, 288
  See also specific entries
Mask error enhancement factor (MEEF), 51
Maskless lithography (ML2), 3, 67–68, 123
Maskless optical lithography, 8, 67–72
MBE, *see* Molecular beam epitaxy (MBE)
Mean free path, 217, 244–245
Mechanical scratching, 150–153
  patterning, 151
MEEF, *see* Mask error enhancement factor (MEEF)
MEMS, *see* Micro-electro-mechanical-system (MEMS)
Mercury lamps, 7, 14, 44
Micro-aspiration assisted lithography (MAAL), 204–205
Microcolumn, 104
Microcontact printing (μCP), 166, 197, 199–202, 235, 302, 307
Microdomain, 321–323
Micro-electro-mechanical-system (MEMS), 69, 155, 250
Micro moulding in capillaries (MIMIC), 197
  principle of, 202
Microphase separation, 321–324
Microstereolithography, 71
Microtrenching, 178, 255–256
Microtrenching effect, 255
Micro zone plate (MZP), 227
Millipede, 155–156
MIMIC, *see* Micro moulding in capillaries (MIMIC)
Model-based OPC, 41
Moiré patterns, 188–189
Molecular beam epitaxy (MBE), 2, 282, 297
Molecular combing, 313
Molecular lithography, 315
Molecular recognition, 302–303, 314–315, 317
Molecular self-assembly, 4–5, 295–296, 298, 313
Molecular weight, 109, 117, 166, 169, 171, 182, 220, 302, 322

Monte Carlo simulation, 90, 100, 105, 220, 255
Moore's Law, 2
Multi-beam lithography, 123
Multilayer resist systems, 51–53, 175
Multi-photon lithography (MPL), 71–72
Multiple probes, 153–154
Multistep processing, 285–287

N
Nanocrystals, 300–301
Nanofabrication, 2–4
Nanoimprinting, 3, 5, 27, 161–164
  lithography (NIL), 3, 161–162, 172, 219
Nanoimprint polymers, 166–169
Nanoimprint stamps, 164–166
Nanomanufacturing, 3–4
Nanoparticles, 1–2, 305–306, 308, 310, 312, 321, 323, 326
Nanoscale pattern transfer, 211–212
  additive
    damascene process, 229–230
    pattern transfer by lift-off, 218–225
    pattern transfer by plating, 226–229
    pattern transfer by stencil mask, 231–233
    thin film deposition, 212–218
  subtractive, 234
    anisotropic wet etching, 236–238
    ion milling, 258–262
    isotropic wet etching, 234–236
    process control in nanoscale RIE, 242–249
    RIE, 238–242
    RIE by inductively coupled plasma, 249–253
    RIE, critical issues in, 253–258
Nanoscience, 1, 3–4, 130, 147, 156
Nanosphere lens array, 284
Nanosphere lithography, 283–285
Nanosphere optical lithography, 284
Nanostencil lithography, 231
Nanotechnology, 1–2
Nanowires, 2
Near-field optical lithography (ENFOL), 8, 57–62, 133
Near-field scanning optical microscope (NSOM), 132, 139–140
  principle of, 133
NEB-22, 109, 113, 117, 177–178
Negative tone lift-off, 223–225
"Next-generation lithography" (NGL) technologies, 3, 7, 27, 123, 161, 185
Non-covalent bonds, 297, 304
Notching, 256–257
Notching effect, 256–257
NSCOM, *see* Near-field scanning optical microscope (NSOM)
Nucleotides, 313
Numerical aperture, 11, 123, 129, 227

O
OAI, *see* Off-axis illumination (OAI)
Octadecyltrichlorosilane (OTS), 153
Off-axis illumination (OAI), 33–34, 38
OPC, *see* Optical proximity correction (OPC)
Optical lithography, 2–5
  at high NA, 22–30
  at low k_1 factor, 30–32
    DFM, 53–54
    double processing, 54–57
    OAI, 33–34
    OPC, 38–44
    photoresists, 44–53
    PSM, 34–38
  at shorter wavelengths
    deep UV, 13–15
    extreme UV, 16–20
    X-ray, 20–22
Optical projection lithography, principle of, 8–13
Optical proximity correction (OPC), 38–44, 54, 66, 67, 71, 185, 235
  masks, 235
Optical proximity effect, 39–40, 94
Optical tweezers, 313
Oxidation sharpening, 274–275

P
PAG, *see* Photoacid generator (PAG)
Pattern collapse, 51–53, 197
Pattern density, 108
Pattern transfer, 5, 51, 62, 104, 112, 136, 165–166, 168, 176, 181–182, 184, 202, 211–268, 286–287, 303, 323–324
  by lift-off, 218–225
  by plating, 226–229
  by stencil mask, 231–233
PDMS, *see* Polydimethylsiloxane (PDMS)
Pellicle, 14–15
PGMEA, *see* Propylene glycol mono ether acetate (PGMEA)
Phase-shifting masks (PSM), 34–38, 199, 235
Photoacid generator (PAG), 44, 48–49, 56

Photolithography, *see* Optical lithography
Photomasks, 3, 7, 15, 31, 33, 34, 42, 61, 66–67, 70, 101, 103, 121, 124, 135, 176–178, 191, 236
Photons, nanofabrication by, 7–8
  interferometric optical lithography, 63–67
  maskless optical lithography, 67–72
  near-field optical lithography, 57–62
  optical lithography at high NA, 22–30
  optical lithography at low k_1 factor, 30–32
    design for manufacturing (DFM), 53–54
    double processing, 54–57
    off-axis illumination (OAI), 33–34
    optical proximity correction (OPC), 38–44
    phase-shifting mask (PSM), 34–38
    photoresists, 44–53
  optical lithography at shorter wavelengths
    deep UV, 13–15
    extreme UV, 16–20
    X-ray, 20–22
  optical projection lithography, principle of, 8–13
Photopolymerization, 162, 180, 185
Photoresists, 7, 44
  contrast, 45–48
  etch resistance, 50–53
  LER, 48–49
  sensitivity, 44–45
Physical self-assembly, 297
Physical vapor deposition (PVD), 213
Plasma sputtering deposition, 213, 216–217, 219–220, 229
Point spread function (PSF), 92–93, 99, 101, 105, 110, 116
Polarization, 15, 26–27, 63
Polydimethylsiloxane (PDMS), 178, 204
Polymer stamps, 179
Polymer synthesis, 181
Polymethylmethacrylate (PMMA), 61, 77, 135, 161
Polystyrene (PS), 168, 283–284, 298–299, 308–310, 322–323
Porous alumina, 313, 324–329
Post-exposure baking (PEB), 44–45, 48, 49
Potassium hydroxide (KOH), 238
p-polarization, 26–27
Process conditions, 108
Process control in nanoscale RIE, 242–249
Projection maskless patterning (PMLP) tool, 124
Projection optical lithography, 8–13, 26, 57, 66–68, 187, 268
Projection reduction exposure with variable axis immersion lens (PREVAIL) system, 121
Propylene glycol mono ether acetate (PGMEA), 203
Proximity correction, 37–41, 95–99, 185, 235
Proximity effect, 22, 39–40, 89, 94–95, 97, 99, 103, 105–106, 112–113, 123
  and correction, 94–99
PS, *see* Polystyrene (PS)
PSM, *see* Phase-shifting masks (PSM)
PVD, *see* Physical vapor deposition (PVD)

Q
Quartz stamps, 178, 193

R
Rayleigh Criterion, 30
Reactive-ion etching (RIE), 5, 18, 104, 112, 136–137, 229, 238–242, 268, 306
  critical issues in, 253–258
  by inductively coupled plasma, 249–253
Receding liquid meniscus, 313–314
Redeposition, 118–119, 241, 261–262
Replication, nanofabrication by, 161–162
  reverse nanoimprint, 192–196
  room temperature nanoimprint, 173–176
  soft lithography, 196–197
    by capillary force, 202–206
    microcontact printing, 199–202
    soft stamps, 197–199
  thermal press nanoimprint, 162–164
    alignment, 172–173
    demolding, 169–172
    nanoimprint polymers, 166–169
    nanoimprint stamps, 164–166
  UV-cured nanoimprint, 176
    alignment through transparent stamps, 187–189
    combined nanoimprint and photolithography, 189–192
    step-and-flash imprint lithography, 183–186
    transparent stamps, 176–180
    UV-curable polymers, 180–183
Resist contrast, 108

Resist materials and processes, 106–107
contrast of, 112–114
resolution enhancement processes, 114–118
sensitivity of, 103, 107–111, 114
Resist thickness, 108
Resolution enhancement processes, 114–118
Resolution enhancement techniques (RET), 8, 32–33, 53, 66–69, 235
Reticle, 7, 68
Reverse nanoimprint, 192–196
RIE, *see* Reactive-ion etching (RIE)
RIE lag, 255
RIE trimming, 276
Room temperature nanoimprint, 173–176
lithography (RT-NIL), 173–176
Rule-based OPC, 41

S

Sacrificial layer, 280, 299
SAL601, 113, 136–137, 154–155
SAM, *see* Self-assembled monolayer (SAM)
SAMIM, *see* Solvent-assisted microcontact molding (SAMIM)
Scaffold, 295, 312–316
Scalloping effect, 250–252
SCALPEL, 121
Scanning near-field optical microscope (SNOM), 132
Scanning optical microscope (NSOM), 132
Scanning probe lithography (SPL), 129–130, 143, 147, 151, 153–154, 156
Scanning probe microscope (SPM), 129, 131
principles of, 130–133
Scanning probes, nanofabrication by, 129–130
additive nanofabrication
Dip-pen nanolithography, 145–147
field-induced deposition, 143–145
exposure of resists
field electron emission, 133–135
by NSOM, 139–140
by STM, 135–139
high-throughput SPL, 153–156
local oxidation lithography, 140–143
SPM, principles of, 130–133
subtractive nanofabrication
electrochemical etching, 147–148
field-induced decomposition, 148–149
mechanical scratching, 150–153
thermomechanical indentation, 149–150
Scanning tunneling microscope (STM), 1, 129–132, 135–141, 143–145, 147–148, 153–154, 236
Scattering and proximity effect, 89
electron scattering, 89–94
ion scattering, 105–106
low-energy e-beam lithography, 101–105
proximity effect and correction, 94–99
secondary electrons, effect of, 99–101
Scattering bars, 35, 40–41
Scattering mask, 121
Scattering with angular limitation projection e-beam lithography (SCALPEL) system, 121
Schottky cathode, 83
Secondary electrons, effect of, 99–101
Selectivity, 50, 136, 143, 154, 164, 178, 181–183, 185, 234–235, 238, 240, 242, 244–246, 249–250, 252–253, 258, 260–261, 310
Self-assembled monolayer (SAM), 140, 199, 235–236, 303
Self-assembly, nanofabrication by, 295–296
building blocks of future nanosystems, 312–313
block copolymers, 321–324
CNT, 316–320
DNA scaffold, 313–316
porous alumina, 324–329
guided, 304–305
electrostatic force, 308–310
magnetic force, 310–312
surface topography, 305
surface wetting, 305–308
processes, 296–304
Self-bias field, 244
Self-organization, 295–296
Sensitivity, 18–19, 44–45, 49, 65, 103–106, 113–114, 118, 120, 154, 177–178, 182, 220, 288, 290, 313
resist materials, 107–111
Shadowmask patterning, 231
Shaped e-beam lithography, 121
SHRINC, *see* Super High Resolution by Illumi-Nation Control (SHRINC)
Sidewall image transfer (SIT), 269
Sidewall lithography, 268–272
Silicon micromachining process, 237–238
Silicon nanowire, 248, 272–273
Silicon-on-insulator (SOI), 155, 256, 280
Silicon surface micromachining, 69, 299
Silylation, 270
SLM, *see* Spatial light modulator (SLM)

Soft lithography, 5, 162, 196–197, 199
by capillary force, 202–206
microcontact printing, 199–202
soft stamps, 197–199
Soft stamps, 166, 179, 197–199, 302, 307
Soft X-ray, 16
Solvent-assisted microcontact molding (SAMIM), 197, 203
Sources, 81–84
Space charge effect, 87–88, 104, 123
Spatial frequency, 22–24, 39
Spatial light modulator (SLM), 70
Spherical aberration, 85–87
Spin on glass (SOG), 154, 176
SPM, *see* Scanning probe microscope (SPM)
s-polarization, 26
SSIL, *see* Step-and-stamp imprinting lithography (SSIL)
Stencil mask, 105–106, 121–124, 231–233, 281
Step-and-flash imprint lithography (SFIL), 183–186
Step-and-flash nanoimprinting lithography, 3
Step-and-stamp imprinting lithography (SSIL), 173
Step coverage, 213, 219
Step & stamp imprinting lithography (SSIL), 173
Step & stamp nanoimprinting, 165
STM-induced chemical vapor deposition (CVD), 145
Stoichiometry, 219
Stripe, 68
SU-8, 66, 111–112, 150, 179, 181, 189–190, 193
Sub-100 nm-scale, 3, 236
Sub-resolution assist features (SRAF), 40, 53, 67, 71
Substrate material, 108
Subtractive nanofabrication
electrochemical etching, 147–148
field-induced decomposition, 148–149
mechanical scratching, 150–153
thermomechanical indentation, 149–150
Super High Resolution by Illumi-Nation Control (SHRINC), 33
Superlattice, 312
Superlens, 61–62
Superparamagentic colloidal particles, 310
Super resolution patterning, 287–292
Supramolecular engineering, 303
Surface plasmon polariton (SPP), 59–62
Surface tension, 29, 53, 193, 298–300, 308
Surface topography, 305
Surface wetting, 305–308
Sylgard 184, 178
Synchrotron radiation source (SRS), 16, 22

T
Teflon, 170
Template, 124, 143, 185–186, 296, 305, 310–311, 313, 315, 322–323, 326
Thermal evaporation deposition, 213, 217, 219–220, 231
Thermal oxidation, 270, 274, 280–281
Thermal press nanoimprint, 162–165, 173, 176, 181, 193
alignment, 172–173
alignment accuracy for, limits in, 173
demolding, 169–172
nanoimprint polymers, 166–169
nanoimprint stamps, 164–166
Thermionic cathode, 82, 87, 259
Thermomechanical indentation, 149–150
Thermoplastic polymers, 168–169
Thin film, 2, 15, 17, 37, 61, 94, 109
deposition, 212–218
materials, 236
pattern, 231
Thin-film deposition, 216
TMAH, *see* Trimethyl ammonium hydroxide (TMAH)
Topcoat, 47
Top-down approach, 295
Top-down meets bottomup, 296
Top surface imaging (TSI), 47, 151, 286
Transfer layer, 184
Transparent stamps, 166, 176–180, 184, 187–189
Trilayer stamp, 199
Trimethyl ammonium hydroxide (TMAH), 238
TSI, *see* Top surface imaging (TSI)
T-topping, 45

U
Ultrasonic assisted development, 114–115, 118
UV3, 113–114, 117
UV-curable polymers, 3, 176, 179–184, 189
UV-cured nanoimprint, 176
alignment through transparent stamps, 187–189
combined nanoimprint and photolithography, 189–192

lithography (UV-NIL), 162, 166, 176–179, 181, 184, 186–189, 191
step-and-flash imprint lithography, 183–186
transparent stamps, 176–180
UV-curable polymers, 180–183

V
Vacuum UV (VUV), 7
Van der Waals force, 297, 302, 305, 311, 317
Vibrating (tapping) mode, 131
Virtual source, 82–83, 86
Volatility, 240, 242–243, 245
of halogen compounds, 242

W
Wet chemical etching, 201, 234–236, 238, 242, 272, 274–275, 281–282
Wet imaging, *see* Immersion imaging

X
X-ray lithography, 4, 16, 20–22, 122, 227, 288
X-ray mask, 4, 20–22, 227, 288

Z
ZEP-520, 103, 105, 109, 116–118